REGULATION OF PHAGE LYSIS:

DIVERSE STRATEGIES AND TARGETS

A Dissertation

by

YI CHEN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Ryland F. Young
Committee Members,	Deborah Siegele
	Ping He
	Lanying Zeng
Head of Department,	Gregory D. Reinhart

August 2016

Major Subject: Biochemistry

ABSTRACT

Most of the large, double-stranded DNA (dsDNA) bacteriophages infecting Gram-negative hosts utilize multi-protein strategy to lyse the three layers of the host cell envelop: the inner membrane (IM), the peptidoglycan (PG), and the outer membrane (OM). Four sets of proteins were illustrated in the phage lysis paradigm: holin (IM lesion formation), antiholin (inhibition of holin), endolysin (PG degradation), and spanin complex (OM disruption).

The classic myophage T4 has a very unique feature called Lysis Inhibition (LIN) for regulating the lysis time in response to the superinfection (when a second T4 or T-even phage infects a T4-infected cell). The latest results showed that in addition to the main antiholin RI, which binds to the periplasmic C-terminal domain of T4 holin T, a T4-encoded soluble protein RIII is also required for LIN. Bacterial two-hybrid and Pull-down assays showed that RIII interacts with the cytoplasmic N-terminus of T. Therefore, RIII is considered as a cytoplasmic antiholin of T4.

For another well-studied coliphage Mu, the detail of its lysis system was obscure for decades. The new results suggested that Mu adopts the classic lysis paradigm, consists of the newly-identified holin-antiholin pair Gp19 and Gp20, the SAR-endolysin Lys, and the spanin complex Gp23/23a. Lys has an N-terminal SAR (Signal-Anchor-Release) domain, which makes Lys capable of escaping the IM gradually and spontaneously. However, unlike other previously studied SAR-endolysins, Lys expressed from the plasmid in low level did not lead to cell lysis unless it was co-

expressed with Gp25, the product of Mu gene *25*. The most reasonable explanation is that Gp25 is required for the release of Lys from the IM by neutralizing the basic residues in the cytoplasmic domain of Lys. Therefore, Gp25 is a novel regulatory factor in the phage lysis paradigm that targets SAR-endolysin.

In sum, these studies proposed new models for lysis regulation. In T4, the current model for LIN involves the stabilization of a complex of three proteins in two compartments of the cell. In Mu, the discovery of the Gp25 suggested the involvement of a new class of protein factor that regulates lysis by targeting endolysin.

ACKNOWLEDGEMENTS

I would like to thank my adviser Dr. Ry Young, for his patient guidance, encouragement and support. I was very fortunate to join the Young Lab and I learned more than I have imagined. I want to thank everyone in the Young Lab and the Center for Phage Technology, past and present, for their idea and feedback, especially for helping me shaping up my presentation, and also for being cheerful friends. I would like to thank my committee members, Dr. Siegele, Dr. He, and Dr. Zeng for their constructive comments and advice. I also want to thank two undergraduate researchers, Stan Miller and Jacob Chamblee, for their contribution to our research projects. I thank Ms. Daisy Wilbert for helping me getting things done on time.

I want to thank my colleagues and friends in the Bio/Bio department for their helpful discussion and support, especially members of Herman Lab, Straight Lab, and Zeng lab. A special thanks to Dr. Daniel Shoup and Rye Lab for their help in the protein purification work included in this dissertation.

I cannot express my appreciation enough to my family. I want to thank my parents for their unconditional love over the years. Also, I want to thank my fiancé Ilan, who supported and inspired me during the hardest time of my grad school. I could not have come this far without y'all.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

Page

Page

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

A brief history of bacteriophage study

Bacteriophages, also known as phages, are a kingdom of viruses that use bacteria as their host (1). The first observation of phage infection was reported by E. H. Hankin in 1896, when he noticed the water from Ganga and Yamuna River in India contained some sort of biological materials that can go through millipore filters and can kill the cholera-causing bacteria (2, 3). A few years after Hankin's publication, F. W. Twort found the micrococcus colonies became transparent after prolonged incubation, and this phenomenon seemed contagious since the healthy-looking colonies turned into transparent and watery form after being contacted by "transparent" colonies (1, 2). His discovery was further backed up by an independent researcher, Felix d'Herelle, who became the first scientist to observe "clear zones" form on the lawn of bacteria (*Coccobacilli*), and named these α-microbe "bacteriophage", literary meaning bacteria-eating (1).

D'Herelle's work contributed tremendously to the establishment of the idea "phage therapy" and the foundation for studying the nature of phages, which became one of the most important model organisms in modern molecular biology (1). Methodology and terminology of phage study developed by D'Herelle include the concept of phage "plaque," which means the clear zone formed by phages on the host bacteria lawn; "lysate," which refers to the filtrate of phage-containing culture; "titer,"

which is the estimate of phage concentration in the liquid by counting plaque forming unit (pfu) (4). He also discovered that phage infection requires a stage of virus particles attaching to the surface of the specific bacteria host. This host-specific attachment is now called “phage adsorption”. These preliminary works conducted by d’Herelle, which today we can still benefit from, had greatly inspired the discovery of important biological issues in successive studies.

On the other hand, due to a series of problems phage investigators were facing (1), research related to the therapeutic usage of phage was stagnated. Following the discovery of penicillin in 1928, only a few years after the first phage report, scientists’ interest in phages had quickly shifted from using them as therapeutic reagent against bacteria related diseases to adopting them as a model organism for studying broader biological questions.

Phage life cycle

Just like eukaryote-infecting viruses, bacteriophages are intracellular parasites with protein and/or membrane coats (capsids) packed with their genetic materials (1, 5) . This means they require chemicals, energy, and biological machinery of the host cell to reproduce. The life cycle of phages includes mainly five stages: 1) Adsorption to a new host and penetration of host cell envelope to inject phage genome; 2) Expression of the phage genes; 3) Phage genome replication; 4) Assembly of the progeny virions; 5) Escape from the host cell via cell lysis (Figure 1.1) (6).

The adsorption stage allows phages to transfer their genome (DNA or RNA) into the cytoplasm of the host, where they can utilize the host machinery (5). Normally,

phages can only adsorb into their specific host, by recognizing receptors on the host cell envelop. The form and number of host receptors required for infection depends on the host and the phage. For example, a Gram-negative bacteria infecting phage might require receptors on the outer membrane (OM) (7, 8), and different receptors on the inner membrane (IM) of the cell (9, 10).

There are two pathways after successful adsorption into host and phage genome injection. If the phage genome did not get destroyed by the bacterial defense system, such as restriction digestion or CRISPER-Cas (11), some phages might not choose to lyse the host cell immediately (12) . Instead, they can co-exist with the host by either integrating their genome into the host chromosome and replicate along with it (e.g., λ or Mu), or becoming stable plasmids, which replicate separately from the host genome (e.g. N15 or P1). This pathway is termed as “lysogeny”, and phages that can undergo this pathway are called “lysogenic” or “temperate” (12). Phages that do not have lysogenic stage and thus propagate exclusively by vegetative growth, such as T4 or T7, are called “lytic” or “virulent” phages.

It is important to note that lysogenic phages can undergo the lytic pathway as well. The decision making process between lysis or lysogeny has been described at the molecular level for phage λ in recent studies (13, 14). It is also possible to force lysogenic phages to choose the lytic pathway by mutating critical proteins for lysogeny such as lysogenic repressor or integrase. On the other hand, no virulent phage has been successfully engineered into a lysogenic phage.

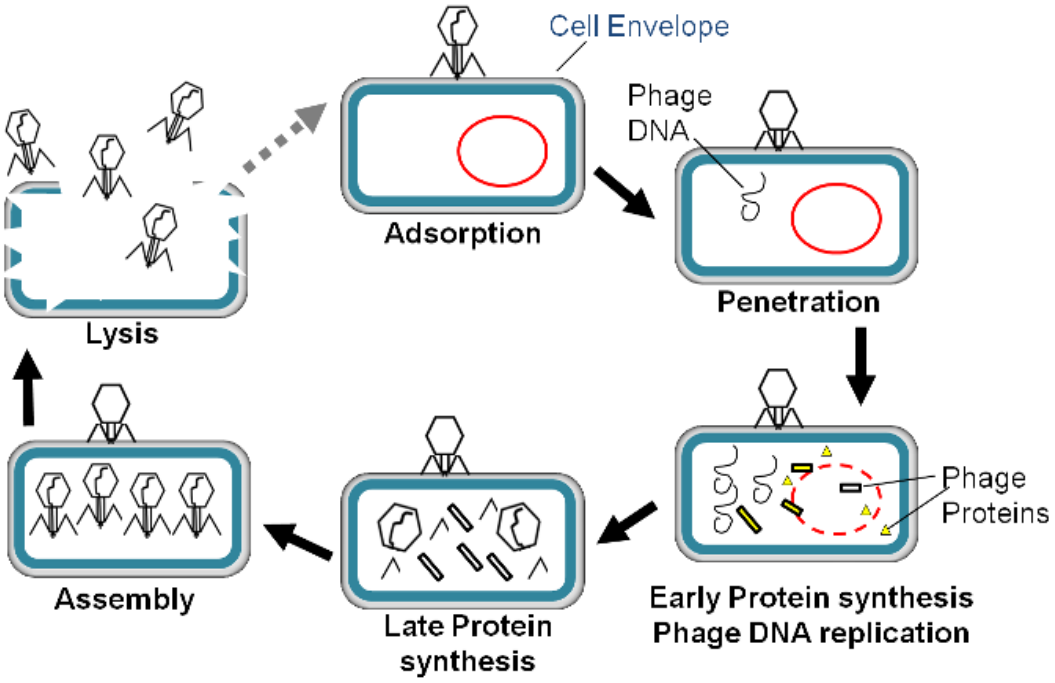

Figure 1.1 Life cycle of virulent phage T4.
The OM or IM of host *E. coli* cell is represented by grey or blue line, respectively.

Phage lysis

For both lytic and lysogenic phages, host lysis is the last event of the phage vegetative cycle, resulting in the release of the progeny particles from the host cells into the environment. Numerous studies have attempted to elucidate the phage lysis process almost since the first discovery of bacteriophage (15). Here we discuss the Gram-negative bacteria infecting double-stranded DNA phages, which are the most well known and well studied phages, to explain their basic strategy for lysing the host cells.

dsDNA phages: lysis paradigm

The cell envelope of Gram-negative bacteria includes three layers: the inner membrane (IM), the peptidoglycan layer (PG), and the outer membrane (OM), which contains lipopolysaccharide (LPS) on its outer surface (16). To break each of these layers, dsDNA phages of Gram-negative bacteria use a multi-gene strategy (17). This involves three classes of proteins: *holin*, the membrane protein which makes lesions ("holes") in the IM; *endolysin*, the phage lysozyme, which enters into the periplasmic space and rapidly degrades peptidoglycan; and *spanin*, which disrupts the OM to complete the lysis (Figure 1.2).

Holin

Holins are small IM proteins that have at least one transmembrane domain (TMD) and have ability to form lesions or "holes" in the IM (17). Some holins make very large holes with a diameters ranging to more than ~300 nm, which is large enough to allow the passage of cytoplasmic macromolecules (Figure 1.3) (18). Other holins can only form small holes

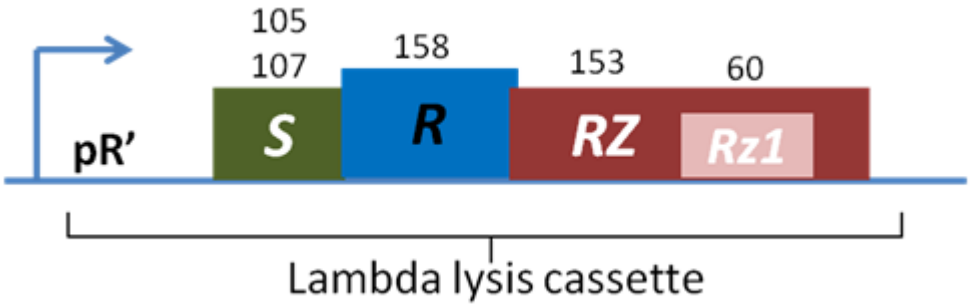

Figure 1.2 The lysis cassette of lysogenic phage λ.
The size of the products of lysis genes (holin S105, antiholin S107, endolysin R, i-spanin Rz, and o-spanin Rz1) are labeled with the number of residues. Adapted from (19).

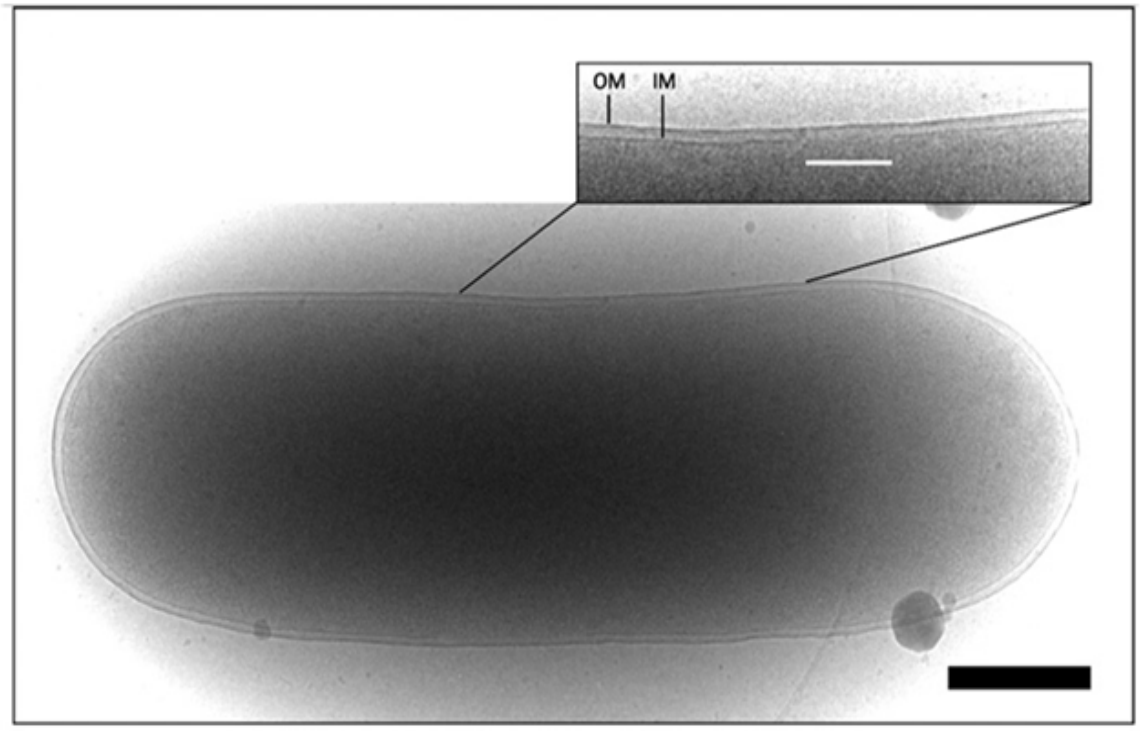

Figure 1.3 Visualization of λ holin membrane leision.
An induced prophage observed by cryo-TEM. Inset: close-up view of the area in which a lesion is observed in the inner membrane. OM: Outer membrane. IM: Inner membrane. The white line indicates the location and extent of the lesion. Scale bar corresponds to 500 nm. Reprinted with permission from (18).

called “pinholes”, with a central lumen of ~2 nm diameter, too small to release macromolecules but capable of depolarizing the IM. Holins that make large holes are called “canonical holins”, while holins that only makes pinholes are termed “pinholins” (20-22). Recent study suggested that these IM holes are lined up by the hydrophilic surfaces of at least one amphipathic TMD from holins (23).

Phage holins which have been studied experimentally to date can be classified into three classes by their topology: the class I holin has three TMDs and adopts N-in, C-out topology; the class II holin has two TMDs and makes pinholes; the class III holin has only one TMD and has N-in, C-out topology. (Figure 1.4) (18, 24).

The best studied holin is that of phage λ, S105 (17). λ has the holin, endolysin, and spanin genes in a small genome section, termed “lysis cassette” (17). In brief, in phage λ infections the holin protein S105 encoded from gene *S* accumulates in the IM and forms holes, releasing the cytoplasmic endolysin R into the periplasmic space, where it attacks and degrades the PG. As the final step, the spanin complex, which consists of Rz and Rz1 proteins, disrupts the OM (19).

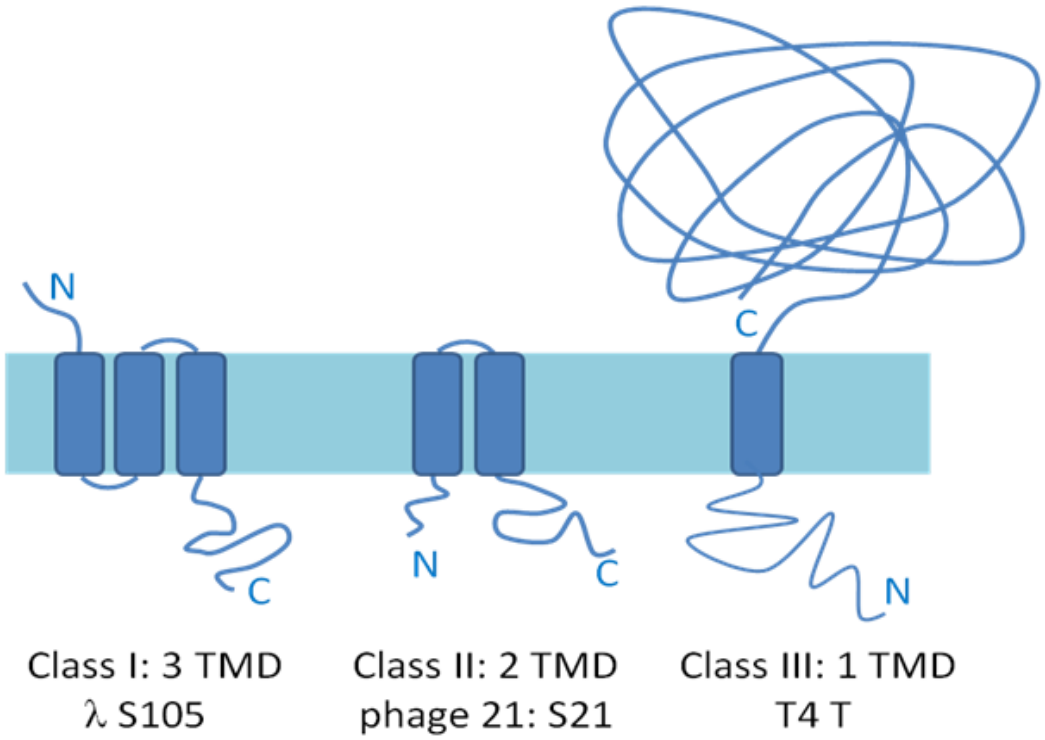

Figure 1.4 Topologies of holins.
Prototype class I holin λ S105, class II holin phage 21 S^{21}, and class III holin T4 T. Rectangle, TMD; Lines, soluble domains and turns.

Despite the existence of three steps, the actual lethal event is the first step mediated by holin (19), which is defined as “holin triggering”. The triggering time is allele-specific. For λ S105, nearly every single amino acid replacement throughout the sequence of the holin altered lysis timing (25, 26). The lesions opened by the holin at the triggering time in the IM will collapse the proton motive force (PMF) and release the cytoplasmic contents into the periplasm (19). This step immediately ends macromolecular synthesis in the cytoplasm and is in effect the termination of the productive infection cycle. Therefore, the holin is the molecule that determines the lysis time and the length of the infection cycle, thus functioniing as the “clock” of the lysis event (25, 27).

Antiholin

As the lysis clock, holin defines the timing of the IM hole formation, thus decidingthe length of the whole infection cycle (25). To adjust the lysis timing, some phages encode an “antiholin”, a protein that interacts directly and specifically with the holin and either blocks or delays the triggering time (19). There are two types of antiholins which have been experimentally studied so far: an alternative translation product from the holin gene itself (28, 29), or the product of distinct gene either in the same lysis cassette (24, 30) or in a distant location of the phage genome (31). Alternatively translated antiholins have a primary structure identical to the with holin except for a few additional amino acids at the N-terminus, like the S107/S105 antiholin/holin pair in λ (Figure 1.5) (19, 28). With this type of antiholin, there is

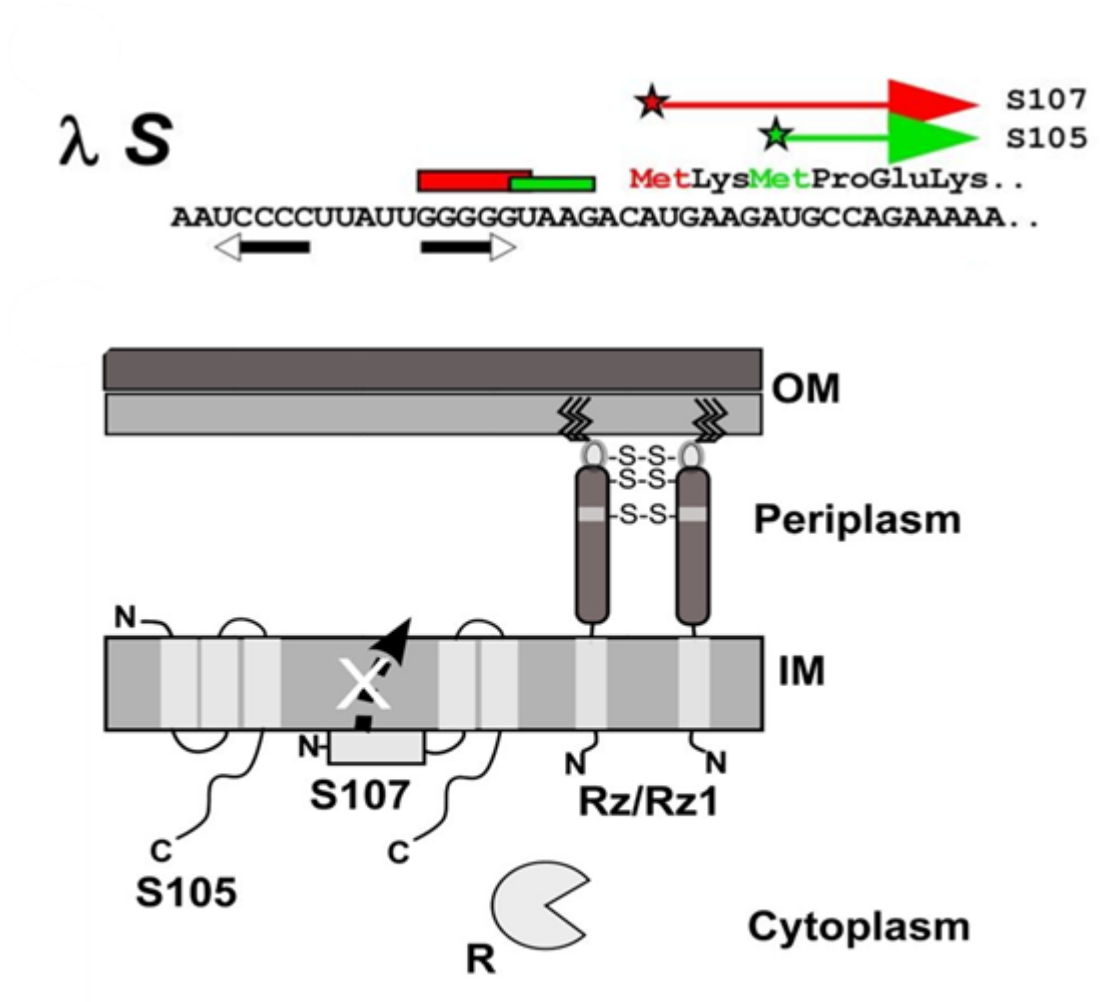

Figure 1.5 The dual start model of holin and antiholin.
Top panel: Dual start motifs for λ holin gene. Start codons for each product indicated by star; red =antiholin, green = holin (S105). Red and green rectangles indicate Shine-Dalgarno sequences. Inverted arrows indicate RNA stem-loops that control choice of start codons. Reprinted with permission from (32).
Bottom panel: Topological dynamics of lysis proteins in λ lysis cassette. In the energized membrane (PMF = ~180 mV, positive outside), TMD1 of the antiholin form (S107) is inhibited from entering (S107) the IM. The R endolysin is shown as a fully active muralytic enzyme. The Rz-Rz1 complex is shown with the Rz i-spanin embedded in the IM with its N-terminal TMD and its periplasmic domain (elongated oval) disposed in the periplasm. Reprinted with permission from (19).

normally a set ratio of holin to antiholin, usually 2:1 to 3:1(28, 33), under standard growth conditions. This system appears to be involved in fine-tuning the triggering time of the holin and avoiding premature lysis, based on the observation that eliminating the translational start of the antiholin in the λ system accelerates timing by only a few minutes. Mutations that increase the ratio of antiholin to holin can dramatically retard the triggering time and even block lysis entirely if an excess of antiholin occurs (28). In paradigm phages P1 and P2, antiholins are encoded by distinct genes in their lysis cassette (24, 30). In these cases, the elimination of the antiholin had only a small effect in terms of accelerating lysis.

The second type of antiholin is constituted by the *r*-proteins of T4 and T4-like phages (34). This type of antiholin binds to and inhibits the holin, but does so only under activating conditions defined by the infection environment (35). This unique "real-time" lysis regulation will be described later in this chapter.

Endolysin

The term "endolysin" was first used by Jacob et al in 1958 to describe "lytic substance" correlated with λ induced bacterial lysis (15), and was later adopted for any phage-encoded enzymes involved in cell lysis. Several distinct enzymes that target either peptide bonds or glycosidic bonds of peptidoglycan layers can be encoded by phage as endolysin (17). For example, the T4 E protein, which sometimes referred as "true phage lysozyme" (17), is a glycosidase that hydrolyzes β(1, 4)-glycosidic bonds (17, 36), while λ endolysin R is a transglycosylase that also breaks the glycosidic bonds but forms a cyclic 1,6-disaccharide product (17). The product of T7 gene *3.5* is an

amidase, which cleaves the amide bond between aminoglycosidic subunits and the tetrapeptide chain (37). The fact that T4 E can complement the defect of λ endolysin indicates that the differences between enzymatic activities of endolysins are unimportant to holin function and the overall cell lysis event (17, 38) (Figure 1.6).

Endolysins can be divided into two classes by the method of getting into the periplasm (38, 39). Cytoplasmic endolysins, such as λ R or T4 E, are trapped inside the cytoplasm until the holins form IM holes that are large enough for them to pass through the membrane bilayer (17). The other type of endolysin, which can escape from the cytoplasm without the assistance from holin, is called a SAR-endolysin (SAR represents Signal-Anchor-Release). With the N-terminal SAR domain, which is a TMD consists of mostly weakly-hydrophobic, polar amino acids residues (e.g. Ala, Gly, Ser, and Thr) (40), this type of endolysin is exported to the periplasm in a *sec* system-dependent manner, resulting in being‘anchored’ to the IM in an enzymatically inactive form . SAR-endo lysin then can be released into the periplasm and be activated by structural rearrangement (Figure 1.7). The detailed molecular mechanism of this spontaneous release remains unclear (39).

Experimentally-studied SAR-endolysins include Lyz from phage P1 (40), R^{21} from phage 21(41), and Lyz^{103} (42)from *Erwinia amylovora* phage ERA103. Based on different strategies of enzymatic activation upon release from the membrane, they can be classified into two classes (38, 39). The first class of SAR-endolysin, which includes Lyz and Lyz^{103}, requires disulfide bond isomerization for

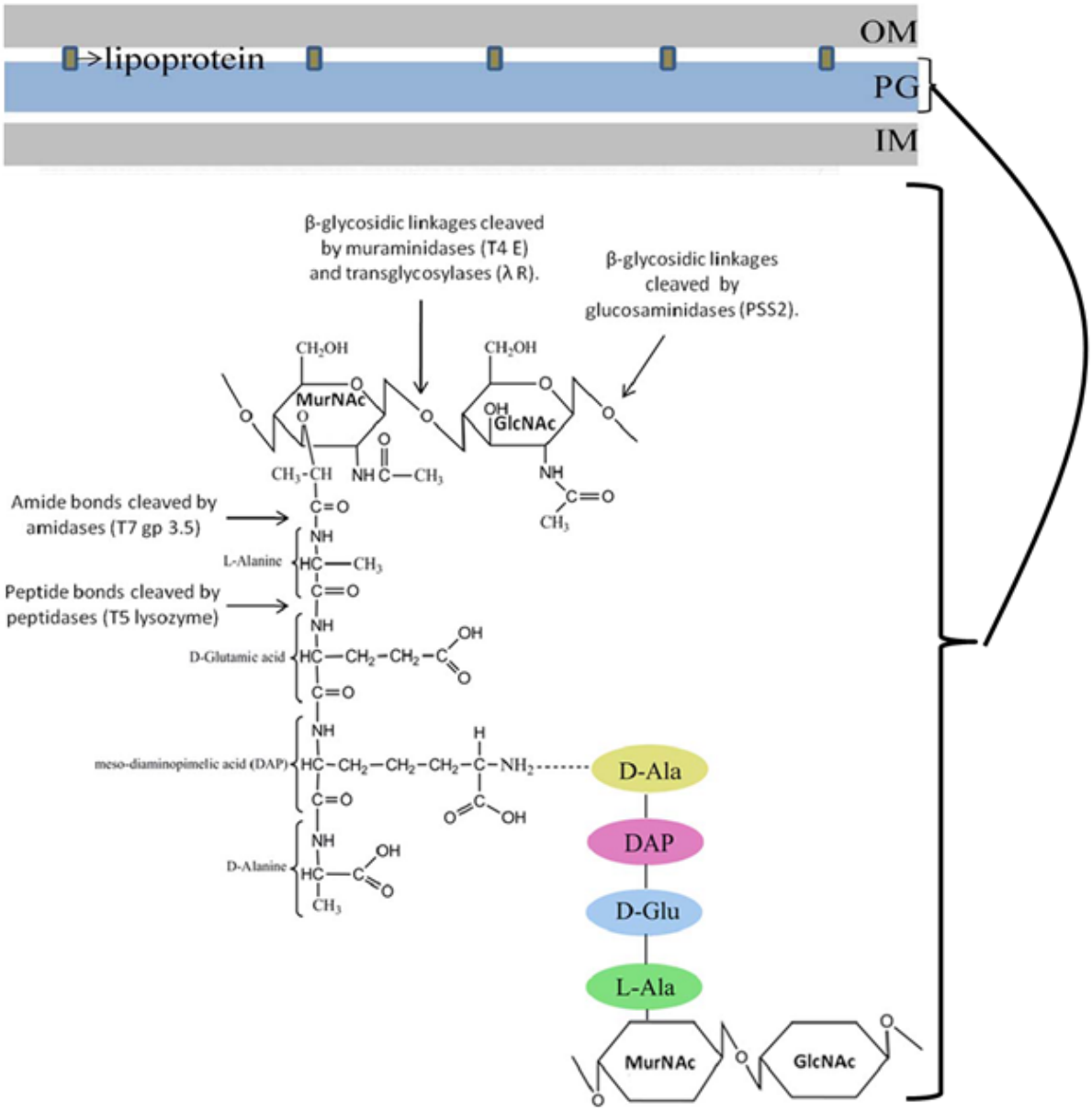

Figure 1.6 The Structure of PG.
The composition of Gram-negative bacteria peptidoglycan and the cleavage sites for different phage endolysins. PG is made up of repeating units of N-acetyl muramic acid (MurNAc)-N-acetyl glucosamine (GlcNAc) disaccharide. The GlcNAc and MurNAc units are connected by a β-glycosidic linkage. Tetrapeptide crosslinks composed of L-Ala-D-Glu-DAP-D-Ala connect the repeating GlcNAc - MurNAc chains. Endolysins cleave various bonds within the PG.OM, outer membrane; PG, peptidoglycan; IM, inner membrane; GlcNac, N-acetylglucosamine; MurNac, N-acetylmuramic acid; m-DAP, meso-diaminopimelic acid; LPP, lipoprotein. Reprinted with permission from (38).

activation (42, 43). In these cases, the SAR-domain is not only required for tethering the protein in the IM but also, upon liberation from the membrane, to contribute a free thiol and thus cause disulfide bond isomerization and refolding of the protein to form an active site cleft (38). In the second class, which is typified by R^{21}, relays on the non-covalent refolding of the SAR-domain after extraction from the membrane, again resulting in the formation of the active site cleft, centered on a catalytic glutamate residue (38). Bioinformatic analysis indicated the R^{21}-like activation mechanism is prevalent in dsDNA SAR-endolysins (38).

Even though SAR-endolysins can lyse the cell without holin function, the induction of SAR-endolysin alone can only cause gradual and significantly delayed cell lysis (22, 40, 43), dependent on the slow release of the SAR endolysins from the bilayer. However, the collapse of membrane potential drastically accelerates the release of SAR-endolysin and leads to a much accelerated lysis process (20, 22). Therefore, even though SAR-endolysin does not rely on the holin-holes to get into the periplasm, the holin still plays a significant role for controlling the lysis time by collapsing the membrane potential and triggering the immediate release and activation of the SAR endolysins (39) (Figure 1.7 and 1.8).

Spanin

The long-established holin-endolysin model of phage lysis held that the destruction of the PG by the phage-encoded lysozyme is necessary and sufficient for the cell lysis (32). In this model, holin only plays a role in assisting the externalization of endolysin, and the disruption of the OM of Gram-negative hosts is not considered as the

critical step for releasing the phage progeny virions. The holin-endolysin model was first challenged by a study showed that *Tn*903 insertion on λ gene *Rz* caused lysis defects in the presence of Mg^{2+} (44). In 1993, a smaller gene was discovered to be located in the +1 register within *Rz* and named *Rz1* (45). A few years later, Zhang *et al* (46) introduced amber nonsense mutations into *Rz* or *Rz1* gene without changing the amino acid sequence of the other gene. This study revealed that without functional gene *Rz or Rz1*, which were found to encode a 153 aa or 60 aa protein, respectively, the infected cells had been transformed in the presence of millimolar scale divalent cations (Mg^{2+}) from rod shape to spherical shape instead of been lysed (46). No phage virion is released from the infected cells in the absence of *Rz/Rz1*. These proteins are named "spanin" to reflect the prediction that the Rz-Rz1 complex could span from IM to OM (47). This study also showed that, in the standard growth condition of *E. coli*, where OM is not stabilized by a high level of divalent cations, the shearing force of shaking conditions (e.g. using of orbital shaker for aeration) in the experiments is sufficient for lysis to be observed without spanins (47).

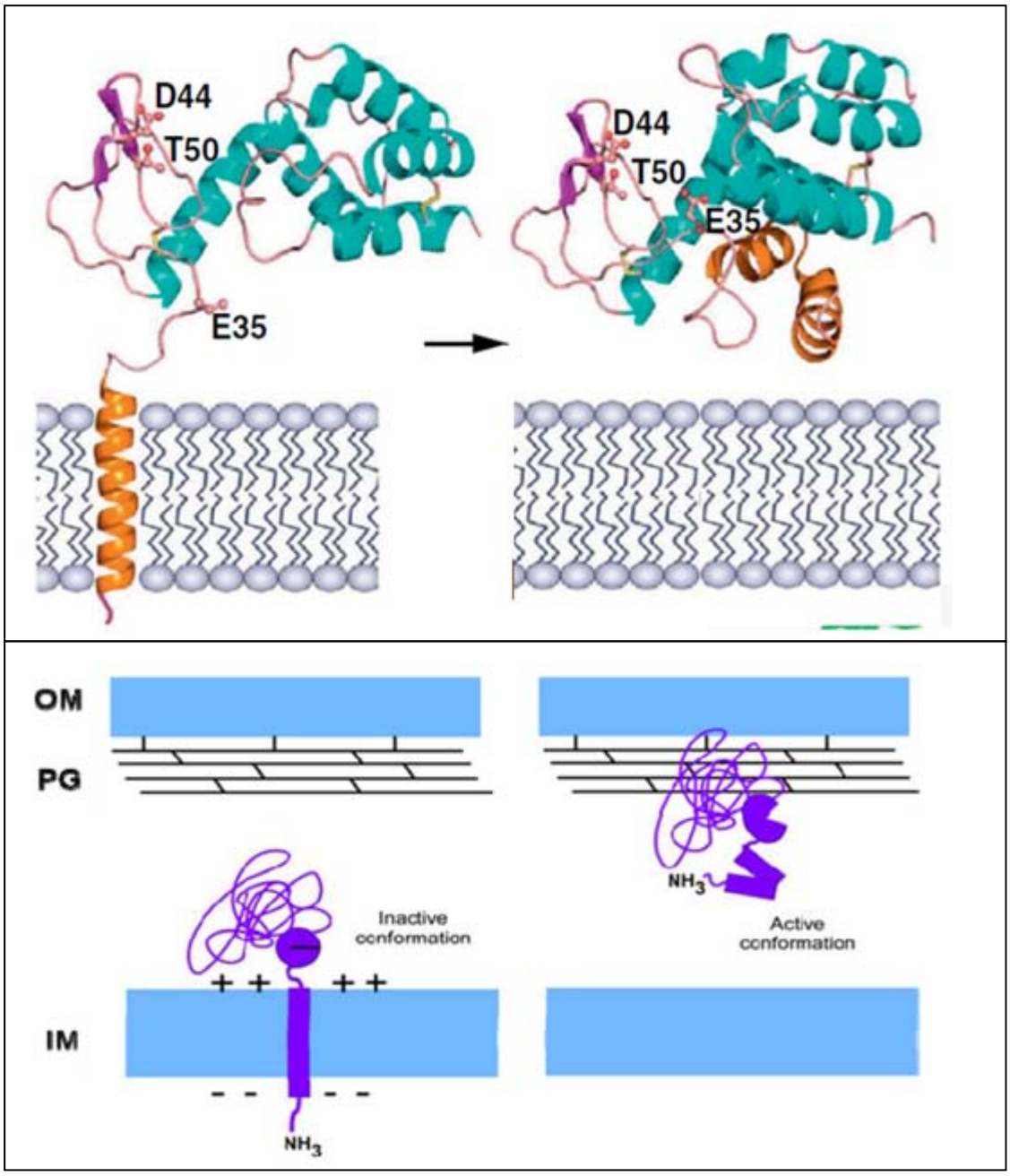

Figure 1.7 Activation of SAR-endolysin.
Top: Crystal structure of phage 21 SAR-endolysin R^{21} in its inactive (in the IM) and active (in the periplasm) form. The catalytic triad (E-D-T) is labeled. Reprinted with permission from (19, 39).
Bottom: The schematic view of SAR-endolysin activation. Reprinted with permission from (38).

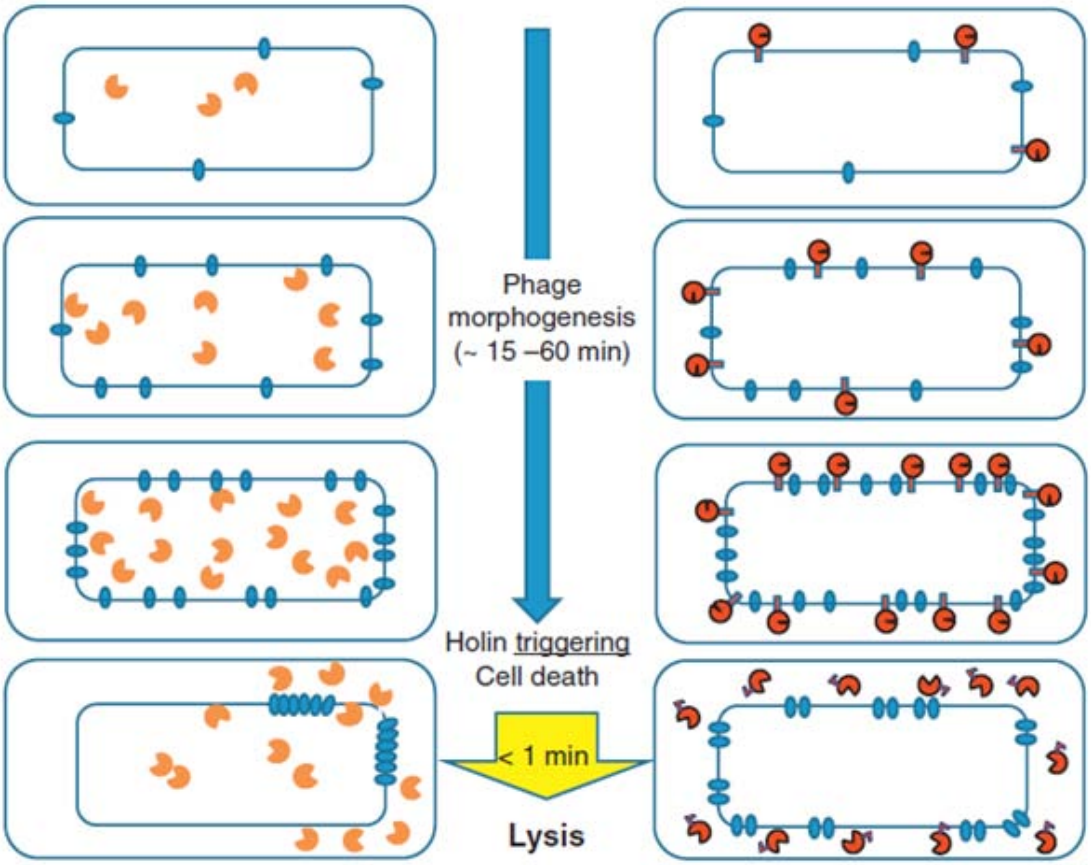

Figure 1.8 Comparison of hoin and pinholin system.
Schematic views of the holin–endolysin (left) and pinholin–SAR endolysin (right) pathways to lysis, beginning at the onset of late gene expression (phage morphogenesis) period. The inner (IM) and outer membranes (OM) of cells are shown, with holin or pinholin (blue ovals) accumulating in the IM. In the bottom figures, upon reaching a critical concentration, the holin triggers to form a large micron-scale hole (left). In contrast, the pinholin triggers to from many heptameric pinholes (represented by double ovals with a channel in between). The orange symbols with the open 'active site' represent the enzymatically active canonical endolysin accumulating in the cytosol. The red symbols with the closed and open 'active sites' represent the inactive SAR endolysin accumulating in the IM and the activated SAR endolysin released into the periplasm, respectively. Reprinted with permission from (19).

Rz/Rz1 were not considered as a crucial factor in phage lysis paradigm until Rz/Rz1 homologs are found to be widespread in Gram-negative bacteria infecting phages (47). Results showing complementation of *Rz./Rz1.* lysis defect by phage P2 genes *lysB/lysC* (48). T4 genes *pseT.3/pseT* (47), and T1 gene 11 (47) suggested that the role of *Rz-Rz1* gene pair is more important than previously thought. Later on, studies conducted by Berry *et al* (49, 50) demonstrated that Rz, an IM protein with an N-terminal TMD, interacts with Rz1, an OM lipoprotein, and their interaction is required for the disruption of the OM. Therefore, Rz and Rz1 homologs are termed i-spanin (inner membrane spanin) or o-spanin (outer membrane spanin), respectively. There are also unimolecular spanin gene products such as T1 *gp11*, which has an N-terminal lipoprotein signal peptide and a C-terminal TMD, is named u-spanin (for unimolecular -spanin). More recent reports on spanin function showed that regardless of the spanin type, disruption of OM was achieved by the fusion of the IM and OM (51) (Figure 1.9).

Three-step lysis of dsDNA phages

For dsDNA phages infecting Gram-negative bacteria, a three-step lysis model, which involves multiple phage proteins as factors, was suggested by the most up-to-date results (32). First, holins form holes in the inner membrane, which control the release of endolysins into the periplasm by collapsing the membrane potential and/or providing the channel in the IM. Second, endolysins attack and destroy the peptidoglycan. Finally, spanins fuse IM and OM, removing the last topological barrier to lysis.

This model also suggests that the lytic event is blocked until holin triggering, to prevent the premature lysis, which significantly affects the robustness of the progeny virions. Instead of relying on the accumulation of a sole muralytic enzyme for lysis, this three step controlling system can adjust the lysis timing for maximizing the yield of the progeny virus particles without losing the competition to other phages that infecting the similar host. The three-step lysis model reflects how lysis proteins cooperate to achieve an efficient yet carefully controlled system to ensure rapid destruction of each layer of the cellular envelope when triggered, and also prevents the premature lysis (32).

T4 lysis and lysis inhibition

T4 is a classic coliphage belonging to the *Myoviridae* family. As a dsDNA phage, it also requires holin (T, the product of gene *t*), cytoplasmic endolysin (E, product of *e*), and spanin (PseT.3 - PseT.2 complex, products of *pseT.3* and *pseT.2*) for the lysis, but unlike λ-like phages, where the lysis genes are clustered together and under the control of a single late promoter, T4 lysis proteins are unlinked in the 170 kb phage genome (Figure 1.10). T4 holin T makes big holes (~300nm in diameter) which let soluble endolysin E escape into the periplasm through the membrane.

In a normal T4 infection cycle, the T4 holin T triggers at ~25 min, and lysis ensues rapidly, liberating an average burst of ~200 progeny virions. If, however, an infected cell undergoes superinfection by another T4 virion after the first 5 min of the infection cycle, the infection enters into the lysis inhibited (LIN) state, in which the T holin is blocked from triggering (Figure 1.11). LIN was first described by Doermann in 1948 (52) when he observed the different lysis profiles between the wild type T4 and

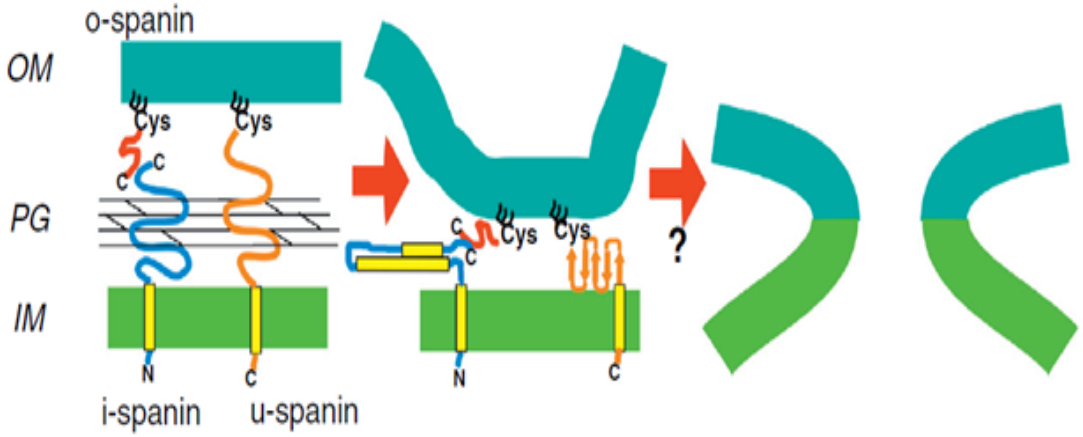

Figure 1.9 Model for spanin structure and function.
Two component spanin complexes consist of an integral IM subunit, the i-spanin (Rz) and an OM lipoprotein, the lipoprotein o-spanin (Rz1), that interact via a C-terminal/C-terminal interaction, spanning the periplasm through the meshwork of the peptidoglycan (PG). Alternatively, a unimolecular spanin, u-spanin (T1 gp11), which is an OM lipoprotein but also has a C-terminal TMD, spans the periplasm as a covalent chain. The Rz/Rz1 complex is shown as a heterodimer for simplicity; in vivo, each subunit is actually a disulfide-linked homodimer. Reprinted with permission from (32).

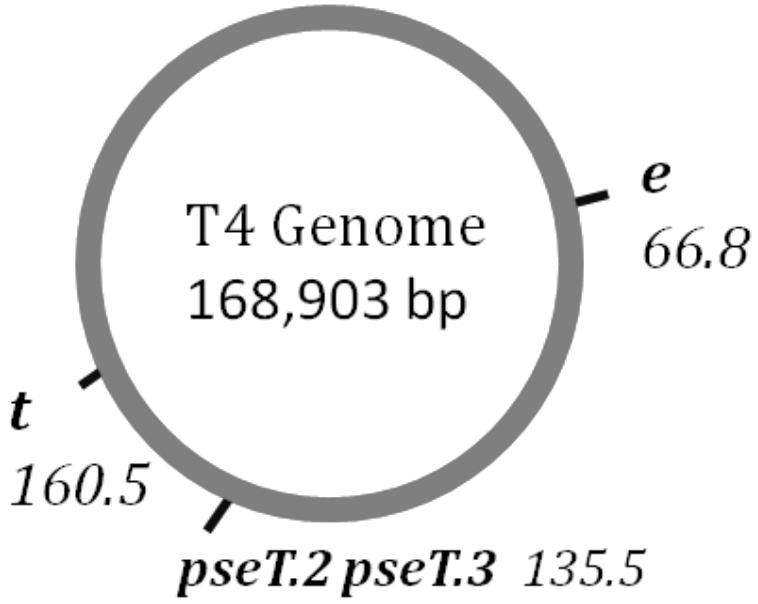

Figure 1.10 Location of lysis genes on the T4 genome.

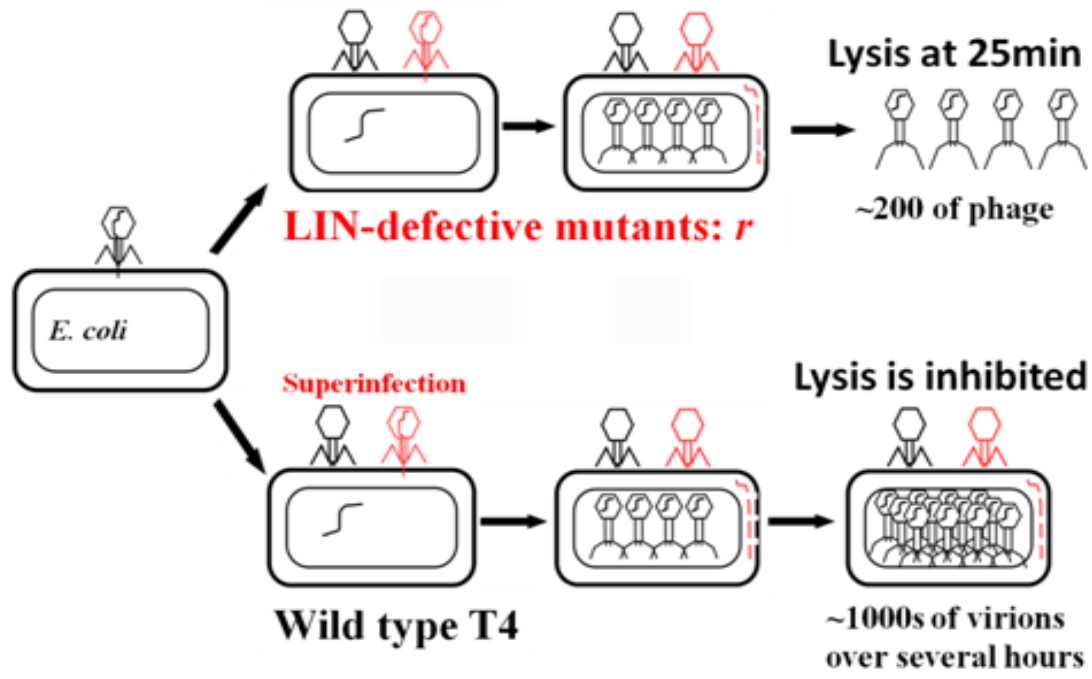

Figure 1.11 Schematic view of T4 superinfection and Lysis Inhibition (LIN). The lysis curves show representative LIN defective mutants (top) and T4wt (bottom) lysis profiles (53).

T4 lysis mutant (***r*** for rapid lysis) infections. While *r* mutants completed lysis within ~30 min after infection, the cell cultures infected with wild type T4 did not lyse for hours. This is because a small fraction of the infected bacterial cells lysed a few minutes earlier than others and released ~200 virions per cell. These newly released phage virions then superinfect the rest of the infected culture and send an undefined LIN "signal" that imposes the LIN state on each superinfected cell. The LIN state can be maintained for hours, allowing phage virion assembly to continue in the infected cells and accumulate more than 1000 virions per cell.

T4 holin

T4 holin, which has 218 amino acids, is the largest holin among all the holins that have been studied in detail. The name *t* was inspired by the tragic story of Tithonus in Greek mythology, where Tithonus was granted eternal life but only as an old man (54). The topology of T is very different from that of λ holin (three transmembrane domains (TMDs), class I holin) or phage 21 holin (two TMDs, class II holin), and thus has been defined as the prototype class III holin (Figure 1.4). T consists of a single TMD, a large, globular shaped C-terminal periplasmic domain (162 aa), and a N-terminal cytoplasmic domain (34 aa) (Figure 1.12). Since T has only one TMD, it is reasonable to presume that this TMD participates in lining the edges of the ultimate membrane lesion formed after triggering. However, both the N-terminal (cytoplasmic) and C-terminal (periplasmic) domains of T are required for holin function (35). The periplasmic domain of T (sT) was co-purified in a complex with the periplasmic domain of the T4 main antiholin, RI (sRI) (53, 55). The crystal structure of sRI-sT

complex was solved in 2011 by Vladimir Kuznetsov from the laboratory of James Sacchettini at Texas A&M university (55), which showed that sT is in a globular shape with a single alpha-helix at the N-terminus (close to the TMD) and five α-parallel beta-sheets connected with loops and turns (53). (Figure 1.13)

T4 holin genetics

Missense mutations conferring defective lysis or altered lysis timing are isolated in all 3 domains of T (56) (Figure 1.14). The cytoplasmic N-terminal domain (NTD) consists of 34 residues, which is by far the largest cytoplasmic domain among all the experimentally confirmed holins. The deletion of the first 28 aa of T results in the loss of holin function, and four non-functional alleles have been isolated with single amino acid change in N-terminus. These residues map to a predicted amphipathic α-helix (Figure 1.15) (56).

In CTD, 13 lysis-defective single missense mutations were found within the 161 residues, as described by Moussa *et al* in 2014. These non-functional alleles include both dramatic (charge or polarity change) and conservative amino acid changes. They are spread widely across the entire domain from residue 63 to the very distal C-terminal residue 213. In addition to these alleles, conservative changes of an essential periplasmic cysteine pair (C175S, C207S, C175S/C207S) result in absolute lysis defect, which indicates that intramolecular disulfide bond formation is required for the proper folding and holin function of T (56).

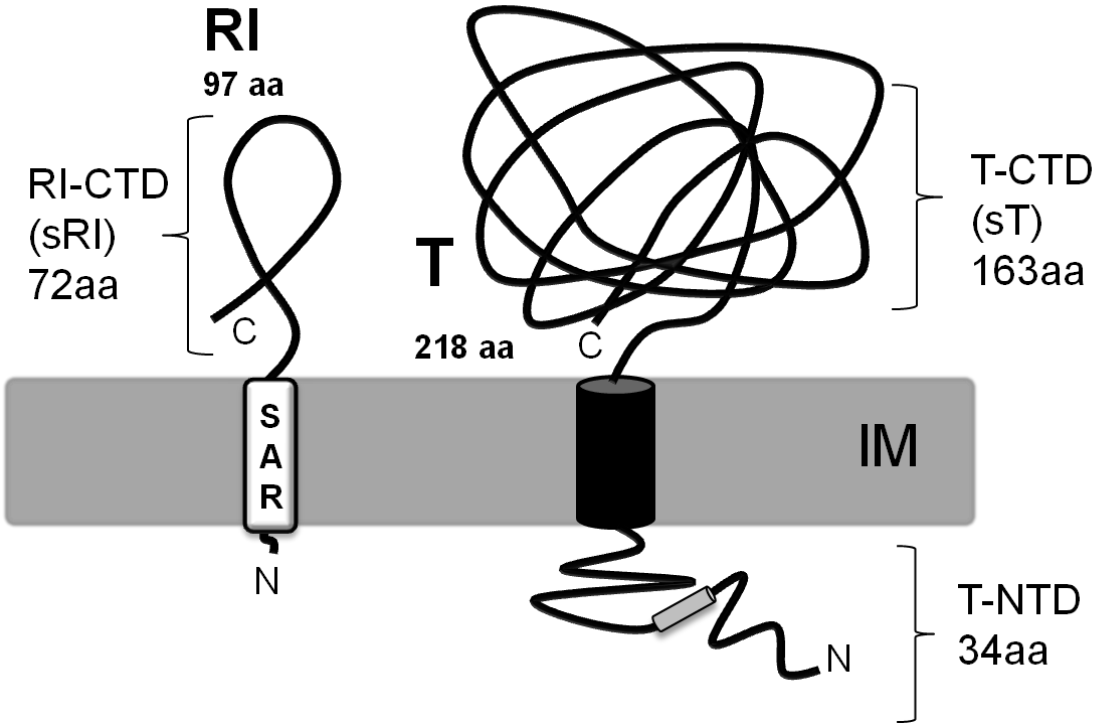

Figure 1.12 Topology of T4 holin T -antiholin RI interaction
T is an inner membrane protein with a single TMD (shown as a black solid cylinder) and an amphipathic helix (shown as a grey cylinder). RI has a SAR (Signal Anchor-Release) domain (shown as a white rectangle) which allows RI to be spontaneously released in to the periplasm. IM, inner membrane (57).

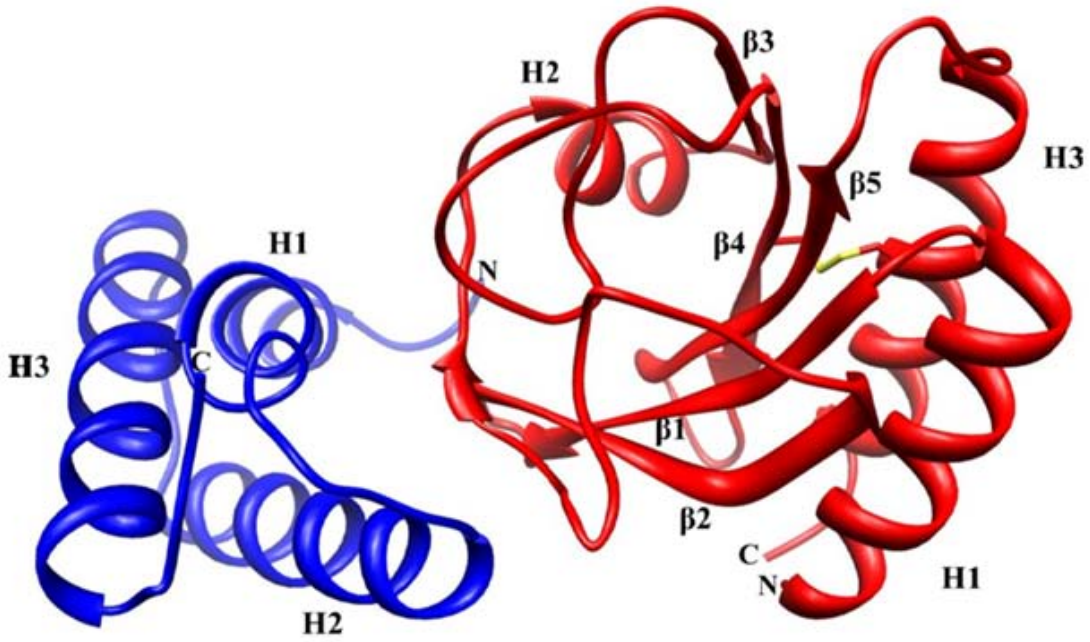

Figure 1.13 Structure of RI-T periplasmic domains complex during LIN
The structure was solved by Vladimir Kuznetsov from the laboratory of James Sacchettini at Texas A&M university (55). The periplasmic domain of T (red) binds directly with the periplasmic domain of RI (blue). The secondary structure elements are labeled for each domain (55, 57). Reprinted with permission from (53).

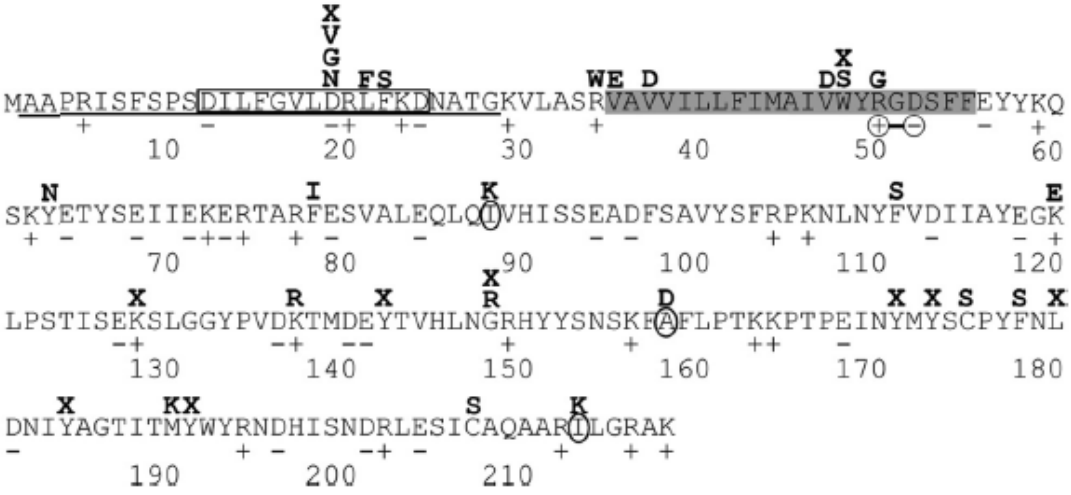

Figure 1.14 Primary structure of the T4 holin.
Lysis-defective missense alleles are indicated in bold. X indicates positions of stop codon mutations that inactivate the protein. The N-terminal domain is underlined and the TMD is highlighted in grey. Reprinted with permission from (57).

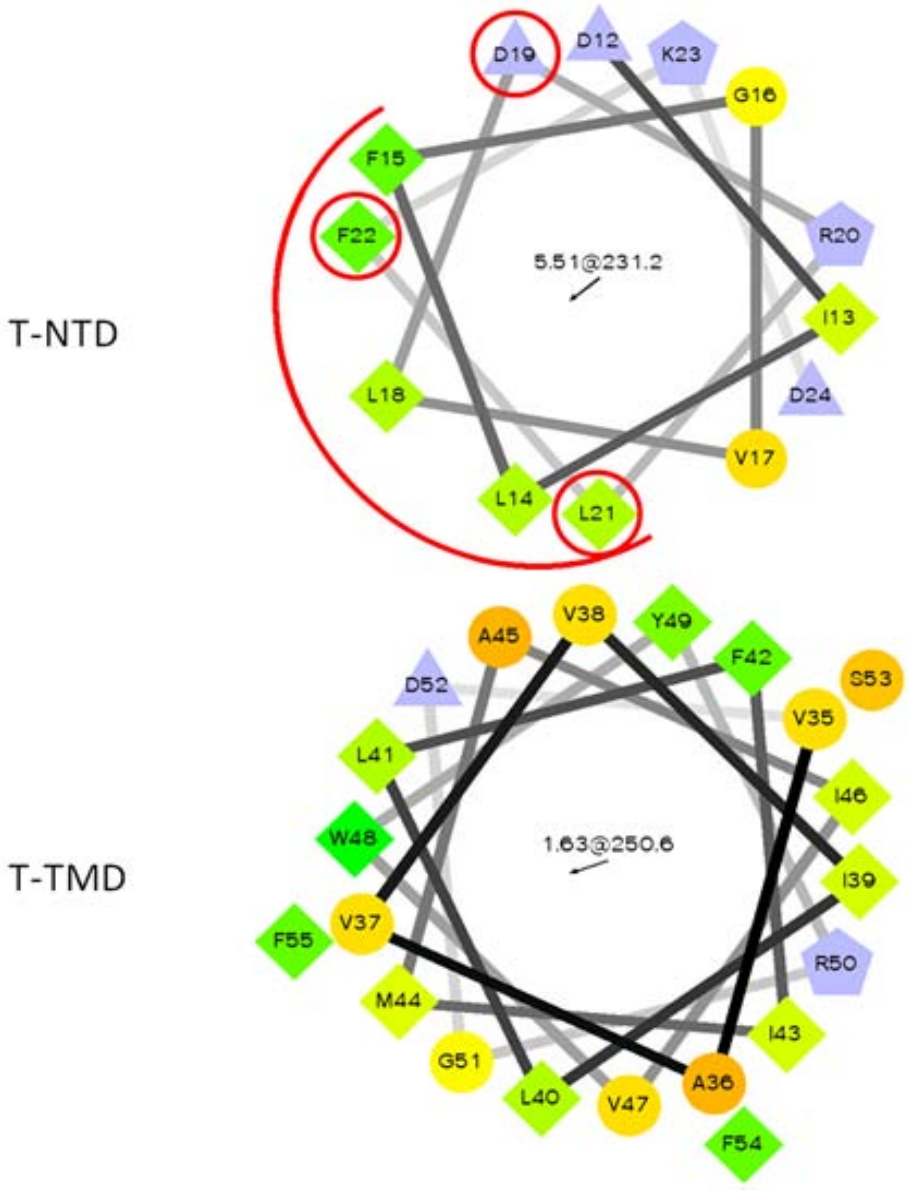

Figure 1.15 Helical wheel projection of T.
Hydrophilic residues as circles, hydrophobic residues as diamonds, potentially negatively charged as triangles, and potentially positively charged as pentagons. Hydrophobicity is color coded as well: the most hydrophobic residue is green, and the amount of green is decreasing proportionally to the hydrophobicity, with zero hydrophobicity coded as yellow. Hydrophilic residues are coded red with pure red being the most hydrophilic (uncharged) residue, and the amount of red decreasing proportionally to the hydrophilicity. The potentially charged residues are light blue. The number shows the hydrophobic moment. The red curve indicates the predicted hydrophobic surface. Reprinted with permission from (56, 57).

Two aspects of the T-TMD are unusual. First, two studies have shown that in the holes formed by the well-studied S105 and $S^{21}68$ holins, the aqueous lumen of the holes are lined by the most hydrophilic surfaces of the holin TMDs. However, the TMD of T has no significant hydrophilic surface (Figure 1.15) (56). Second, the presence of only a single TMD in a canonical holin is unusual. Most holins contain more than one TMD and the only holin that forms its lethal hole with a single TMD is the $S^{21}68$, the pinholin of lambdoid phage 21, which forms small heptameric holes with lumenal diameters of ~ 2 nm (21, 29, 58). However, as shown in Figure 1.7, $S^{21}68$ is synthesized as an IM protein contains two TMD at first, then activated by the release of its N-terminal TMD, which is a SAR-domain (22). In contrast, the S105 holin and all other holins have two or more TMDs and make large membrane holes with diameters of micron scale. This has led to the thought that to form a large membrane lesion, two or more TMDs may be required. We hypothesize that in the mature T hole, the predicted helix in the NTD of T, which includes a predicted amphipathic α-helix (56), may be directly involved in the hole formation and become a part of the hole wall (Figure 1.15). However, no molecular evidence has been found to support this model to date.

In the TMD, there are 11 missense mutations have been identified that altered the T lysis profile (34, 59, 60). However, only the mutations that drastically change the charge of the residues (V37E, V39D, V47D, and R50G), or significantly change the hydrophobicity of the TMD (W48S) debilitate the holin function (Figure 1.14). Unlike other two classes of holins, of which conservative amino acid changes in TMD can lead to the lysis-defective phenotype (56), all T-TMD mutations with conservative changes

retain the lytic function with delayed lysis time. This may suggest the TMD of T is relatively less important for the hole formation compared to other two domains: the N-terminal and C-terminal soluble domains.

LIN and T4 superinfection exclusion

As described above, the LIN state is imposed when a T4 infected *E. coli* was super-infected by another T-even phage (17). Instead of passing into the host cytoplasm, the DNA of the secondary infecting phage is ectopically ejected into the periplasm, where it is degraded by *E. coli* endonuclease I (61). The exclusion of the secondary infection involves two T4 genes: *sp* and *imm*. T4 *sp* (for "Spackle") mutants were originally isolated as the suppressor of T4 endolysin-defective mutants (62) . Later, it turned out that *sp* encodes a 97 aa periplasmic protein which provides the defense against the disrupting of host cell wall by T4 tail baseplate lysozyme, gp5. The product of *imm* gene is an 83 aa membrane protein (63), and is essential for the exclusion of secondary phage DNA. However, the mechanism of *sp-imm* T4 superinfection exclusion remains unknown.

The superinfection could be an indicator of the shortage of uninfected host cells in the environment. During the LIN state, superinfected host cells accumulate progeny virions intracellularly for hours (17), which is beneficial for phages in an environment without available hosts. This makes T4 LIN a unique example of phages adjusting their lysis timing in response to the environmental condition (64).

r *gene history and impact*

It is known that LIN has consequences for T4 plaque morphology. Because there is a high phage content in the developing plaque, cells around the edges of the growing plaque are presumably always being super-infected and are thus in the LIN state (17). As a result, wild type T4 forms small and fuzzy-edged plaques. Mutants incapable of LIN form larger, sharply-defined clear plaques that are expected from phages that do not have a lysogenic cycle (Figure 1.16). These mutants were termed ***r*** mutants, for "rapid lysis" (64), although this is actually the normal, non-inhibited lysis. Because of the dramatic differences in plaque morphology, it was easy to isolate T4 mutants defective in LIN. Beginning with Hershey's work in 1946 (65, 66), seven different genetic loci giving *r* plaques were identified: *rI, rIIA/rIIB, rIII, rIV, rV,* and *rVI* (64, 67). When the T4 genome was sequenced, it turned out that, besides *rI* and *rIII*, two of these loci corresponded to genes with known roles in lysis or LIN: *rIV* is *sp* (68) and *rV* is the T4 holin gene *t* (69). The *rVI* locus is lost to history and the most recent r-plaque screen did not identify any mutations correspond to *rVI* (67) .The *rII* genes (*rIIA* and *rIIB*) are only required for T4 plating on λ lysogens of *E. coli* B strains, and appear to be involved in preventing abortive infections (64) and fundamentally unrelated to the lysis regulation pathway

r *genes directly involved in the regulation of holin mediated lysis*

Gene *rI*

Previous studies from the Young laboratory showed that RI, the product of *rI* gene, is a 97 aa membrane protein with one TMD (Figure 1.17) (34, 70). However, the

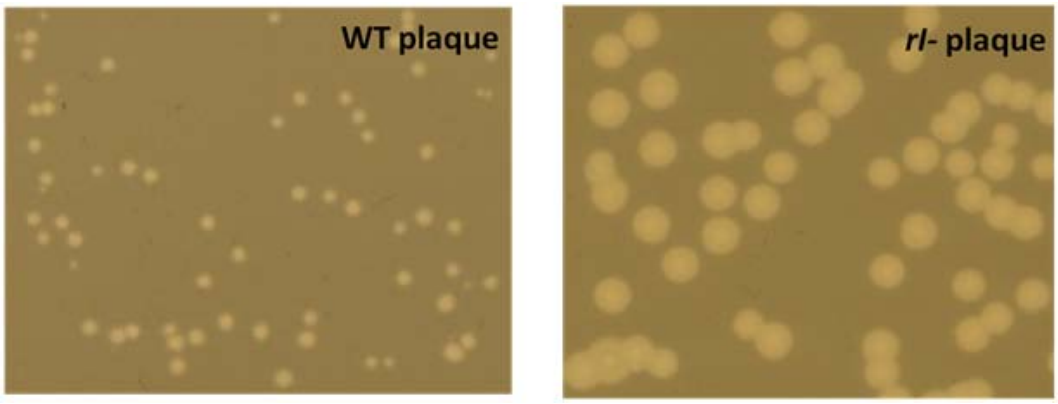

Figure 1.16 Plaque morphology of T4.
Representative picture of T4D (left) and T4*rI*- mutants (right) on the lawn of *E. coli* B834.

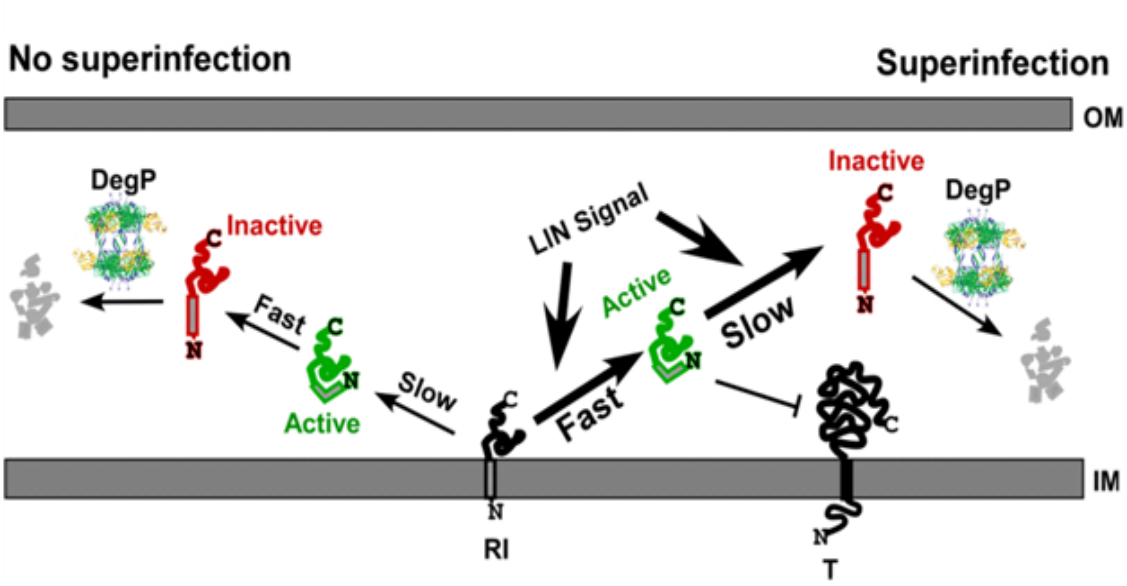

Figure 1.17 Schematic model for T4 lysis inhibition involving antiholin RI.
RI, initially tethered to the inner membrane (IM), is spontaneously released into the periplasm, where it is rapidly inactivated both functionally and proteolytically,when no superinfection/LIN signal exists. Also, the membrane-tethered form of RI is not competent for inhibiting the holin. When superinfection occurs and a LIN signal exists in the periplasmic space, the released RI is stabilized, and is active and competent for LIN. Reprinted with permission from (57).

RI TMD is a Signal Anchor Release (SAR) domain, which is rich in weakly hydrophobic residues, causing it to be gradually and spontaneously released from IM into the periplasm (35). Normally, RI is quickly inactivated and subsequently degraded by DegP, an *E.coli* periplasmic protease (35). Therefore it has an average half-life of ~90 sec *in vivo*. If the SAR domain is replaced by the cleavable secretion signal of the periplasmic protein PhoA, the processed RI protein is stable and accumulates in the periplasmic space (35). Induction of the gene encoding of this stable version of RI, $phoA_{ss\text{-}RI}$, results in stable LIN independent of a superinfection signal and also independent of *rIII*. Therefore, SAR domain is considered to be responsible for the instability. These results suggest that the "signal" from the superinfecting virion acts by stabilizing RI functionally, allowing it to bind to T and block its triggering (35, 70). The periplasmic domain of RI, called sRI, and periplasmic domain of T, called sT, have been successfully overexpressed in an *E. coli* strain which is especially engineered to allow the cytoplasmic formation of disulfide bonds (Figure 1.13) (57). The sRI protein has been purified and characterized by Samir Moussa (57), and crystallized in 2011 by Vladimir Kuznetosov (53, 55) . sRI, which forms a dimer when purified alone, has three helices (H1, H2 and H3) connected by short flexible loops. The structure of H2 and H3 is fixed by a crucial disulfide bond formed between C69 and C75, while the loop between H1 and H2 undergoes drastic conformational change when sRI binds to sT (Figure 1.13) (53, 55). sT was found to oligomerize and precipitate when eluted from affinity columns. Therefore sT was purified in a soluble 1:1 complex with sRI and co-crystallized with sRI. High-resolution structures of sRI and sRI:sT complex solved by

Vladimir Kuznetsov (55) backed up a model suggests that T-RI-RI-T tetramers are formed during the LIN state (53).

Gene *rV*

Lytic but LIN-insensitive T holins were originally isolated as *rV* mutants (*18*). Five *rV* mutants, including three classic *alleles*, have been mapped and sequenced to date. As summarized by Burch *et al*, those mutations were located in three different topological domains of T: R5K in N-terminal cytoplasmic domain, I39V in TMD, and T75I, S80P, and V81G in the CTD (67). These T mutants all trigger at an earlier time than the wild type T. In the condition where RI is expressed from a medium-copy number plasmid (presumably in a higher expression level than physiological level), these *rV* mutants are sensitive to RI-mediated LIN, but the length of the inhibition is significantly shorter than WT (15). Other LIN-defective T mutants associated with significantly late-lysis phenotype all map to the CTD (59, 60), but the detailed profile of these mutants was not defined.

Gene *rIII*

Compared to other *r* genes, the role of *rIII* gene in LIN has been ignored, even though T4*rIII* mutants were first isolated in the 1950s and the gene was mapped and sequenced in the early 1990s. The major reason is that, as also found for the *rIIA/rIIB* genes, *rIII* was thought to be required for LIN only in certain *E. coli* B strains, but not for *E. coli* K-12, the most widely used laboratory strain. Therefore the participation of *rIII* is not considered as an important part of LIN. However, the most recent report about T4*rIII* phenotype suggests that mutations in *rIII* gene lead to LIN-defective

phenotype in all host backgrounds (67). The result of BLAST search and data analysis showed that, with few exceptions, the phages that have RIII homologs also have T homologs and RI homologs. These new data suggested that *rIII* may play an important role for LIN.

rIII was considered to be controlled by early, middle and late promoters (31). However, microarray analysis showed that it is transcribed between 5 min to 8 min after infection (71). The *rIII* gene encodes a 82 aa protein which has been predicted to reside in the cytoplasm since it has no TMD or secretory signal sequence. The comparison of RIII with its homologs shows it is highly-conserved in its C-terminus. From many years of isolating T4 mutants with *r* plaques, Bruch *et al* acquired a number of non-functional *rIII* mutants includes 12 missense mutations, two nonsense mutations and a flameshift mutation (67). Majority of the single-missense mutations of *rIII* are located in the conserved regions, and interestingly, half of these mutations are mapped to the last 10 aa residues of the RIII, a putative C-terminal β-sheet. This may indicate the C-terminus of RIII is critical to its function. However, no molecular analysis of RIII has been done so far.

The LIN signal

Despite previous efforts for identifying the signal of LIN, the nature of this signal is still not understood. One clue about the LIN signal is that T4 phage “ghosts”, in which the contents of the capsid have been released by osmotic shock (17), are able to adsorb to and kill the uninfected cells with unit efficiency but do not induce LIN in

previously infected cells (72). The T4 capsid contains a single copy of 170 kb dsDNA molecule with glycosylated hydroxymethylcytosine (HMC) instead of cytosine, and about 1000 internal head proteins (IPs) which are required for the inhibition of host restriction systems (64). In a normal T4 infection, these contents are normally ejected from the phage head and go into the host cytoplasm through the T4 tail tube, which makes a temporary pore in the IM. However, by ~5 min after the infection, two T4 early gene products, Imm (product of gene *imm*) and Sp (product of gene *sp*) accumulate in the cytoplasmic membrane and periplasm, respectively, and prevent the tail tube of superinfecting phage from forming the temporary hole in the IM (61). As a result, the contents of the phage capsid are ejected into the host periplasm (17). Since T4 ghosts fail to induce LIN, it is reasonable to presume that at least one of the capsid contents constitute the LIN signal (53). The IPs are not considered likely as the LIN signal because IPs are heterogeneous among the T-even phages (T2, T4, and T6), but all of these phages can cause LIN in a T4 infection (53). Also there is no significant conservation between different IPs. A T4 or T4-like phage genome DNA that ejected from the superinfecting phage heads will end up in the host periplasm and will be degraded by periplasmic endonuclease I (73). This event can significantly change the concentration of DNA in the periplasm. If single T4 genome DNA is ejected into *E. coli* periplasm, which has 8%-16% of the total cell volume (74), the concentration of DNA (nucleotides) in the periplasm could reach ~1000 μg/ml, which is comparable to the cytoplasmic DNA concentration of the host cell (~8000μg/ml) (75). This high concentration of DNA may affect the properties of the periplasmic space and lead to the

imposition of LIN. The absence of HMC-DNA modification has no effect on LIN (53). Therefore any external DNA fragments which have been ectopically ejected into the periplasm could be the LIN signal candidate. Although an electrophoretic mobility shift assay (EMSA) result suggests a weak interaction between purified sRI and 30bp dsDNA (54), there is no enough evidence that shows the strong interaction of DNA and any lysis protein or antiholin. It is also not clear how the DNA could affect the stability and function of the antiholin RI.

Lysis of phage Mu

Mu, the transposable phage

The coliphage Mu was discovered accidentally in 1960s by A. L. Taylor and named 'Mu' for its ability of causing mutations of the host *E.coli* after lysogenization (76). Mu is a myophage that contains an icosahedral head (~50 nm diameter), a long contractile tail, and a baseplate with six short tail fibers attached to it (77). It has a 36,717kb linear dsDNA genome; however each Mu phage particle contains ~40kb of dsDNA in its head capsid, due to ~100bp (left) and ~2,000bp (right) of host DNA attached to the each end of the genome as the result of headful packaging (77). Ever since its discovery, Mu is especially of interest to researchers for its several unique features, including the random integration of phage genome into the host chromosome, the replication of phage DNA via transpositions, the distinctive DNA modifications and regulation of gene expression by DNA methylation (76). Among these features, the most fascinating one is the transposition ability of Mu (78). The study of Mu

transposition proved the existence of insertion elements in bacteria for the first time and later lead to the establishment of the very first *in vitro* transposition system (78), which significantly pushed forward our knowledge of movable elements of genomes.

The two main protein factors of transposition are transposase MuA and MuB. MuA is a recombinase containing an RNaseH-like domain with catalitic triad DDE (39, 79); The target DNA activator MuB is critical for finding of the target host DNA (Figure 1.18) (77). The infectious Mu transposition can be summarized in four steps: 1) During the initial infection, 'transpososome (or intasome)' is formed by pairing the attL and attR end of the Mu genome injected into the host cytoplasm; 2) MuA makes hydrolytic nicks on both phage DNA and the proximate, non-specific target host DNA; 3) The phage genome is ligated to the host DNA, creating the hybrid of the phage-host strand (strand transfer); 4) Transposition is finalized by the removal of the flanking host DNA linked with the Mu genome and the repair of gaps Morphogenetic Structures Present in Lysates of amber mutants of Mu (77, 80). When Mu becomes lytic, the phage and the host enzymes assist the formation of intasome and the creation of the two replication forks at the phage-host joint strands, which allows the copying of the whole phage genome using the 3' hydroxyls of host DNA as replication primers without excision from the host genome (81). Repetition of this 'copy-and-paste' replicative transposition can result in >100 copies of the Mu genome in a single host cell. This makes Mu the most powerful transposon we know of to date (76).

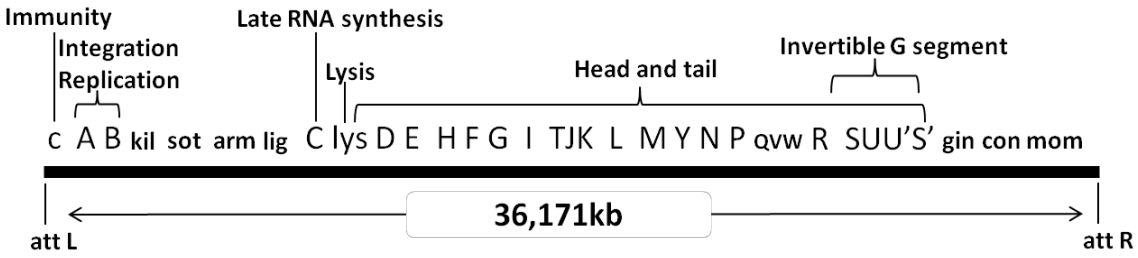

Figure 1.18 Map of phage Mu.
Genes with assigned functions are labeled with their names. Adopted from (77).

The initial DNA integration and transposition are required for both lytic and lysogenic pathway of Mu. The decision of lysogeny or lysis depends on the ratio of Repc and Ner repressors (77). Once the Mu lysogen is formed, it is quite stable and cannot be induced by UV or DNA-damaging chemicals (77). Therefore, the lytic phase of Mu was mainly studied with Mu*c*^{ts}mutants, which allow the deactivation of the main repressor c protein at 42°C. As a repressor, c not only maintains the lysogeny, but also immunizes the lysogen from the phage superinfection (77).

Mu lysis gene: lys

Mu is one of the most well-studied phages besides λ and T4, with which Mu shares the distinction of having a book entirely dedicated to it (77). However, because of the focus and the emphasis on its transposition mechanism, very little study was conducted toward the lysis process of Mu. In 1970s and 1980s, the majority of Mu genes, both essential or nonessential for Mu growth, were mapped and analyzed by either deletion, insertion or nonsense (amber) mutations (82). The most up-to-date published annotation of Mu genome lists 55 predicted genes, among which 18 genes have undefined function (83). As shown in Figure 1.18, *lys* is the only gene which was identified and assigned with any phenotype related to lysis (84). The amber mutations in *lys* gene abolished lysis and cause the failure of virion release, but did not affect the accumulation of the progeny phage particles or the phage morphology (82). However, the molecular function of *lys* or the mechanism of lys-mediated lysis has never been studied.

Taking advantage of the complete genome sequence of Mu published in 2002 (83) and the progress in understanding various phage lysis systems, the most recent bioinformatics study suggested that product of *lys* is a SAR-endolysin. Lys, a protein consists of 171 aa, is predicted to have a N-terminal SAR-domain and a large C-terminal soluble domain includes classic T4 E-like catalytic triad sequence (ExxxxxxxxDxxxxxT) (Figure 1.19) (40). The fact that Lys does not have any Cysteine in it SAR-domain suggests that it is an R^{21}-like SAR-endolysin, which requires the conformational change of SAR-domain to activate the enzymatic function (38). This prediction brings up a possibility that Mu, as a dsDNA phage infecting Gram-negative host, may have a lysis paradigm consists of pinholin, SAR-endolysin and spanins, similar to phage 21. Summer et al reported that the product of gene *23* of Mu is an equivalent of λ i-spanin Rz, and a newly identified gene *23a* which is partially overlapped with gene *23* is the o-spanin Rz1 equivalent (47). The fact that spanin gene candidates are in proximity of *lys* (gene 22) suggests Mu might have λ -like 'lysis cassette'. However, none of the Mu gene products fits the topology profile of any experimentally characterized holin. Other genes involved in Mu lysis are yet to be determined.

Questions addressed in this dissertation

This dissertation describes four major research topics:

(1) Determine the molecular function of *rIII* in T4 lysis inhibition. The function of the product of *rIII* gene, an 82aa protein, was studied in a λ context that mimics the wt T4 infection process, and the interaction partner of RIII was identified. A new model for

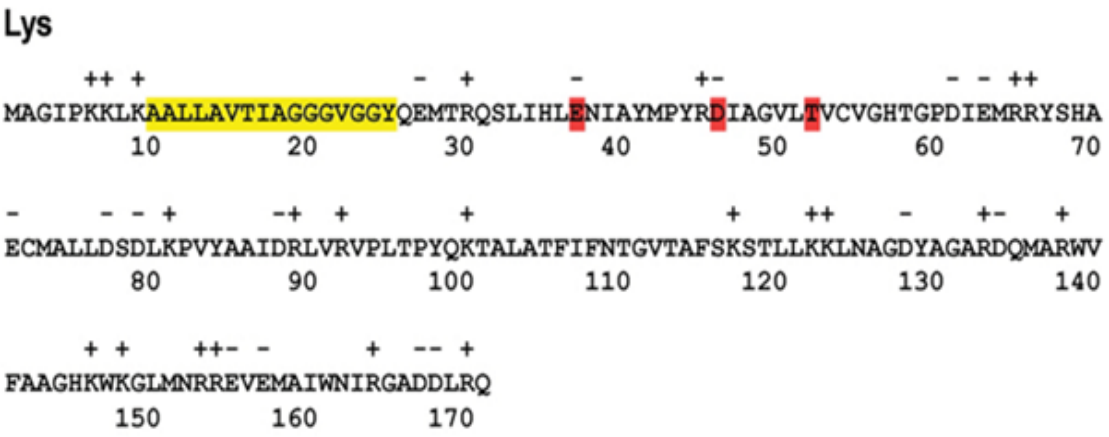

Figure 1.19 Primary structure of Mu SAR-endolysin Lys.
The SAR-domain is highlighted in Yellow . The catalytic triad residues are marked with red (85)..

T4 LIN that involves the stabilization of a complex of three proteins in two compartments of the cell was proposed.

(2) Biochemical/structural characterization of RIII. Un-tagged native RIII protein was purified to homogeneity. The character and the structure of RIII were studied using various biochemistry methods. The quantification of the RIII protein in the T4 infection was carried out to answer the questions regarding to the stoichiometry of proteins involved in T4 LIN.

(3) Lysis paradigm of phage Mu. The lysis paradigm of a classic phage Mu was studied. The holin-antiholin pair, Gp19 and Gp20, was identified and their functions were experimentally confirmed. In addition, a new class of protein factor that regulates the lysis by activating the endolysin function is discovered and tested by biochemical and genetic approaches.

(4) Dependence of IM receptor in newly isolated lambdoid phages. It is reported that phage λ relies on E. coli mannose transporter complex ManYZ as the IM receptor for its DNA penetration. However, mechanism of the interaction between the IM receptor and the phage tail proteins remains controversial. Two new lambdoid siphophages that strictly require ManYZ complex for their infection were isolated from the environmental samples, and their suppressors for *manYZ* hosts were sequenced to study the phage genes involved in the IM receptor interaction.

CHAPTER II

T4 RIII IS A CYTOPLASMIC CO-ANTIHOLIN

Introduction

The *r* genes of the T-even phages, first identified by laboratories of the Phage Group in the 1940s (52, 65), have a special place in the history of molecular biology. Detailed studies of the first three loci discovered, - *rI*, *rIIAB*, and *rIII* - were foundational in working out the fundamentals of inheritance, the genetic code, mutation, recombination, DNA repair, and gene structure (86-90). These mutable loci were originally discovered by their distinctive plaque morphology: large, clear, sharply-defined plaques, easily distinguished from the small, fuzzy-edged turbid plaques of the parental phages (65). The "*r*" designation meant "rapid lysis", which refers to the observation that the mutant phages isolated from the r-type plaques caused rapid, culture-wide lysis at ~25 min after infection, whereas cultures infected with the parental phages continued to increase in mass and accumulated progeny virions intracellularly for hours, in a state called "lysis inhibition"(LIN)(66). In the ensuing decades, more loci were assigned as *r* genes based on mutant plaque phenotypes; at one point, there were *r* genes numbered up to *rVI* that were assigned map positions (67, 91). In 1998, Paddison *et al*. (64) reviewed this field and concluded that only *rI*, *rIII* and *rV* were directly involved in LIN, with the other genes causing lysis phenotypes through indirect physiological pathways. The *rV* mutants were shown to be missense alleles of gene *t*, which encodes T, the holin of phage T4 (69). Holins are the master lysis control proteins

of Caudovirales (17), acting to terminate the infection cycle by permeabilizing the cytoplasmic or inner membrane (IM) at a programmed time. It followed that the simplest operational model to explain the involvement of the remaining loci associated with direct LIN defects, *rI* and *rIII*, would be that the RI and RIII proteins were required to inhibit the lethal function of T and thus establish the LIN state (64, 67).

More recent studies on T and RI have confirmed aspects of this operational model for LIN and provided molecular details for the lysis pathway of T4 (54, 56, 59). Like other holins, including the well-studied S105 holin of phage λ, the T holin accumulates harmlessly in the host IM until it suddenly forms lethal, micron-scale membrane lesions at an allele-specific time. This event, which is defined as holin triggering, results in the escape of cytoplasmic endolysin E (product of gene *e*) (17) into the periplasm, where it rapidly degrades the cell wall. In turn, the loss of cell wall activates the spanin complex (product of *pseT.2* and *pseT.3*)(47), which then disrupts the outer membrane (OM) and completes the release of the progeny. In single infections, T4 completes this three-step pathway in ~25 min (65). However, if the T4-infected cells are super-infected by other T4 (or T-even phages) after the first five minutes of the infection cycle, LIN is imposed (92). There has been progress on the molecular basis of LIN (34, 35, 56, 70). While most holins have two or more transmembrane domains (TMDs) and only short soluble loops connecting them (17), the T holin has a unique structure, with only a single TMD and significant N- (34 aa) and C-(163 aa) terminal cytoplasmic and periplasmic domains, respectively (56, 57) (Figure 2.1).

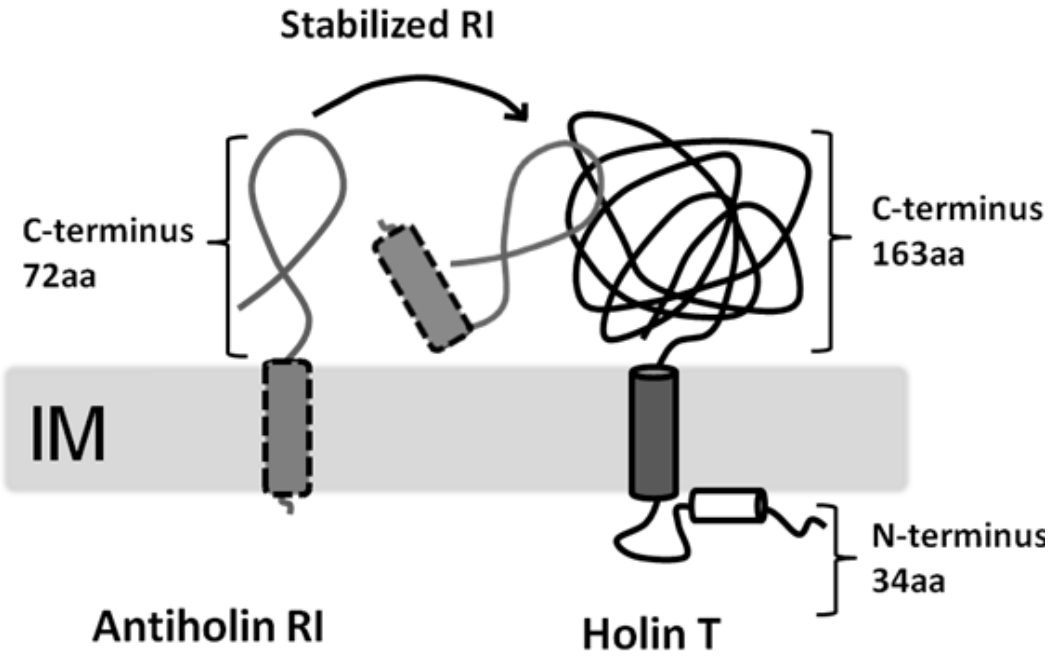

Figure 2.1 Topology of T4 holin T -antiholin RI interaction.
T is an inner membrane protein with a single TMD (shown as a solid cylinder) and an amphipathic helix (shown as a white cylinder). RI has a SAR (Signal Anchor-Release) domain (shown as a dash line rectangle) which allows RI to be spontaneously released in to the periplasm (35). If stabilized by the LIN signal, periplasmic RI binds to the C-terminal globular periplasmic domain of T. IM, inner membrane.

Moreover, the RI protein was shown to have a SAR (signal anchor release) domain, which is a TMD that can escape from the membrane (35). By virtue of this domain, RI is secreted initially as a membrane-tethered periplasmic protein and then releases into the periplasm where, in single infections, it is degraded rapidly (35, 70). However, under LIN conditions, (i.e., when there is superinfection with a second T4 phage particle), RI is stabilized and accumulates in the periplasm, where it forms a complex with the cytoplasmic domain of T and inhibits triggering, thus imposing the LIN state. Additionally, if the SAR domain of RI is replaced by a cleavable Signal Peptidase I signal sequence, the processed RI protein over-accumulates in the periplasm in a stable, mature form, forms the complex with T and imposes LIN without requiring the superinfection activation (35, 70).

Because RI is a specific inhibitor of T, it is formally a member of a diverse class of proteins designated as antiholins (24, 33, 93, 94). Moreover, since RI inhibits T only under certain physiological conditions, it is the only antiholin known that transduces environmental information to effect real-time control of holin function and thus the length and fecundity of the phage infection cycle (34). However, despite these conceptual and mechanistic advances with T and RI, the genetic basis of the LIN phenomenon remains incomplete, some decades after the genetics of the *r* genes were first published, because no role has been found for *rIII* (65, 86). Although it was reported that RIII was not required for LIN on some *E. coli* K-12 strains (89), *rIII* shares with *rI* the feature that neither locus can suppress *t* lysis-null mutations and both loci are transcribed from both early and late promoters (64, 71). Recently, *rIII* was suggested to

play a role in the propagation of T4 in slow-growing host cells (95). Here, we present the preliminary results of *in vivo* and *in vitro* characterization of *rIII*. The results are analyzed in terms of a model that suggests direct molecular involvement of RIII in LIN as a new class of antiholin.

Materials and methods

Bacterial growth and induction

See Table 2.1 for the full list of phages and bacteria strains used in this study. Bacterial strains were plated on standard LB-agar plate supplemented with the appropriate antibiotics (ampicillin, 100 μg mL^{-1}; chloramphenicol, 10 μg mL^{-1}; kanamycin, 40 μg mL^{-1}). A single colony from a LB plate was used to inoculate 3 mL overnight culture at 30°C for λ lysogens and 37°C for non-lysogenic *E. coli* strains, as described before (34). Overnight cultures were diluted to A_{550} ~ 0.03 and grown at 30°C or 37°C with aeration. Bacterial growth and lysis were monitored as described (34) using a Gilford Stasar III sipping spectrophotometer (Gilford Instrument Inc, Oberlin, OH). The λ lysogens were induced as described (34, 59). All plasmid-cloned genes were induced with 1 mM isopropyl β-D-thiogalactosidase (IPTG).

Phage infection and preparation of phage lysates

Phage lysates were prepared by adding 10% $CHCl_3$ (v/v) to the *E. coli* cell culture after lysis , in either induced lysogens or by liquid culture infections, as described previously (35). The lysate was cleared by centrifugation at 5,000 x *g* and the supernatant was filtered through a 0.22 μm syringe filter. Phage infection experiments were carried out as described (34, 35). For liquid culture infections, host *E. coli* cells

Table 2.1 Phages, strains, and plasmids used in Chapter II.

Phages	**Description**	**Source**
T4wt	Bacteriophage T4D	Laboratory Stock
T4*rIII*	T4*r67*. H42 to R (CAU to CGU) mutation in *rIII* locus.	Laboratory Stock
T4Δ*rI*	Complete deletion of *rI* from nt 59204 to nt 59496 in T4D genome	(35)
T4Δ*rIII*	Complete deletion of *rIII* from nt 130779 to nt 131042 in T4D genome	This study
λ -t	□with holin gene *S* replaced with T4 holin gene *t*	(59)
λS_{A52G}	□cI857 carrying Ala52Gly early lysis allele of *S* holin gene.	(26)
T4rBB9	W16 to stop (UGG to UGA) mutation in *rIII* locus	Laboratory Stock
T4rES35	H42 to Q (CAU to CAA) mutation in *rIII* locus	Laboratory Stock
T4rES40	K82 to E (AAG to GAG) mutation in *rIII* locus	Laboratory Stock
Bacteria Strains	**Description**	**Source**
CQ21	*E. coli* K-12 *ara leu lacI^q purE gal his argG rpsL xul mtl ilv*	Laboratory Stock
CQ21λ-t	CQ21 lysogen carrying □-t prophage	(59)
CQ21λS_{A52G}	CQ21 lysogen carrying □SA52G prophage	This study
BL21(DE3) *fhuA*::Tn*10*	*E. coli* B *ompT* r_B^- m_B^- (*P_{lac}UV5::*T7 *gene1*) *slyD::Kan fhuA*::Tn*10*	Laboratory Stock
B834	*E. coli* B *ompT* r_B^- m_B^- *met* $^-$	Laboratory Stock
MG1655	*E. coli* F- λ - *ilvG- rfb-50 rph-1*	Laboratory Stock
MDS12 *tonA*::Tn*10* *lacI*^{q1}	MG1655 with 12 deletions, totaling 376,180 nt including cryptic prophages	(35)
DHP1	*E. coli* F- *cya-99 araD139 galE15, galK16, rpsL1 (Strr) hsdR2 mcrA1 mcrB1*	(96)
Plasmids	**Description**	**Source**
pZE12	ColE1 origin; $P_{LlacO-1}$ (PL promoter with three lacO operators); AmpR	(97)
pZE12-luc	Luciferase gene luc cloned under $P_{LlacO-1}$	(97)
pZE12RI	T4 *rI* cloned under $P_{LlacO-1}$ with native SD	(34)
pZE12RIII$_o$	T4 *rIII* cloned under $P_{LlacO-1}$ with native SD	(34)
pZE12RIII$_s$	T4 *rIII* cloned under $P_{LlacO-1}$ with plasmid SD	This study
pZE12RI-RIII	Tandem clone of *rI- rIII* inserted between KpnI and XbaI site	This study

Table 2.1 Continued

Plasmids	Description	Source
pET11a-RIII	pBR322 origin, T7 promoter, carrying codon 1-82 of *rIII*	This study
pET11a-RIII$_{H42R}$	H42 to R (CAU to CGU) mutation in *rIII*	This study
pTB146	*bla lacI*^{q} PT7::h-sumo	(98)
pTB146-RIII	Codon 2-82 of *rIII* gene inserted between SapI and XhoI site	This study
pTB146-nT	Codon 2-34 of *t* gene inserted between SapI and XhoI site	This study
pCH364	T18-empty (AmpR);N-terminal tag	(96)
pKNT25	Empty-T25 (KanR); C-terminal tag	(96, 99)
pKT25	T25-empty (KanR); N-terminal tag	(96, 99)
pCH364RIII	Codon 2-82 of *rIII* gene inserted between BamHI and EcoRI site	This study
pKNT25RIII	Codon 2-82 of *rIII* gene inserted between XbaI and EcoRI site	This study
pKT25nT	Codon 2-34 of *t* gene inserted between BamHI and EcoRI site	This study

Table 2.2 Primers used in Chapter II.

Primer name	Sequence	Source
$RIII_S$ CLONING F	CGGTACATTAAACAATTACAACACGCTC	This study
$RIII_S$ CLONING R	GGCTCTAGATTACTTCAGTGTTACCACAAAGTG	This study
$RIII_S$ PET F	GGAATTCCATATGATTAAACAATTACAACACGCTC	This study
$RIII_S$ PET R	GCGGGATCCTTACTTCAGTGTTACCACAAAGTG	This study
RIII DEL +500 F	GGGGTACCCATCTGTTAACAAAAAGGAAAAACG	This study
RIII DEL -500 R	GCTCTAGAGCGTTCAGATTAATCGTTTTCA	This study
RIII DEL MIX F	TTTTAATCTCTAACGAGGGAGATTCACTGCCTTAGTGTGAGC	This study
RIII DEL MIX R	CCGAGTTTTAATCTCTAACGAGGGAGATTCACTGCCTTAGT	This study

were grown to A_{550} ~ 0.3 and infected at a multiplicity of infection (MOI) ~ 5. To observe plaque morphology, 100 μL overnight cultures of host cells were added to 3 mL of LB top agar and immediately poured on standard LB-agar plates. 5 μL of phage lysates with proper dilutions were spotted onto the top agar. For complementation experiment, BL21(DE3) *fhuA*::Tn*10* cells carrying pET11a vectors were grown to A_{550} ~ 1 at 37°C, and induced with 1mM IPTG for 2 h before mixed with LB top agar and poured onto LB plates containing proper antibiotics and 1mM IPTG. All plates were incubated ~16h at 37°C. The plaque sizes were analyzed using ImageJ software (NIH, Bethesda, MD).

Standard DNA manipulations and sequencing

All plasmids used in this study are listed in Table 2.1. Isolation of plasmid DNA, DNA amplification by polymerase chain reaction (PCR), DNA transformation, and DNA sequencing were performed as previously described (50, 56, 57). Oligonucleotides (primers) DNA sequences are listed in Table 2.2. All purified oligonucleotides (primers) were purchased from Integrated DNA technologies (Coralville, IA). Restriction and DNA-modifying enzymes were purchased from New England Biolabs (Ipswich, MA). Manufacturer's instructions were followed when performing reactions. The DNA sequence of all constructs was verified by sequencing service provided by Eton Bioscience (San Diego, CA).

PCR and plasmid construction

T4D phage lysate was directly used as the PCR template for cloning out T4 genes. Pfu DNA polymerase was used for all PCR reactions following standard

protocols provided by Promega (Madison, WI). Site-directed mutagenesis was performed as described (57). The *rIII* gene either with its native ribosome binding site (GAG) or a stronger ribosome binding site (AGGAG) was cloned into the medium-copy IPTG-inducible vector pZE12 (97). Plasmid pZE12-$RIII_o$ and pZE12-$RIII_s$ were constructed by inserting T4 DNA from nt 130738 to nt 131080 ($RIII_o$), or from nt 130785 to nt 131033 ($RIII_s$) into pZE12 between KpnI and Xbal sites. Plasmid pET11aRIII has the same insertion as pZE12$RIII_s$ between its Ndel and BamHI sites. Plasmid pZE12RI-RIII was made by inserting a tandem clone of *rI* and *rIII* genes with their original ribosome binding site into plasmid pZE12. These plasmids were transformed into a CQ21λ-t lysogen, in which the λ holin gene *S* has been replaced by T4 gene *t* (59). In this system, RI or/and RIII can be expressed *in trans* to T from pZE12 plasmids by adding 1mM IPTG after lysogenic induction. A λS_{A52G} lysogen was used as a control, since the S_{A52G} confers a ~20 min lysis time, similar to the *t* gene (26). Plasmid pTB146 is a derivative of plasmid pET11a encoding an N-terminal his6-SUMO tag (50, 98). Plasmids encoding His-SUMO-tagged versions of RIII and nT (the N-terminal domain of T), pTB146-RIII and pTB146-nT were constructed by inserting codon 2-81 of the *rIII* gene, or codons 2-34 of the *t* gene (nt 160218 to nt 160322 of T4 genome), respectively, into the pTB146 plasmid between its SapI and XhoI sites.

Constructing T4 rIII *deletion mutant*

T4Δ*rIII* was constructed by homologous recombination between pZE12-Δ*rIII* and T4D, as described previously for T4Δ*rI* (35). pZE12-Δ*rIII* was made by deleting the *rIII* gene from plasmid pZE12-*rIII*-flank, which contains T4 DNA from nt 130231 to nt

131541 between its KpnI and Xbal sites, using our previously described method (100). Plasmid pZE12-Δ*rIII* was transformed into *E.coli* strain MDS12 *tonA*::Tn*10 lac*I^{q1}, and the transformants were grown to $A_{550} \sim 0.4$. The culture was infected with T4D phage at a MOI=10 for 3 h at 37°C with aeration, and then lysed by adding 10% v/v $CHCl_3$. T4 *rIII* recombinants in this lysate were enriched three times for early lysis as described (35). The enriched lysate was plated on *E. coli* B834 and screened for *r* plaque morphology. The Δ*rIII* deletion was confirmed by PCR and sequencing.

SDS-PAGE and Western blotting

SDS-PAGE and Western blotting were conducted as previously described (57). 10% trichloroacetic acid (TCA) was used to precipitate proteins from the whole-cell samples. Reducing sample loading buffer (SLB) supplemented with β-mercaptoethanol was used for resuspending protein samples unless otherwise indicated. RIII proteins transferred onto PVDF membrane were detected using rabbit polyclonal α-RIII antibody purchased from Genscript (Piscataway, NJ). The monoclonal α-his-tag antibody (α-his) was purchased from Sigma-Aldrich (Carlsbad,CA). To detect proteins, blots were incubated overnight at 4°C with α-RIII or α-his at a dilution of 1:4000 in 3% milk-TBS buffer. Blots were developed with the West Femto SuperSignal Chemiluminescence kit purchased from Thermo Fisher Scientific (Rockford, IL). The chemiluminescence signal was detected using a Bio-Rad Chemidoc XRS (Bio-Rad Laboratories, Hercules, CA). Images were obtained and analyzed by Quantity One 1-D Analysis Software (Bio-Rad Laboratories, Philadelphia, PA).

Bacterial two hybrid assay

Bacterial two hybrid (B2H) assays were conducted as described previously (96, 99, 101). Plasmids were constructed by inserting codons 2-81 of the *rIII* gene, or codons 2-34 of the *t* gene into plasmids pCH363, pKNT25, pCH364, and pKT25, as described (96). Different pairs of plasmid were transformed into strain DHP1 and grown to A_{550} ~ 0.3 in LB with 0.2% glucose and appropriate antibiotics (ampicillin, 50 μg mL^{-1}; kanamycin, 25 μg mL^{-1}). For the plate assay, 5 μL of cell cultures were spotted on M9 minimal media plates supplemented with 0.2% glucose, 40□μg mL^{-1} X-Gal, 150 μM IPTG and proper antibiotics and incubated for 48h at 25°C.

Pull-down assays

Plasmid pET11a and pTB146 derivatives described above were transformed into BL21(DE3) *fhuA*::Tn*10* strains. Pull-down assays were conducted as instructed by manufacturer protocol from Dynabeads® His-Tag Isolation and Pulldown kit (Thermo Fisher Scientific, Rockford, IL). All incubation and reaction was carried out in 4°C and beads were collected using DynaMag™-2 Magnet (Thermo Fisher Scientific). All samples were resuspended in SLB and boiled for 5 min to elute proteins, which were then analyzed by SDS-PAGE and Western blotting were performed as described above.

Results

rIII *is required for LIN in both E. coli B and K-12 background*

The role of *rIII* in LIN has been ambiguous, with reports differing in whether *rIII* was required on *E. coli* B but not K-12 strains (67, 89, 102). In our hands, the classic *rIII* alleleT4*r67*, which was used as the standard allele in the early T4 genetic map studies,

formed r-type plaques on the lawn of both *E. coli* B834 and MG1655 (Figure 2.2A), with somewhat smaller plaques compared to those formed by T4*rI*, but significantly larger than wt T4 plaques (Table 2.3). It was also reported that different T4 *rIII* mutants differed in plaque size, suggesting a possible correlation between the location of mutations on *rIII* locus and plaque morphology (103). However, when we compared plaque morphologies of four different T4 *rIII* defective mutants (T4*r67*, T4*rBB9*, T4*rES35*, and T4rES40) on B834, we did not observe significant differences (Table 2.3). Nevertheless, as the first step for interrogating the role of *rIII* in LIN, we constructed an *rIII* deletion allele, T4Δ*rIII*, to eliminate the potential for partial reversion. As shown in Table 2.3 and Figure 2.2A, T4Δ*rIII* formed r-type plaques that were larger than wt plaques, but smaller than those of T4*rI*. Moreover, the wt plaque morphology could be complemented by a plasmid-borne *rIII* gene (Figure 2.2B). In infections of both *E. coli* B834 or K-12 cultures under conditions where the wt T4 exhibited LIN, T4Δ*rIII* infections showed lysis at ~25 min, in both cases reproducibly later than the lysis time of T4Δ*rI* (~18 min), (Figure 2.3A). The simplest notion, based on the established role of RI in LIN, is that RI expressed in the T4Δ*rIII* infection causes transient LIN, and that, by extension, RIII is required for stable LIN in both *E. coli* B and K-12 strains.

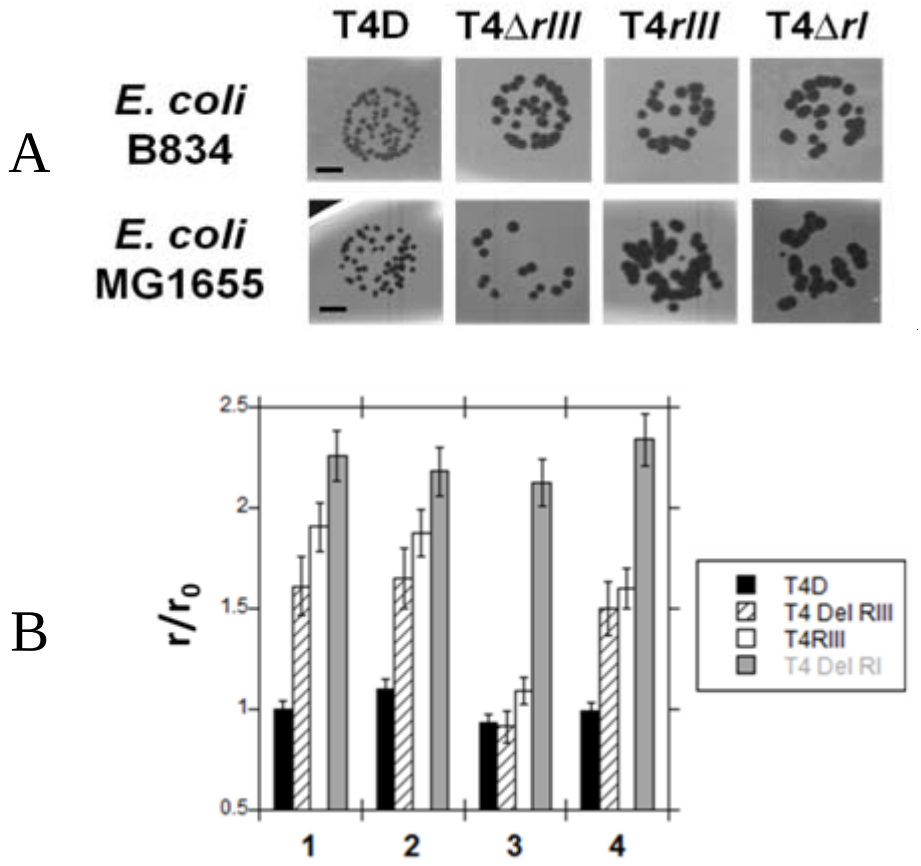

Figure 2.2 Plaque morphologies of T4 and its *r* mutants.
(A) Plaque morphology of T4 wt (T4D) and T4 mutants on either *E. coli* B strain (B834) or *E. coli* K-12 strain (MG1655). The black bar represents 2.5mm. Average plaque sizes of T4D, T4Δ*rIII*, T4*rIII* and T4Δ*rI* on *E. coli* B834 or *E. coli* MG1655 are listed in Table 2.3.
(B and C) Complementation of *r* plaque morphology. T4 *rIII* mutants plated on *E. coli* strains expressing wt RIII restored wt T4 plaque morphology. In B, differences in plaque sizes were shown as the ratio of the average phage plaque radius (r) to the average plaque radius of T4D plated on B834 (r_0). 1, MG1655; 2, BL21(DE3) *fhuA*::Tn*10* no plasmid; 3, BL21(DE3) *fhuA*::Tn*10* pET11a pET11a-RIII; 4, BL21(DE3) *fhuA*::Tn*10* pET11a-$RIII_{H42R}$.

Table 2.3 Mean diameter of phage plaques (mm)

Host / Phage	B834	MG1655	BL21(DE3)	BL21(DE3) pET11a RIII	BL21(DE3) pET11a RIII-H42R	BL21(DE3) pTB146	BL21(DE3) pTB146-nT
T4D	0.57 (±0.04)	0.67 (±0.04)	0.63 (±0.06)	0.53 (±0.05)	0.57 (±0.05)	0.68 (±0.04)	1.07 (±0.05)
T4*rIII*	1.09 (±0.09)	1.43 (±0.07)	1.07 (±0.09)	0.62 (±0.03)	0.92 (±0.08)	1.32 (±0.08)	1.38 (±0.06)
T4□*rIII*	0.92 (±0.08)	1.12 (±0.06)	0.95 (±0.07)	0.52 (±0.06)	0.86 (±0.08)	1.13 (±0.08)	1.17 (±0.07)
T4□*rI*	1.29 (±0.09)	1.88 (±0.11)	1.24 (±0.07)	1.21 (±0.03)	1.34 (±0.06)	-	-
T4rBB9	0.89 (±0.15)	-	-	-	-	-	-
T4rES35	1.02 (±0.05)	-	-	-	-	-	-
T4rES40	1.05 (±0.05)	-	-	-	-	-	-

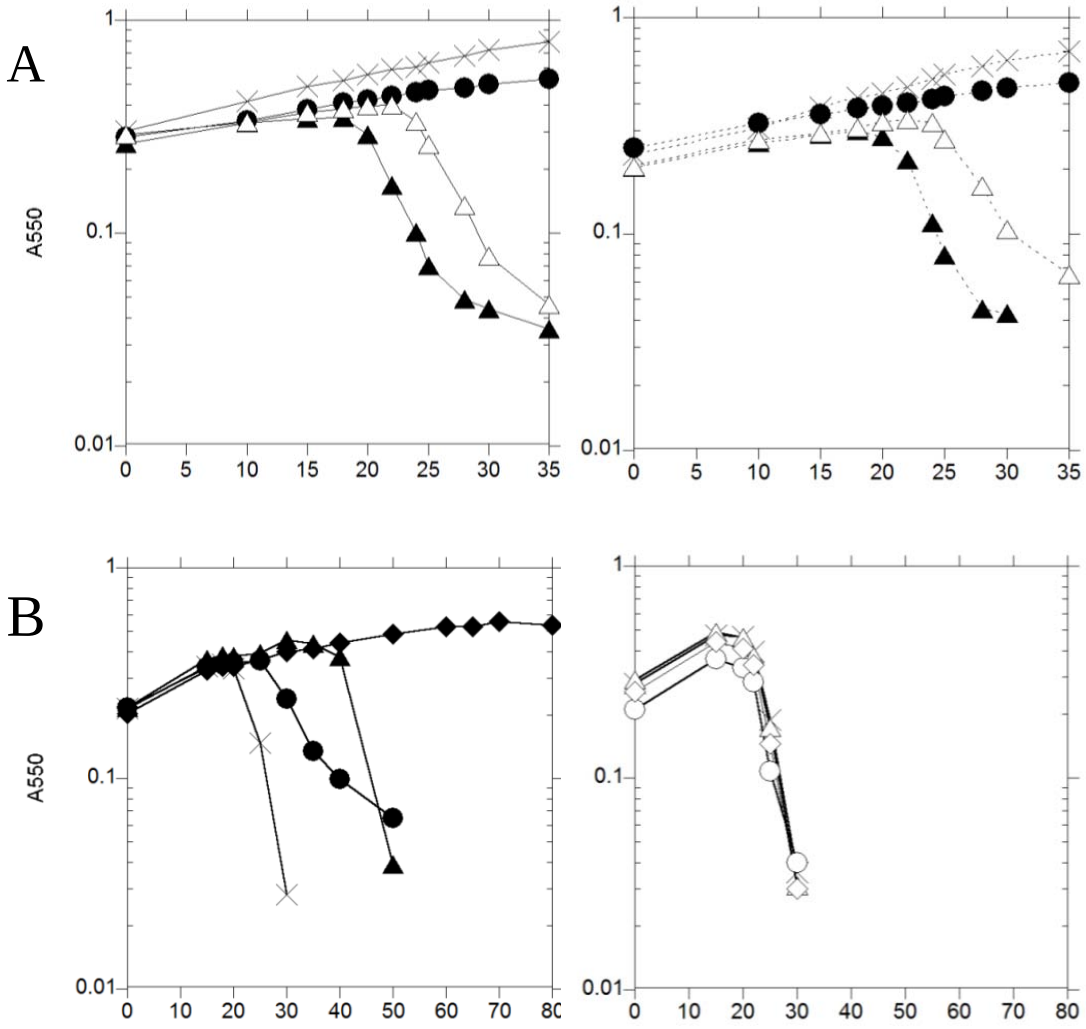

Figure 2.3 Lysis curves.
(A) Lysis in infections of T4 and derivatives infecting *E. coli* B strain B834 (Left, solid line) or K-12 strain MG1655 (right, dotted line). ×, no phage; ●, T4D (wt); ▲, T4Δ*rI*; △, T4Δ*rIII*. Cultures were grown to A_{550} ~ 0.25 at 37°C, then infected with T4 at MOI~5.
(B) Inductions (at t=0) of CQ21λ-t (Left, solid symbols) or CQ21λS_{A52G} (right, open symbols) lysogens carrying indicated genes cloned under IPTG control in the context of the pZE12 plasmid. Plasmids were also induced by addition of 1mM IPTG at t=0. ×, luc (negative control); ▲ and △, RI; ● and ○, RIIIs; ◆ and ◇, RI-RIII.

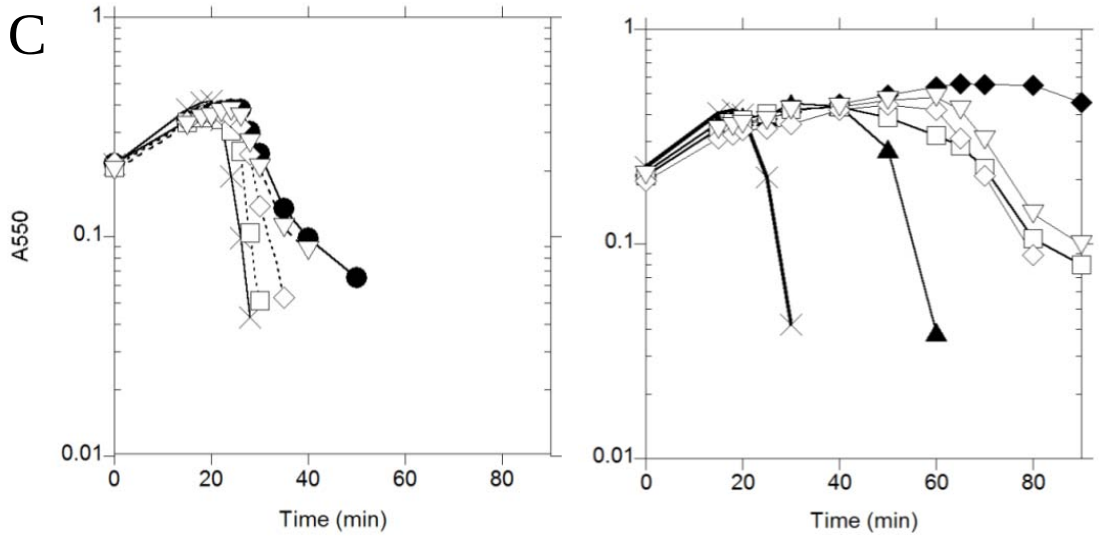

Figure 2.3 Continued.
(C) RIII missense mutants exhibit intermediate LIN phenotype. CQ21λ-t lysogens carrying pZE12 plasmids with indicated genes were induced at t=0. ×, luc; ▲, RI; ●, RIIIs; ◆, RI-RIII. Left pane with dotted lines: □, $RIII_{G24D}$; ◇, $RIII_{H42R}$; ▽, $RIII_{A70V}$. Right panel with solid lines: □, RI-$RIII_{G24D}$; ◇, RI-$RIII_{H42R}$; ▽, RI-$RIII_{A70V}$.

Identification of the RIII protein

rIII encodes an 82aa polypeptide without any secretion signals (Figure 2.4A) and had not been identified as a protein species in T4-infected cultures. We raised a polyclonal antibody against an RIII oligopeptide sequence predicted to be highly immunogenic (Figure 2.4A). The RIII protein could be visualized by immunoblotting in samples taken from cells infected with T4wt, but not with T4Δ*rIII* (Figure 2.4B); the mobility of this RIII species corresponded to a slightly lower molecular mass than predicted (8.9 kDa versus 9.3 kDa predicted), presumably due to the high content of charged residues (24 out of 82).

Recapitulating the role of RIII in LIN in the λ context

To address the role of RIII in LIN, we used a convenient system based on the inducible λ prophage, λ-t, in which the λ holin gene *S* is replaced by T4 *t* (59). Not only was this hybrid phage previously shown to recapitulate T4 lysis timing and LIN at physiological levels of expression, it allows the co-expression of selected T4 genes cloned in inducible plasmid vectors without the confounding effects of T4-mediated host DNA degradation and translational repression(34, 56, 59, 60). This system mimics the T-dependent lysis in the λ context where effects of T4 genes other than *t* are excluded. To provide RIII in trans, the *rIII* gene was cloned into a medium copy-number plasmid vector pZE12 (97). Two isogenic clones were constructed with different Shine-Dalgarno (SD) sequences serving the *rIII* cistron, one the relatively weak native sequence (GAG) and the other with a stronger near-consensus sequence (AGGAG). The resulting plasmids pZE12RIII$_o$ (original SD sequence) and pZE12RIII$_s$ (strong SD sequence) were

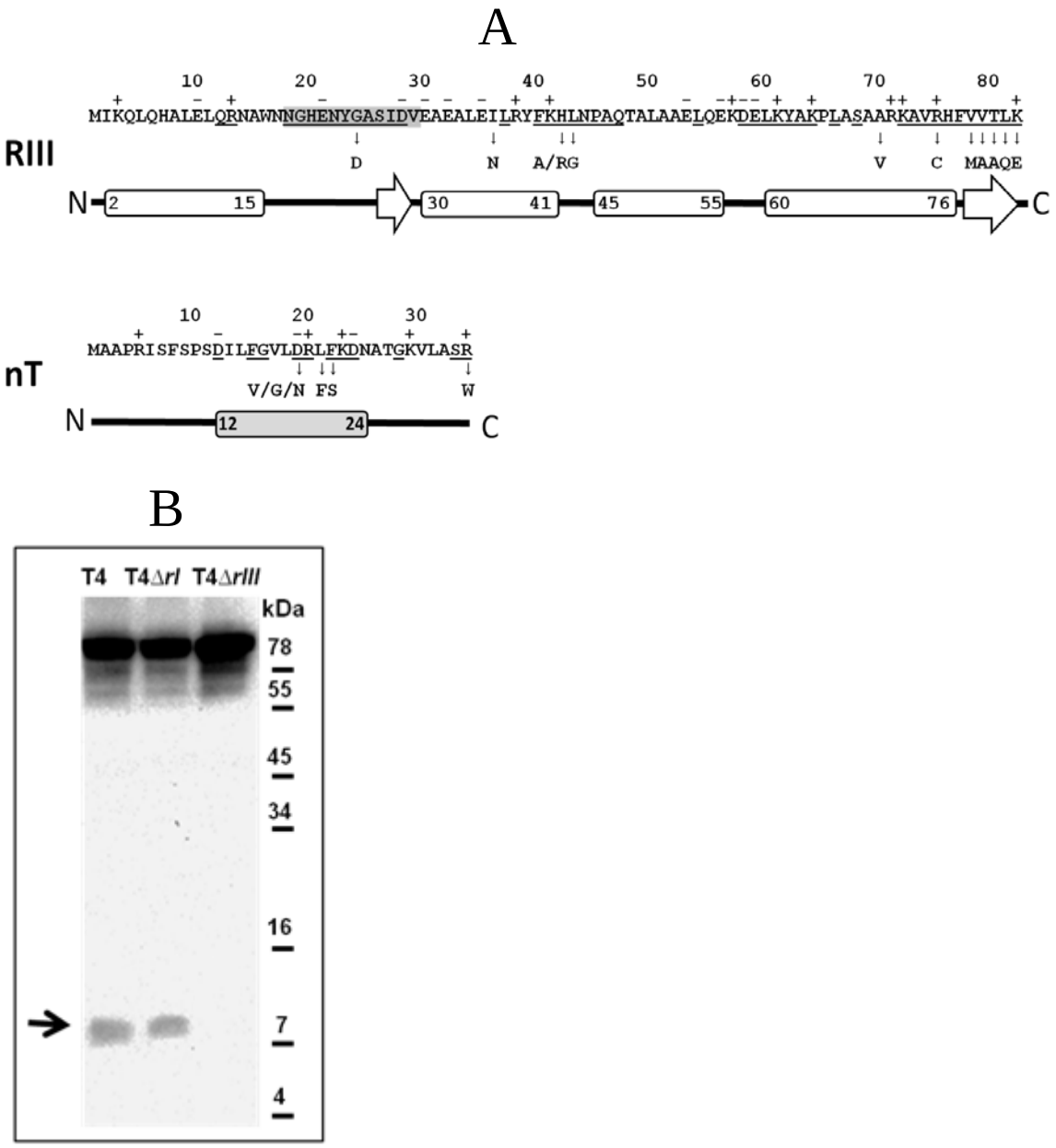

Figure 2.4 Structure of RIII and nT.

(A) Primary structure of RIII and N-terminus of T4 holin T (nT) with LIN-defective and lysis-defective alleles indicated by black arrows. Conserved residues are underlined. The shaded area represents the oligopeptide used to raise theα-RIII antibody. Predicted secondary structure is indicated: white box, helix; solid line, turn; white arrow, beta-sheet; grey box, amphipathic helix.

(B) RIII protein accumulates during infection. For each sample, 1 A600 equivalent of cells was loaded. The α-RIII antibody was used in Western blotting. Black arrow indicates predicted molecular mass (9.3kDa) for RIII monomer.

transformed into the λ-t t lysogen. Induction of the λ-t lysogen resulted in reproducible and sharply defined lysis at ~ 20 min; induction of pZE12RIII$_s$ with IPTG conferred a mild but reproducible lysis delay of ~5 min (Figure 2.3B). In contrast, lysis was much more severely affected by induction of an isogenic clone of the T4 antiholin gene *rI*, as previously shown (34). Induction of pZE12RIII$_o$ did not affect the lysis timing, probably due to the lower protein expression level (Data not shown).

We next asked if RIII can extend RI-mediated LIN. An isogenic plasmid with *rI* and *rIII* cloned in tandem was constructed and introduced into the λ-t lysogen. Induction of this plasmid, pZE12RI-RIII, led to a drastically delayed LIN compared to induction of either pZE12RIII$_s$ or pZE12RI (Figure 2.3B); under these conditions, the LIN state lasted up to 80 min and then gradually deteriorated. Using this system, we tested three *rIII* missense alleles isolated by UV mutagenesis: G24D, H42R, and A70V (67) (Figure 2.4A). In the absence of RI, two of the alleles, G24D and H42R, exhibited a slight but reproducible LIN defect, although the phenotype was subtle due to the relatively small effect of the parental *rIII* allele under these conditions (Figure 2.3C). Co-induction of these *rIII* with *rI*, however, resulted in distinct intermediate LIN defects, with lysis times ranging from 40 min~60 min, indicating these are partially defective alleles, at least in the λ context (Figure 2.3C). The lysis blockage was *t*-specific, as indicated by the fact that isogenic experiments with a λ holin allele, S_{A52G}, which has an early lysis phenotype that matches the normal *t* lysis time (26), did not show lysis delays in inductions of pZE12RIII$_s$, pZE12RI or pZE12RI-RIII (Figure 2.3B). Taken together, these data indicated that RIII has T-specific antiholin activity.

RIII binds to the cytoplasmic N-terminus of T

Next we addressed the molecular basis of RIII participation in LIN. The simplest hypothesis is, like other antiholins, including RI, RIII affects holin triggering by directly binding to holin T and blocking hole formation in the IM. Since RIII has no membrane or export signals (Figure 2.4A), the only possible target for RIII is the N-terminal cytoplasmic domain of T (nT), which has 34aa, and is required for holin function (56). To test this idea, we used the bacterial two hybrid system, based on intragenic complementation of CyaA function (101). We fused the nT, wt RIII, and four RIII mutant allele sequences to various combinations of the T25 and T18 fragments of CyaA. As shown in Figure 2.5A, this system revealed a strong self-interaction of RIII *in vivo*, which was abolished in the G24D allele, partially affected by H42R, and unaffected by L43Q or R75C. The transformants carrying plasmids expressing T18-RIII and T25-nT resulted in light but reproducible signals (Figure 2.5B), suggesting a relatively weak interaction between RIII and nT. Significantly, none of the four RIII mutant fusions retained the nT-binding ability (Figure 2.5B). These results correlate with the liquid culture lysis results (Figure 2.3C) and indicate the nT-RIII interaction is affected by the changes in the lysis-defective alleles.

To address the nT-RIII interaction in vitro, we constructed versions of nT and RIII tagged at the N-terminus with the His6-Sumo moiety (See Materials and Methods). After induction in a T7-based over-expression system, both His6-Sumo-nT and His6-Sumo-RIII accumulated as soluble forms (Figure 2.6, top panel, lanes 2-5). To detect complexes formed *in vitro*, the SUMO-tagged nT and RIII proteins were bound to

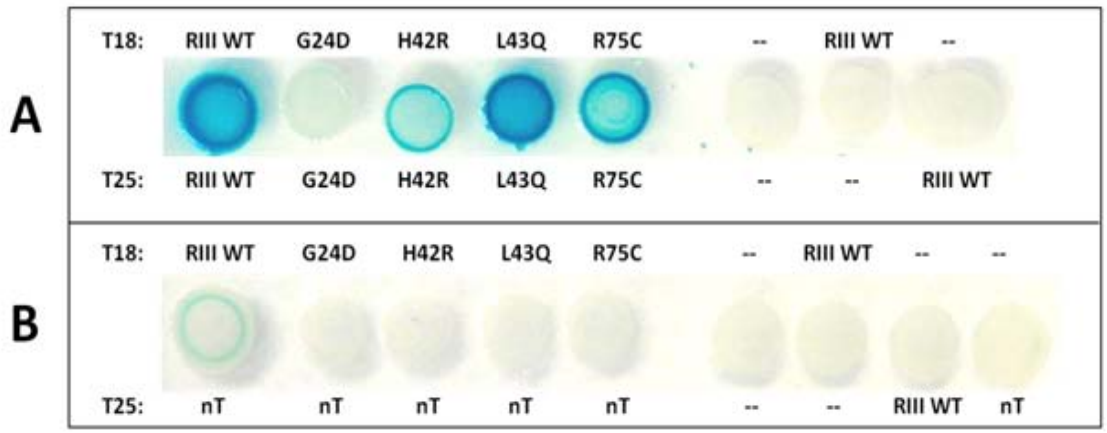

Figure 2.5 Bacterial two hybrid assay.
Bacterial two-hybrid results showing self-interaction of RIII
(A) and interaction of RIII and N-terminus of T (nT) **(B).** T18, protein fused to T18 fragment of CyaA protein; T25, protein fused to T25 fragment of CyaA. Negative control (--) indicates T18 or T25 fragments without RIII or nT fusion.

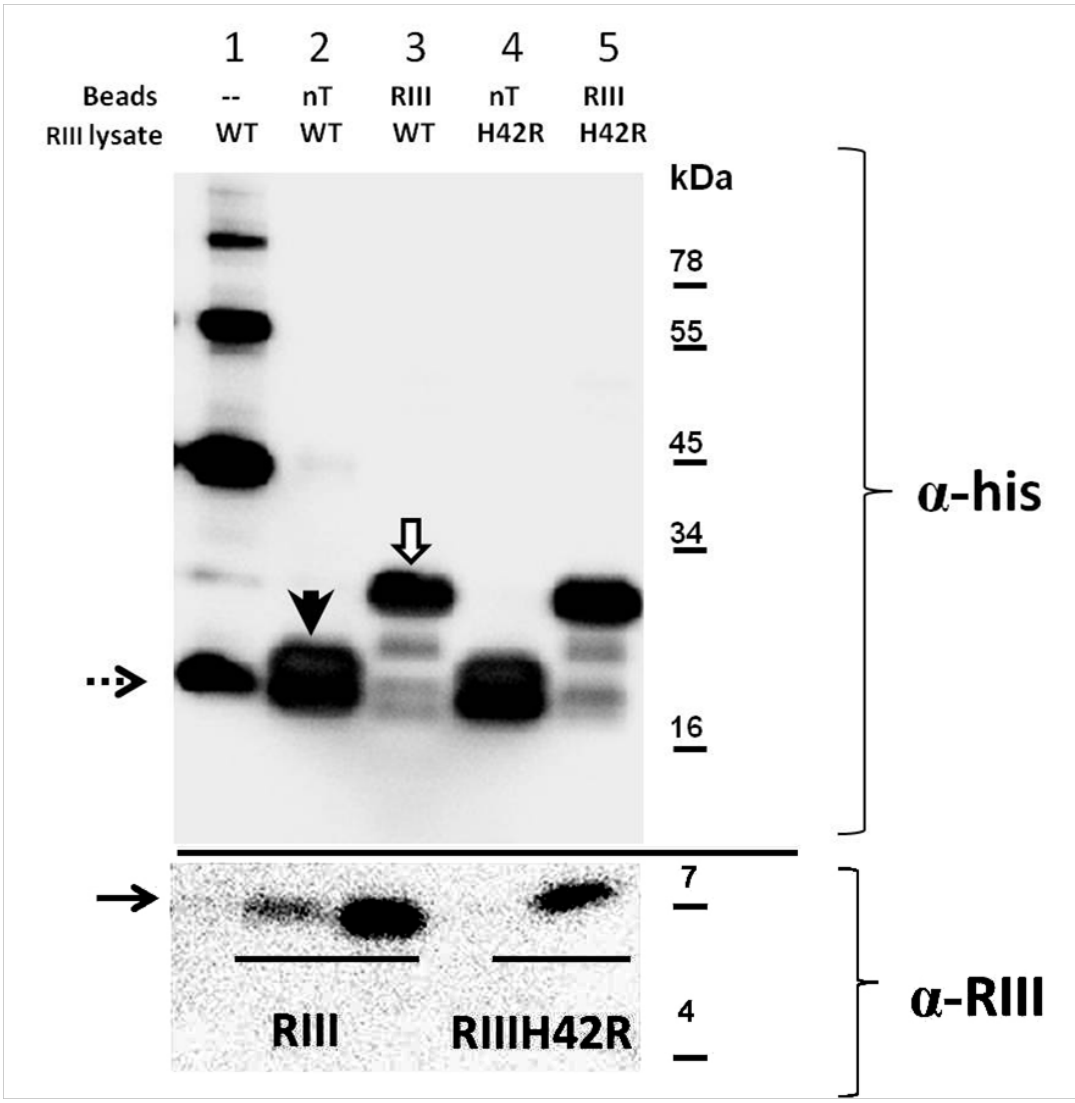

Figure 2.6 Pull-down assay using magnetic beads.
I*n vitro* interaction between nT and RIII. His-Sumo-tagged nT or RIII was bound to α-his Dynabeads, and RIII protein pulled down by Dynabeads was analyzed by Western blotting. His-sumo tag only (lane 1, dash black arrow), His-sumo nT (lane 2 and 4, black arrow head), His-sumo RIII (lane 3 and 5, white arrow) are shown in the upper panel as the result of western blotting using α-his antibody. RIII protein (solid black arrow) is visualized in the bottom panel using α-RIII antibody.

magnetic beads, mixed with cell lysates containing wt RIII or mutant $RIII_{H42R}$ protein, fractionated as bound and unbound, and then analyzed by immunoblotting. The results showed that both wt RIII and mutant $RIII_{H42R}$ proteins form complexes with His6-Sumo-RIII, but only wt RIII complexes with His6-Sumo-nT (Figure 2.6, bottom panel). The simplest interpretation is that the H42R mutation completely abrogates the RIII-nT interface but not the RIII homo-oligomerization interface, which is consistent with the results of the bacterial two-hybrid experiments.

The cytoplasmic N-terminal domain of T can block lysis inhibition in an RIII-specific manner

The finding that RIII binds nT in vitro and in *E. coli* in the context of the two-hybrids suggested that the *r* phenotype could be imposed in vivo by titrating the RIII produced in a T4 infection with the Sumo-tagged nT derivative. To test this idea, we plated T4 on lawns of cells induced for the over-expression of His6-Sumo-nT; under these conditions, T4 wt generated plaques distinctly larger and cleared compared to those generated on the isogenic control strain expressing the His6-Sumo tag (Figure 2.7). Neither T4*rIII* nor T4Δ*rIII* plaque morphology was affected by overexpression of nT (Figure 2.7).

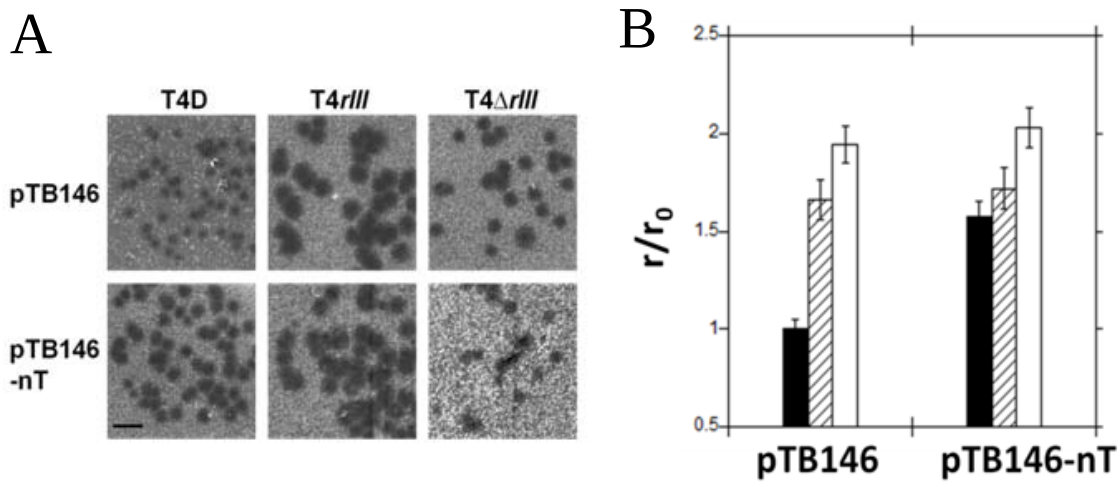

Figure 2.7 Rescue of *r* plaque morphology by overexpression of nT.
(A) T4 phages were plated on lawns of *E. coli* BL21(DE3) *fhuA*::Tn*10* carrying control plasmid expressing His-sumo-nT (pTB146-nT, bottom panels) or His-sumo (pTB146, top panels, neg control). Black bar represents 2.5mm.
(B) Quantification of plaque sizes were shown as the ratio of the average plaque radius (r) to the ratio of T4D plaques plated on pTB146 (r_0). Black bar, T4D; Patterned bar, T4Δ*rIII*; White bar, T4*rIII*.

Discussion

Among the Caudovirales, the lysis timing effected by the holin defines the length and fecundity of the phage infection cycle. Mutational analysis has shown that holin-mediated lysis timing can be drastically altered by single missense changes (56, 60, 104, 105), leading to the suggestion that this extreme mutational sensitivity is an evolutionary fitness factor, allowing phages to mutate rapidly to a radically different length of life cycle in response to altered environmental conditions (19). Despite the implied correlation between lysis timing and the environment, the T4 LIN phenomenon remains the only documented example of real-time regulation of lysis timing. Genetic analysis has shown that mutations in two of the classic T4 plaque-morphology loci, *rI* and *rV*, the latter allelic to the *t* holin gene, confer an absolute defect in LIN. Our work had shown that RI is a secreted protein that is initially synthesized as periplasmic protein tethered to the membrane with an N-terminal signal-anchor-release (SAR) domain (35). The presence of the SAR domain allows it to release into the periplasm and also confers extreme proteolytic instability on RI. Over-expression of the wt *rI* gene was shown to impose a delay on T-holin triggering in the λ context. A chimeric *rI* gene in which the SAR domain was replaced by a cleavable signal sequence generated a proteolytically stable periplasmic RI and, expressed in trans to *t*, imposed a stable LIN state. A model has been proposed in which an unknown LIN signal is generated by a super-infecting T4 virion. Under these conditions, it is suggested that the T4 Spackle and Imm proteins force the superinfecting virion to eject its capsid contents ectopically into the host periplasm (106). Some component of these virion contents, which include both the T4

genomic DNA and ~1000 protein molecules (64, 102), acts as a signal to stabilize the periplasmic RI protein. In this model, RI accumulates and binds to the periplasmic domain of the T4 holin T in a manner that T triggering is inhibited.(35). Significant progress has been made on the RI-T interaction. The soluble domain of RI, sRI, has been purified and shown to be largely alpha-helical in structure (57). In addition, the sRI molecule was able to bind the soluble domain of T, sT (55, 57), and prevent it from aggregation. Crystal structures of sRI and the sRI:sT complex have been determined (55). However, major gaps remained in our understanding of the LIN phenomenon. First of all, the signal provided by the superinfecting phage is completely unknown. In addition, the possible role of other *r* loci, most notably *rIII*, was not reflected in the model.

In this study we have shown that *rIII* is also unambiguously required for LIN on both *E. coli* K-12 and B hosts, resolving a long-standing controversy (64, 67). Moreover, we have shown the *rIII* gene expressed *in trans* to the T4 holin gene *t* can effect a small but reproducible delay lysis in a T-specific manner. In addition, expression of *rIII* significantly stabilized the LIN state imposed by over-expression of wt *rI*, which otherwise imposes a lysis delay that collapses after ~45 min. Since RIII is a cytoplasmic protein, the simplest notion is that RIII acts by binding to the short cytoplasmic domain of T, nT. Evidence supporting this was obtained from bacterial two-hybrid analysis and pull-down assays, which revealed a specific interaction between nT and the full-length RIII polypeptide. Importantly, known dysfunctional *rIII* missense mutations caused a defect in the RIII-mediated stabilization of RI-LIN. Finally, bacterial two-hybrid

evidence was provided showing that RIII has dimerizing or oligomerizing propensity, which may be functionally important in view of the fact that one of the known *rIII* defective missense mutations abrogates the response.

Taken together, these results indicate that both RI and RIII are, strictly defined, specific antiholins of the T4 holin T, and suggest an expansion of our previous model to include inhibitory interactions on both sides of the cytoplasmic membrane (Figure 2.8). In this scenario, RI acts as the LIN master regulator by receiving the signal generated by the super-infecting virion. Stabilization of RI leads to the formation of RI-T complexes that prevent the T protein from participating in the holin triggering pathway. The available evidence indicates that holin triggering occurs when the holin reaches a critical two-dimensional concentration and forms large oligomers, or rafts, within which the lethal holes are formed (19, 23). The simplest notion is that RI may simply block homo-oligomerization of T and thus T-triggering, which is consistent with the ability of sRI to prevent aggregation of sT (70). In our new model, we suggest that RIII participates in LIN by stabilizing the RI-T complexes. Indeed, the sRI:sT crystal structures were in the form of sT:sRI:sRI:sT hetero-tetramers (55). Thus an attractive notion is that in the onset of LIN, T-RI-RI-T heterotetramers are formed providing a symmetric binding site for RIII dimers to bind to the cytoplasmic nT domains (Figure 2.8). It should be noted that, in this perspective, RIII is the first example of an antiholin with no secretory or membrane signal, and also that the RI-RIII combination is the first example of a multiple-antiholin system. Since stabilizing RI by removal of the SAR domain can lead

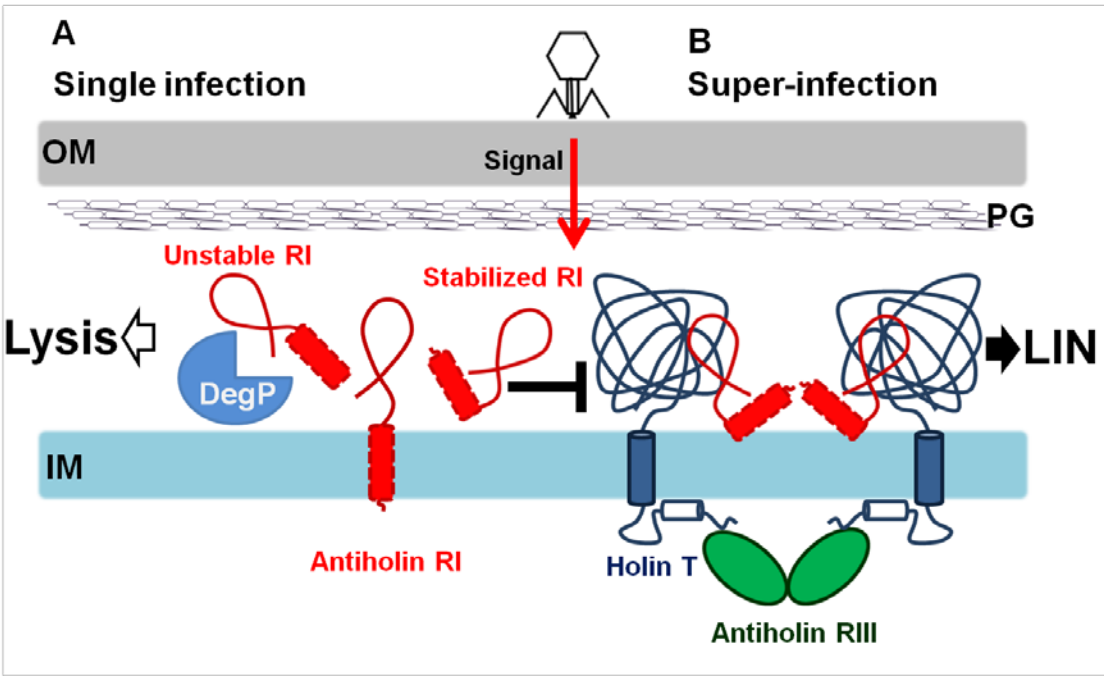

Figure 2.8 The current model of LIN involving two antiholins.
Both RI and RIII are required for the stable LIN.
(A) In a single phage infection, antiholin RI will be degraded by periplasmic protease DegP after spontaneous release into the periplasm. Cell lysis occurs at ~ 25 min.
(B) In a superinfection, the DNA of a superinfecting T4 phage will be ectopically ejected into the periplasm, generating the "signal" to stabilize the periplasmic antiholin RI. This leads to accumulation of RI, which then binds the periplasmic domain of T, in a T-RI-RI-T heterotetramer. This facilitates the binding of cytoplasmic antiholin RIII to the N-terminus of T. This unique, sandwich-like structure spanning two cell compartments robustly blocks participation of T in hole-formation.

to stable LIN without the participation of RIII, we propose to designate RI as the antiholin with RIII as a co-antiholin.

As noted above, major gaps remain in our understanding of the T4 LIN phenomenon, which deserves attention not only because of its historical status, but as a richly-documented phenomenon that may be important in our understanding of phage propagation in liquid culture and in environmental scenarios that may be relevant to phage-based the rapeutics . Immediate future efforts will be directed at determining the nature of the RIII-nT interaction at the structural level.

CHAPTER III

BIOCHEMICAL AND STRUCTURAL CHARACTERIZATION OF RIII

Introduction

For double-stranded DNA phages infecting Gram-negative hosts, the length of phage infection cycle is determined by the phage-encoded holin, a small integral membrane protein (17, 27) . Holins accumulate harmlessly in the inner membrane (IM) until suddenly forming lethal membrane lesions, or holes, at an allele-specific time. This event, which is defined as "triggering" (17), causes the release of the phage-encoded endolysin into the periplasm, resulting in the degradation of the cell wall and lysis. In addition to the holin, some phages encode another protein, an antiholin, to regulate the lysis timing by inhibiting the holin through direct protein-protein interaction. Unlike the well-studied holin-antiholin system of phage λ S105 and S107 (28, 33), where the main function of antiholin seems to be to prevent premature lysis, phage T4 uses two antiholins, the proteins RI (34, 35, 70) and RIII, to delay lysis and does so under specific physiological condition; i.e., when the infected cells are superinfected by other T4 or other T-even phages (Figure 3.1A) (64, 67, 107). The prolongation of infection cycle can last for several hours and result in the intracellular hyper-accumulation of T4 progeny virions. The lysis-inhibited state, designated as LIN (Lysis Inhibition), can be viewed as a signal transduction event, with the signal being the super-infection event, indicating a deficit of available hosts in the environment, and the response being an inhibition of holin-triggering (56) . T4 LIN is the only experimentally-determined

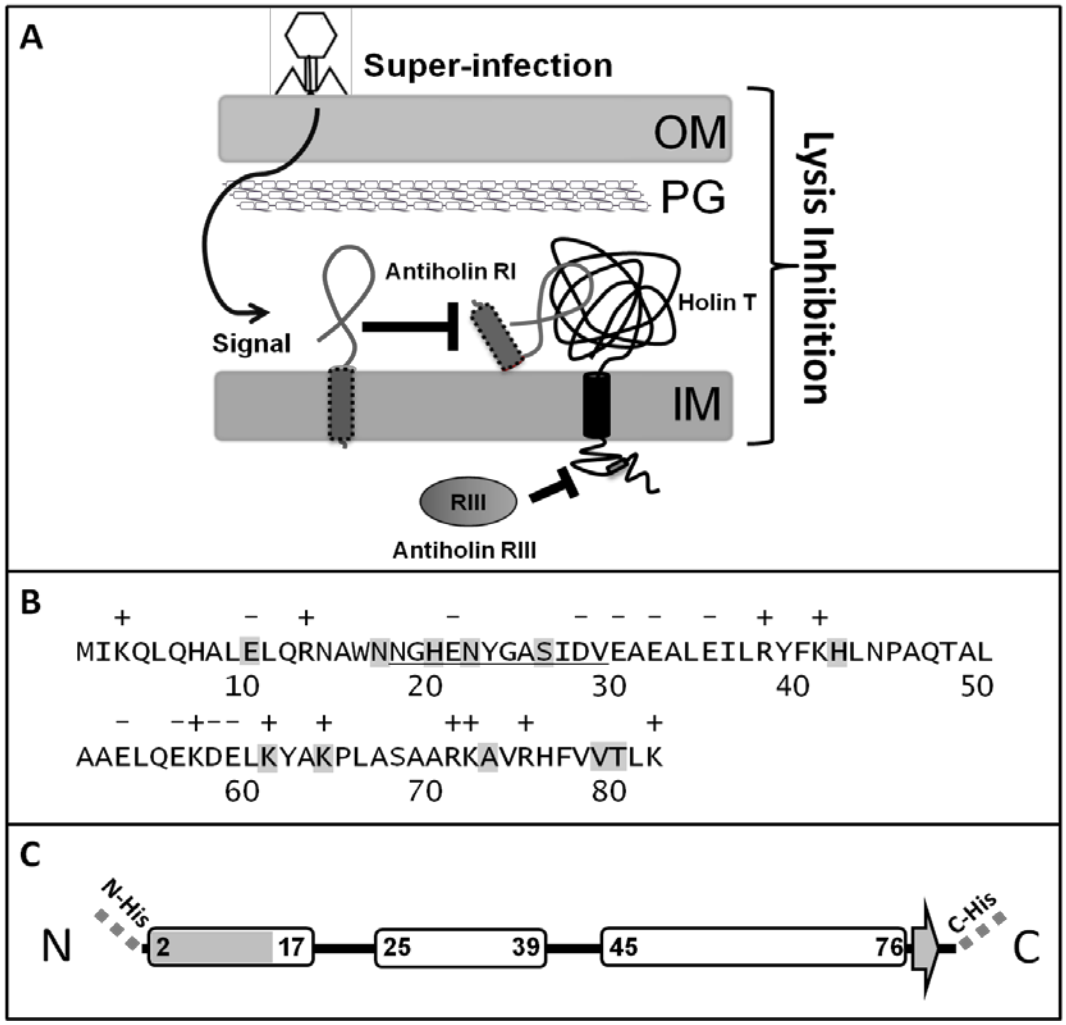

Figure 3.1 Model for LIN involving RI and RIII.
(A) The model of LIN involving two antiholins. Both RI and RIII are required for the stable LIN. The main antiholin RI is stabilized by the LIN signal before binding to the C-terminus of holin T. The cytoplasmic antiholin RIII then binds to the N-terminus of T, forms a stable inhibitory complex to stop holin T from triggering.
(B) The primary structure of RIII is shown with residue charges above the sequence. Conserved residues are highlighted. The epitope region for the α-RIII antibody is underlined.
(C) Secondary structure of RIII predicted by Jpred (108) . RIII is predicted to be 77% α-helix including 17% of coiled-coil helix, and 3.6% of β-strand. Helical structures are shown as rectangles, with the N-terminal coiled-coil helix region shaded; sheets are shown as grey arrows; unstructured regions are indicated by solid lines. The N-terminal (N-His) or C-terminal (C-His) his-tag is indicated by grey dashed lines.

example of a phage adjusting the length of the infection cycle in response to an environmental condition (67).

The holins that have been studied experimentally consist of two or more transmembrane domains (TMDs) separated by short linkers and ending in a C-terminal tail rich in basic residues (17, 19). The unique capacity of the T4 holin T (product of gene *t*) for LIN is reflected in its unique primary structure: a small (34 aa) N-terminal cytoplasmic domain, a single transmembrane domain (TMD), and a large, globular periplasmic C-terminal domain (56, 59). The periplasmic domain of T is crucial for both its hole-formation function and its interaction with the main antiholin (34, 70, 109), RI, the product of gene *rI*, one of the T4 *r* (rapid lysis) loci , which played a significant role as a model genetic system during the development of modern molecular genetics more than seven decades ago (67). RI has a SAR domain (signal-anchor release), which causes it to be exported to the periplasm in a membrane-tethered form and then released from the membrane as a soluble periplasmic protein (35). Formation of a complex between the soluble form of RI and the periplasmic domain of T is necessary for the imposition of LIN (35, 57). Recently we showed that in addition to RI, the product of one of the other classic *r* genes, *rIII*, is required for the stable maintenance of the LIN state (Chapter II).

Since RIII is a polypeptide of 82 amino acids (aa) lacking any membrane-localization or export signal (Figure 3.1B), its likely target for directly effecting lysis would be the short N-terminal cytoplasmic domain of T. Indeed, our recent results demonstrated that RIII binds to the N-terminus of T (Chapter II). This makes RIII the

first antiholin with no membrane domain. Here we report the results of efforts to purify RIII and characterize its structure. The results are discussed in terms of a model for the RI-T-RIII complex and its role in LIN.

Materials and methods

Bacterial growth and induction

See Table 3.1 for the list of phages and bacterial strains used in this study. Bacterial strains were plated on standard LB-agar plate supplemented with antibiotics (ampicillin,100 mg mL^{-1}; chloramphenicol, 10 mg mL^{-1}; kanamycin, 40 mg mL^{-1}) when appropriate, as described (50, 57). Preparation of cultures, induction of thermoinducible λ lysogens, and induction of strains harboring pET11a over-expression plasmids were performed as described (23, 56, 57). Bacterial growth and lysis were monitored using a sipping Gilford Stasar III Spectrophotometer (Nova biotech, El Cajon, CA) (34, 56).

DNA manipulations, sequencing, and plasmid construction

All plasmids used and generated in this study are listed in Table 3.1. Isolation of plasmid DNA, designing of primers, DNA amplification by polymerase chain reaction (PCR), DNA transformation, and DNA sequencing were performed as previously described (50, 57, 70, 110). Restriction and DNA-modifying enzymes were purchased from New England Biolabs (Ipswich, MA). Manufacturer's instructions were followed when performing restriction digestion, ligation and all other reactions using New England Biolabs enzymes. The DNA sequence of all constructs was verified by Eton Bioscience (San Diego, CA) plasmid sequencing service. The RNA stem-loop structure was analyzed with UNAFold (111). Plasmid pZE12-RIIIo, pZE12-RIIIs, pET11a-RIII

Table 3.1 Phages, strains, and plasmids used in this study

Phages	**Description**	**Source**
T4wt	Bacteriophage T4D	Laboratory Stock
T4Δ*rI*	Complete deletion of *rI* from nt 59204 to nt 59496 in T4D genome	(35)
T4Δ*rIII*	Complete deletion of *rIII* from nt 130779 to nt 131042 in T4D genome	This study
λΔSR	cI857 ΔSR *bor::Kan*	Laboratory Stock
Bacteria Strains	**Description**	**Source**
BL21(DE3) *fhuA*::Tn*10*	*E. coli* B *ompT* r_B^- m_B^- (*P*$_{lac}$*UV5::*T7 *gene1*) *slyD::Kan* *fhuA*::Tn*10*	Laboratory Stock
B834	*E. coli* B *ompT* r_B^- m_B^- *met*$^-$	Laboratory Stock
MDS12 *tonA*::Tn*10* *lac*I^{q1}	MG1655 with 12 deletions, totaling 376,180 nt including cryptic prophages	(35)
MQλΔSR	MC4100 lacI^q λΔSR lysogen	This study
Plasmids	**Description**	**Source**
pZE12	ColE1 origin; $P_{LlacO-1}$ (PL promoter with three lacO operators); AmpR	(97)
pZE12-luc	Luciferase gene luc cloned under $P_{LlacO-1}$	(97)
pZE12RIII$_o$	T4 *rIII* cloned under $P_{LlacO-1}$ with native SD	(34)
pZE12RIII$_s$	T4 *rIII* cloned under $P_{LlacO-1}$ with plasmid SD	This study
pZA32RIII$_s$	p15A origin; $P_{LlacO-1}$ (PL promoter with three lacO operators); Carries same *rIII* insertion as pZE12RIII$_s$. CamR	This study
pET11a-RIII	pBR322 origin, T7 promoter, carrying codon 1-82 of *rIII*	This study
pET-HisRIII	His6 tag inserted to the N-terminus of *rIII*	This study
pET-RIIIHis	His6 tag inserted to the C-terminus of *rIII*	This study
pT4T(R-RzRz1)am	pBR322 derivative carrying late promoter and lysis cassette of λ with *S* gene replaced by T4 *t* and all other lysis genes carry amber mutation. AmpR.	(53)

Table 3.1 Continued

Primers	Sequence	Source
RIII N-his F	CACCATCACGGAGGTATTAAACAATTACAACACGCTC	This study
RIII N-his R	ATGGTGATGACCTCCCATTTAAAAATTCTCGTTAG	This study
RIII C-his F	CACCATCACGGAGGTTAATTTATTGGAGATTCACTG	This study
RIII C-his R	ATGGTGATGACCTCCCTTCAGTGTTACCACAAAGTGAC	This study

were constructed as described (Chapter II). pET-HisRIII and pET-RIIIHis were created by the insertion of hexahistidine codons (CATCATCACCATCACCAC) before the start codon (pET-HisRIII)or the stop codon (pET-RIIIHis) of the wild-type *rIII* allele in the pET11a-RIII using primers listed in Table 3.1.

TCA precipitation, subcellular fractionation, SDS-PAGE and Western blotting

Trichloroacetic acid (TCA) precipitation and subcellular fractionation was performed as previously described (57). Briefly, cells were collected and concentrated by centrifugation, and then disrupted in a small SML-Aminco French pressure cell (Spectronic Instruments, Rochester, NY) at 4 °C. The cell lysates were then centrifuged at 100,000 x g for 1 h in a Beckman TLA100.3 rotor to separate the soluble and membrane fractions. Equivalent amounts of each samples were precipitated by ice-cold TCA (10% final volume) and loaded onto SDS-PAGE gels and analyzed by Coomassie blue staining or Western blotting using rabbit polyclonal α-RIII antibody as the primary antibody. The epitope for a-RIII antibody is shown in Figure 3.1B. For western blotting, the chemiluminescence signal was detected using a Bio-Rad Chemidoc XRS Imager (Bio-Rad Laboratories, Hercules, CA). Images were obtained and analyzed by Quantity One 1-D Analysis Software (Bio-Rad Laboratories, Philadelphia, PA) and ImageJ (NIH, Bethesda, MD).

Testing stability of RIII protein

The stability of RIII protein was tested as described (35). Briefly, the cultures of MDS12 *tonA*::Tn*10 lacI*^{q1} strains carrying pZE12-RIIIs or pZE12-luc (negative control) were grown to A_{550} ~ 0.4 at 37°C in 20 mL of LB supplemented with ampicillin, and

then divided into two equal aliquots. Each aliquot was diluted into 30 mL with LB-Amp, and then grown in 37°C again until A_{550} ~ 0.4 and induced with IPTG, as described above. After aerating for 30 min, chloramphenicol was added into each culture with a final concentration of 300 μg mL^{-1} to stop the protein synthesis (time =0). 5 mL TCA precipitation samples were taken from each cultures at 5 min intervals until time = 20 min. RIII protein was visualized by Western Blotting using a-RIII antibody and analyzed by ImageJ (112) .

Protein purification

RIII protein was purified using cation-exchange and gel filtration column chromatography. ÄKTA FPLC system, resins, gel filtration columns were purchased from GE Healthcare Bio-science (Marlborough, MA). Rotors were purchased from Beckman Coulter (Brea, CA). Four liter cultures of BL21(DE3) *fhuA*::Tn*10* carrying the plasmid pET11a-RIII were grown in LB at 37°C until A_{550}=0.6, immediately cooled on ice for 30 min, supplemented with 1 mM IPTG and aerated for 24 h at 16°C . The induced cells were harvested by centrifugation in a Beckman JA-10 rotor at 4000 rpm for 45 min at 4°C. Cell pellets were resuspended in 5 mL Buffer A (50 m*M* Tris, 0.5 m*M* EDTA, pH7.0), supplemented with Protease Inhibitor Cocktail (Sigma, St. Louis) and 100 mg mL^{-1} final concentrations of DNase I from New England Biolabs (Ipswich, MA) and RNase I from Thermo Fisher Scientific (Waltham, MA). Cells were disrupted in a French pressure cells and the resulting cell lysates were cleared by centrifugation at 100,000 x g for 1 h using Beckman TLA100.3 (57) . The supernatant was then rapidly filtered through a 0.22 mm syringe filter and passed over SP Sepharose Fast Flow resin

(GE Healthcare, Marlborough, MA) calibrated with Buffer A and packed in 24 mL bed volume column on ÄKTA FPLC system. After the column was washed with 15 column volumes of Buffer A, RIII protein was eluted using Buffer B (50 m*M* Tris, 0.5 m*M* EDTA, 2 *M* NaCl, pH7.0). Elution fractions were pooled and concentrated to 500 μL final volume with a Vivaspin Turbo 15 MWCO 3000 (Sartorius Corporation, Bohemia,NY) and was loaded onto a Superdex200 10/300 GL column which is pre-equilibrated at room temperature with S200 buffer (50 m*M* Tris, 0.5m*M* EDTA, 500 m*M* NaCl, pH7.0). SDS-PAGE, Coomassie blue staining and western blotting were used for assessing the peak A280 fractions. Concentration of purified proteins was determined by Bradford assay using calculated extinction coefficients, or by amino acid analysis service performed by the Protein Chemistry Laboratory, Texas A&M University.

CD spectroscopy

Circular dichroism spectra data were obtained using a Jasco J-815 CD spectrometer (Jasco, Eaton, MD). The RIII protein was eluted from Superdex200 in CD buffer (50m*M* Tris, 0.5m*M* EDTA, 150m*M* NaCl; pH7.0), and concentrated using a Vivaspin Turbo 15 MWCO 3000 to 5m*M*. Protein concentrations of RIII were determined by amino acid analysis. 300 μL samples were loaded into 1.0 mm quartz cuvettes and equilibrated to 25°C for 3 min before being scanned. Scans from 250 to 190 nm with a 1 nm data pitch and 50nm per min scanning speed were applied to all samples. To estimate the protein secondary structure, K2D3 server (113) was used to fit all CDs data from 190 to 240 nm that were converted to differential absorption units (Δε), as described (57).

Interaction between T and RIII

The cytoplasmic membrane vesicles were prepared from MQλΔSR carrying pT4T(R-RzRz1)am plasmid, which is a pBR322 derivative carrying late promoter and lysis cassette of λ with the holin *s* gene replaced by T4 *t* (35), but have amber mutation on all other lysis genes (endolysin R, spanin Rz and Rz1). A previously described protocol was adopted for preparing the cytoplasmic invert vesicles (114) with modification. 15 mL cultures were grown until A_{550} ~ 0.3 (~1.1 10^8 cells mL^{-1}) at 30°C, induced by shifting to 42°C water bath shaker for 15 min and then were kept in 37°C for 1 h. Cells were harvested by centrifugation at 1500 x g then washed twice with 50m*M* Tris-HCl buffer, pH=7.0. Bacteria were resuspeded in 2 mL and then passed through a small French Press cell for two times. Unlysed cells were collected by a low-speed centrifugation. Then, 500 μL of the resulting supernatants were mixed with purified RIII (1μ*M* final concentration) and incubated at room temperature for 10 min. The mixtures were then centrifuged for 1 h at 100,000 x g in a Beckman TLA100.3 rotor to pellet the membrane fractions. The resulting supernatants, membrane vesicle pellets, and the whole-cell control were resuspended in 50m*M* Tris-HCl buffer, pH=7.0, and precipitated by adding ice-cold TCA. Equivalent amount of each samples were resuspended in SLB and analyzed by SDS-PAGE and Western blotting.

Quantification of RIII in T4 infection

The amount of RIII protein expressed during a T4 infection was quantified as described (57) with modification. Cultures of *E. coli* B834 were grown in LB until A_{550} ~ 0.3, and then were infected with T4D, T4Δ*rI* or T4Δ*rIII* at a multiplicity of infection

(MOI) ~ 10. After 30 min of infection, 10 mL samples were collected and precipitated with ice-cold TCA. The samples were collected and resuspended in protein sample buffer with reducing reagent (β-mercaptoethanol). Purified RIII protein standards were prepared by diluting with S200 buffer, then 50, 25, 12.5, and 6.25 pmol of RIII protein were loaded in each well of an SDS-PAGE gel. All samples and standards were analyzed by SDS-PAGE gels and Western blotting. Resulting images were analyzed using the software ImageJ (NIH).

DSP Crosslinking

A p15A origin, IPTG inducible plasmid pZA32RIIIs was constructed by replacing the *luc* gene in plasmid pZA32-luc (97) with the T4 *rIII* gene. Dithiobis (succinimidyl propionate) (DSP) was obtained from Thermo Fisher (Thermo Scientific, Rockford, IL) and a previously described method for crosslinking experiment was followed (21). 30 mL of cultures of *E. coli* MQλΔSR harboring plasmids pZA32RIIIs and pT4T(R-RzRz1)am were pelleted by centrifugation after the induction. The cell pellets were washed with 10 mL PBS-Na^+ buffer (0.1 M Phosphate, 0.15 M NaCl, pH 7.2) for three times and then resuspended in 1 mL. DSP was prepared in DMSO and 80 μL of 10mM DSP solution was added to 320 μL of resuspended cells. The mixture was incubated at room temperature for 30 min with agitation. The reaction was stopped by adding 8ul of 1 M Tris pH7.5 buffer. 80 μL of DMSO without DSP was used for negative control. The cells were then precipitated by 10% TCA and resuspended in protein sample loading buffer, with or without reducing reagent (β-mercaptoethanol). The results were analyzed by SDS-PAGE and Western blotting, as described above.

Results

RIII accumulates as a soluble protein in the host cytoplasm

Our first studies directed at understanding the role of RIII in T4 LIN at a molecular level exploited a convenient inducible lysis system based on a λ prophage in which the T4 *t* holin gene replaced the native λ *S* gene and the RIII protein was supplied *in trans* from an inducible plasmid vector (Ramanculov, 2001). With this plasmid, in which the *rIII* gene was served by its native Shine-Dalgarno sequence (S-D) (Figure. 3.2A), no effect of *rIII* expression was observed on the lysis of induced λ-*t*. In Chapter II we showed that *rIII* expressed from an isogenic construct with a strong, artificial S-D, could temporarily and specifically block λ-*t* lysis. Moreover, this temporary RIII-mediated lysis inhibition could be stabilized by also supplying the main antiholin RI *in trans*. Using an antibody against an epitope of native RIII, we interrogated *rIII* expression levels under these conditions. The results showed that under the modified SD sequence improved RIII accumulation significantly (approximately 50-fold) (Figure. 3.2B). This increased expression may also be due to the disruption of an RNA secondary structure that is predicted to occlude the native SD (Figure. 3.2A). The constructs with the artificial SD sequences were designated as pRIIIs and used for all experiments described in this study.

The main antiholin RI is an unstable protein containing Signal-Anchor-Release domain (SAR) that allows RI to be spontaneously secreted into the periplasm (35). Unlike RI, RIII is predicted to be cytoplasmic protein without any TMD or signal sequence. Subcellular fractionation and Western blotting revealed that RIII does indeed

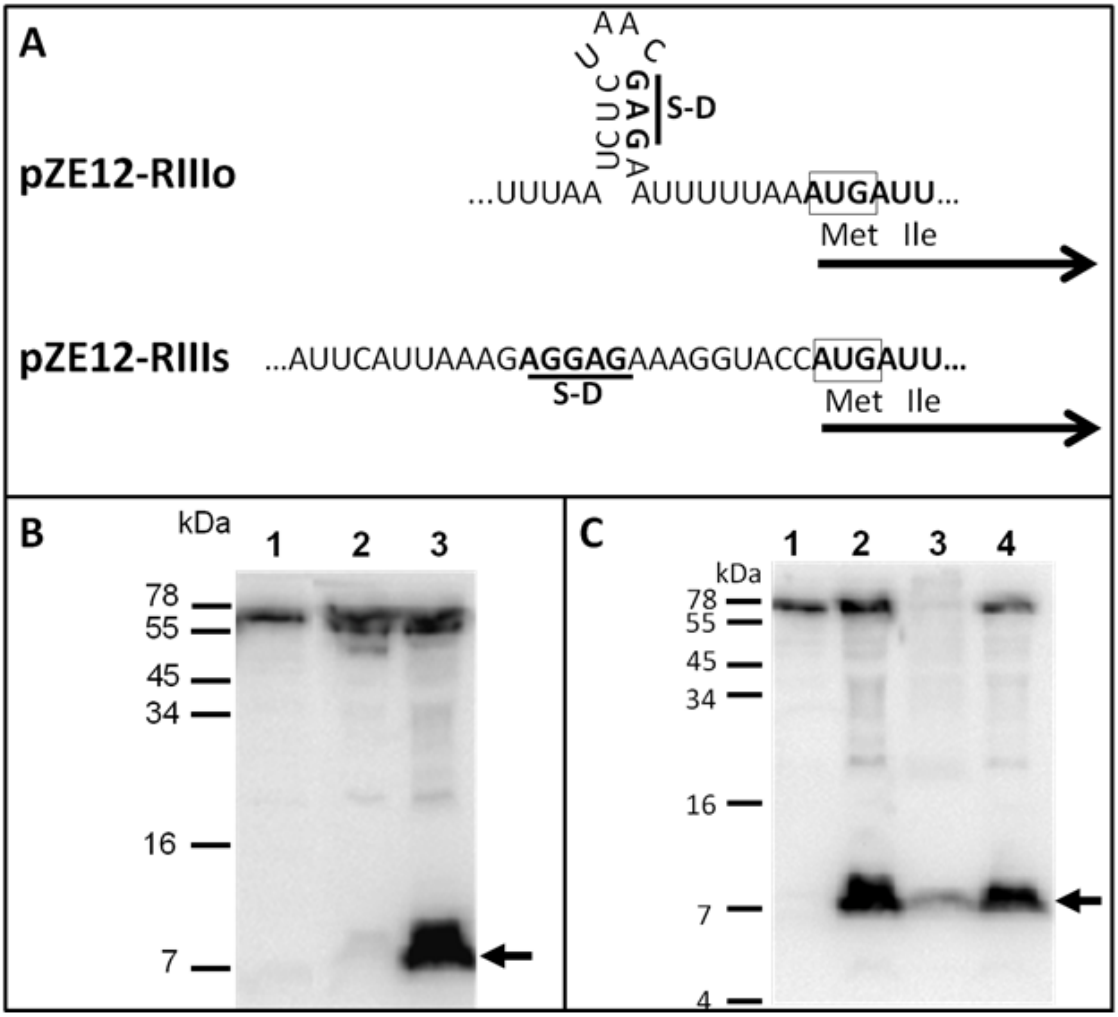

Figure 3.2 Expression level of RIII.

(A) The region upstream of the *rIII* start codon in the RNA transcripts of pZE12-RIIIo and pZE12-RIIIs are shown with predicted stem loops. The native (GAG; RIIIo construct) and artificial (AGGAG; RIIIs construct) S-D regions are labeled with solid lines. The start codon of *rIII* (AUG) is boxed. The direction of transcription is indicated with black arrows.

B) Western blotting results of RIII expressed under different S-D region. 1, pZE12-luc (negative control); 2, pZE12-RIIIo, 3, pZE12-RIIIs. Samples were collected after 1h of IPTG induction. The black arrow head indicates RIII protein on the blot. Each well was loaded with one A550 unit of cells.

(C) RIII is a soluble protein. MDS12 *tonA*::Tn*10 lac*I^{q1}cells carrying pZE12-RIIIs culture samples were collected after induced for 1h with 1mM IPTG at 37°C, then lysed byp French press. 1, Uninduced negative control; 2, Whole cell sample; 3, Cell pellet after centrifugation; 4, Supernatant. The black arrow indicates RIII protein on the blot. Wells loaded as in B.

accumulate as a cytosolic species (Figure 3.2C); moreover, RIII is stable, with no detectable proteolytic instability after >20 min of inhibition of protein synthesis by chloramphenicol (Figure 3.2D), compared to the 90 sec half-life of RI under the same conditions (6).

RIII is a predominantly α-helical monomer

As the first step in obtaining purified RIII protein, we constructed plasmids based on the hyper-expression vector pET11a carrying *rIII* alleles encoding wild type (wt; pET-RIII) RIII, an N-terminal His-tagged RIII (pET-HisRIII), and a C-terminal His-tagged RIII (pET-RIIIHis). Neither of the His-tagged RIII proteins accumulated as monomeric species in SDS-PAGE. Instead, for inductions of both pET-HisRIII and pET-RIIIHis, nearly all of detectable product accumulated in SDS-resistant oligomers of a wide range of molecular masses up to >200 kDa (Figure 3.3A). The predicted secondary structure of RIII consists of three alpha-helical domains spanning >75% of the sequence, with the N-terminal domain predicted to be a coiled-coil helix domain (Figure. 3.1C). The intolerance of RIII for either N- or C- terminal oligohistidine tags suggests that the coiled-coil N-terminal helix and the C-terminal extended segments are critical for tertiary structure. In contrast, the induction of pET11a-RIII resulted in the high-level accumulation of a ~10 kDa species, with an approximate yield of 11 mg L-1 after 2 h of induction, based on the intensity of the Coomassie-blue stained band (Figure. 3.3A). However, in this over-expression condition only ~50% of RIII was found in soluble fraction, indicating that a substantial fraction of the protein was not folded properly and thus accumulated in insoluble aggregates. To minimize misfolding, we aerated cultures

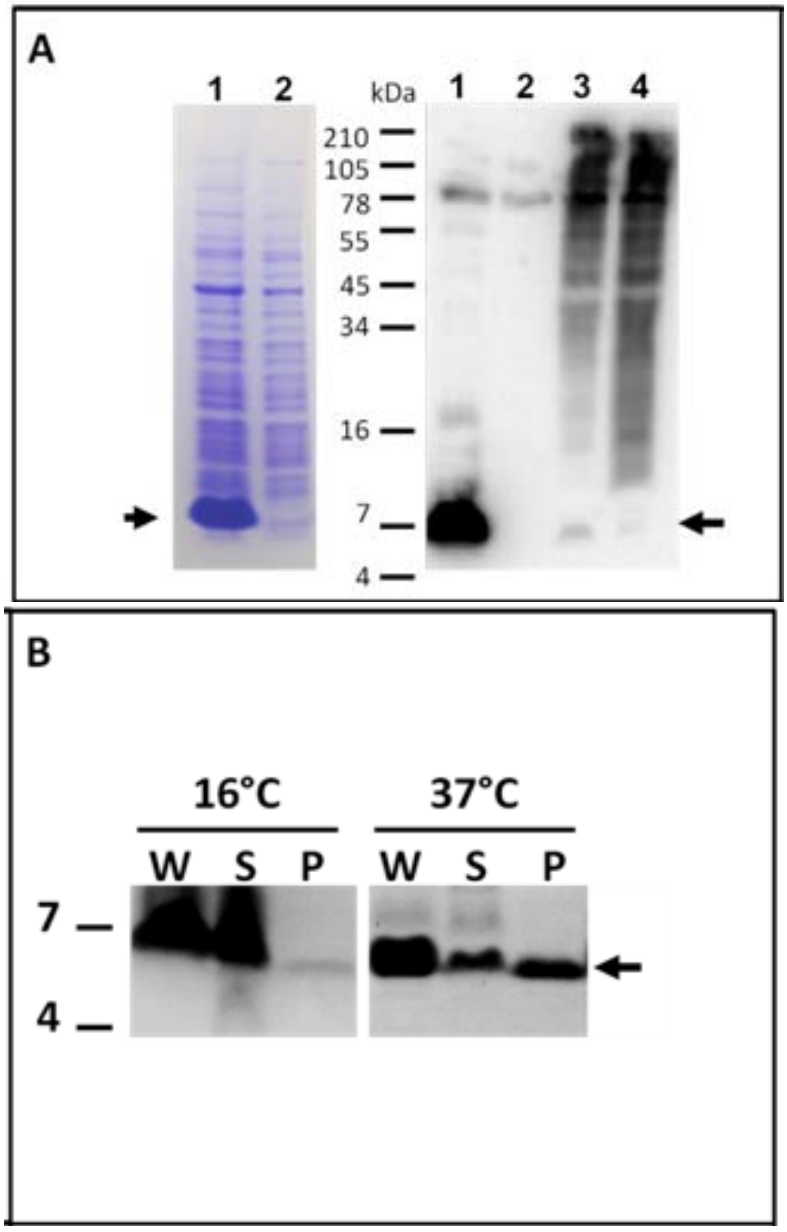

Figure 3.3 Solubility of RIII.
(A) His-tagged RIII did not accumulate in the solution. 1-4 shows the whole-cell TCA precipitated samples 1, pET11a-RIII; 2, pET11a vector control; 3, pET11a N-his RIII; 4, pET11a C-his RIII. The black arrow indicates the location of RIII monomer on the blot.
(B) Solubility of overexpressed RIII in different incubation temperature. pET11a-RIII plasmid was induced by 1mM of IPTG in BL21(DE3) *fhuA*::Tn*10* cells at 16°C or 37°C, as described in Materials and Methods. W, whole cell sample; S, supernatant sample; P, pellet sample. Loading was normalized to O.D. unit of 0.3 per each well.

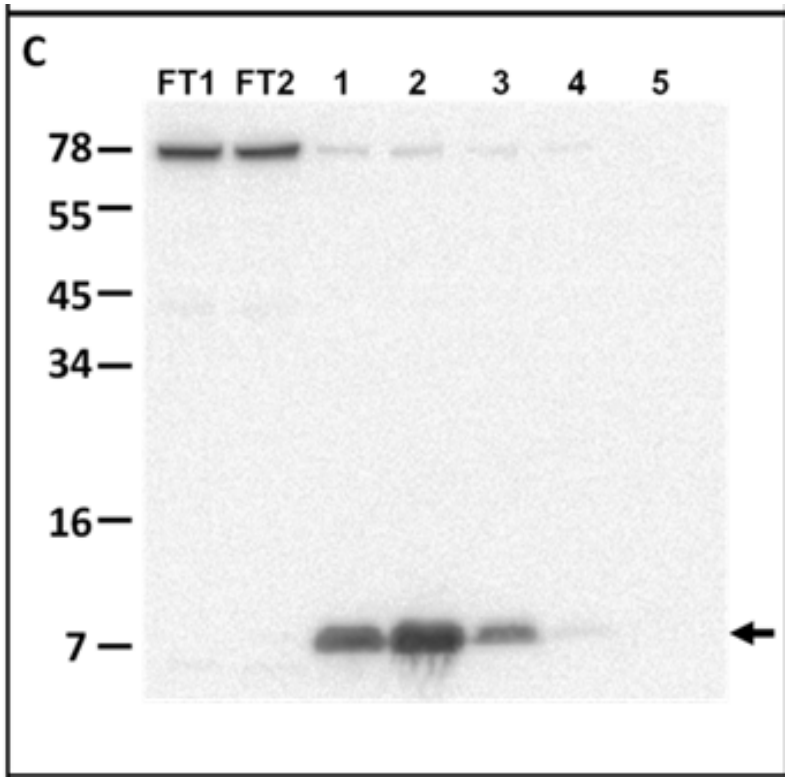

Figure 3.3 Continued.
(C) RIII protein after went through cation exchange column. 5 ul of each fraction was collected and mixed with 2XSLB and loaded on each well. FT1 and FT2, representative flow through fraction before the elution of RIII protein. 1-5, fraction of RIII eluted from SP Fast Flow column.

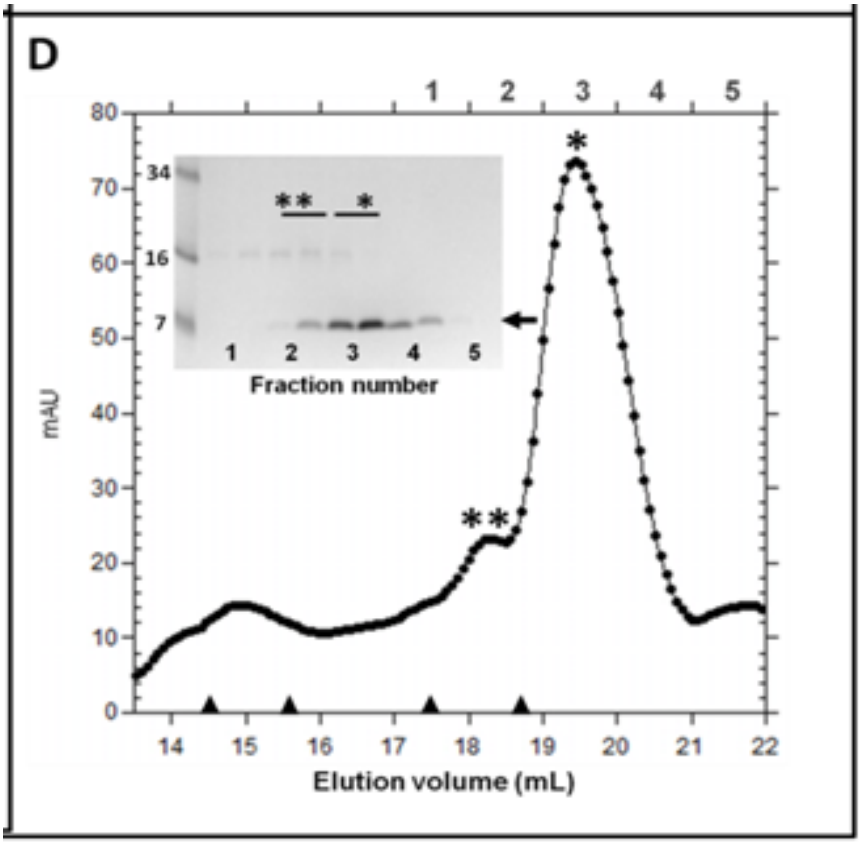

Figure 3.3 Continued.
(D) Superdex 200 gel filtration of RIII. Elution profile of sRI with major monomer peak and minor dimer peak are indicated as * and **, respectively. Protein standards are indicated by solid rectangles at the bottom, from left to right: 150, 66, 29, and 13.7 kDa. Inset shows the Coomassie blue stained SDS-PAGE gel of samples collected from fraction 1-5, as indicated on the top right. The black arrow indicates RIII monomer.

at 16°C overnight after IPTG induction. This resulted in ~ 10 mg/ L accumulation, based on Coomassie-blue staining, with >90% soluble (Figure 3.3B). RIII was then purified by cation exchange and gel filtration chromatography (Figure 3.3C and Figure 3.4), as described in Materials and Methods. The majority of the RIII protein eluted from a gel filtration column in a peak corresponding to 10 kDa, which is consistent with 9.3 kDa molecular mass predicted for a RIII monomer. A minor peak was observed at ~20 kDa, indicating a small fraction of purified RIII eluted as a dimer (Figure. 3.3D). To examine if the purified RIII was properly folded, we performed circular dichroism spectroscopy (CDs) to analyze the secondary structure of RIII (Figure. 3.5). The CD spectrum was consistent with ~ 95% of a-helix, which is higher in α-helix percentage than the secondary structure prediction made by Jpred (75%; Figure. 3.1C) (108) , CFSSP (88% a-helix) (115), or I-TASSER (80% a-helix) (116). Taken together, we conclude that RIII exists as a highly helical monomer in solution, with a slight tendency to dimerize.

Interaction between the T Holin and RIII

The antiholin RI was shown to form equimolar complexes with the C-terminal periplasmic domain of the T holin (55, 60, 70). Results of yeast two-hybrid assays and co-precipitation experiments with Sumo-nT hybrids indicated that RIII binds to the 34 aa N-terminal domain of T (nT), which is predicted to be cytoplasmic *in vivo* (Figure. 3.1A). However, direct interaction of RIII and T in the context of the full length, membrane-spanning T protein has not been demonstrated. We addressed the RIII-T interaction by DSP

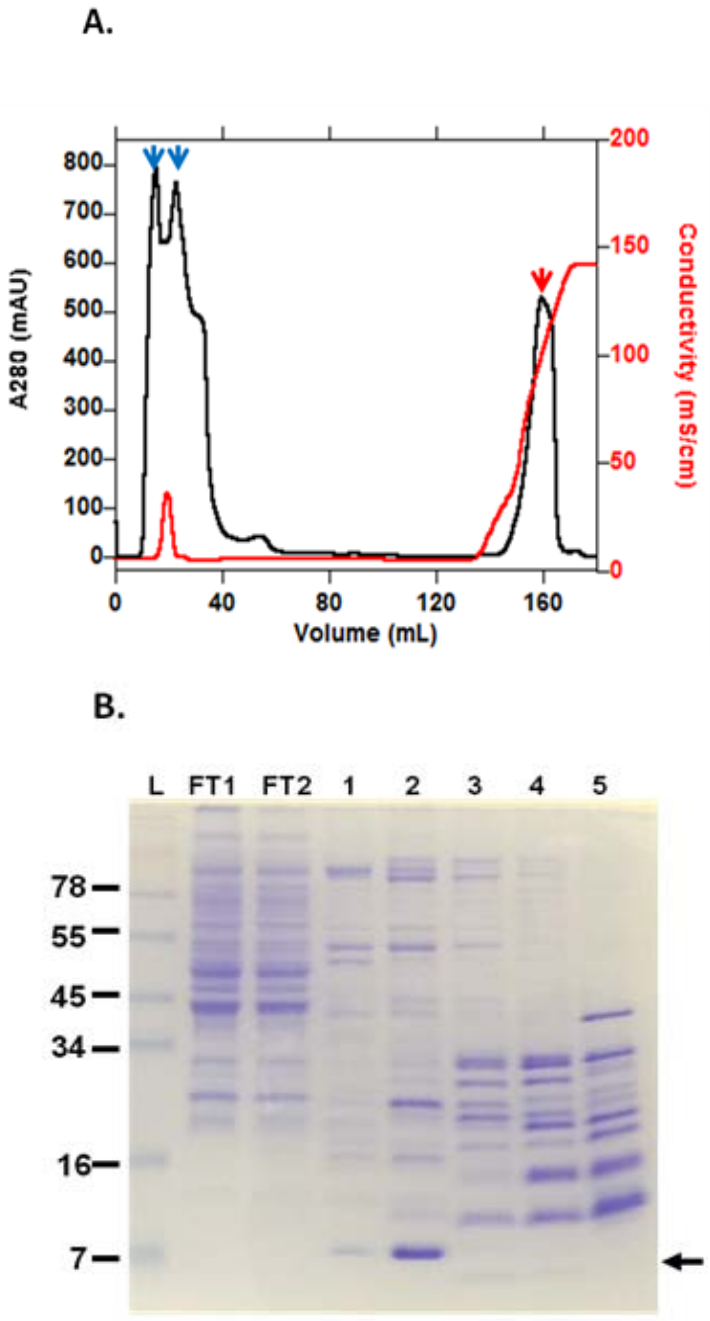

Figure 3.4 Purification of RIII.
(A):Cation exchange chromatography of RIII using SP Fast Flow resin. The A280 adsorption and Conductivity are shown as black and red line, respectively. Blue arrows indicate when the FT samples were taken. The red arrow shows when fractions 1-5 were taken. **(B)** Coomassie blue staining gel of SP Fast Flow column elution fractions same as Figure 3C. Black arrow indicates RIII monomer (9.2 kDa) on the SDS-PAGE gel.

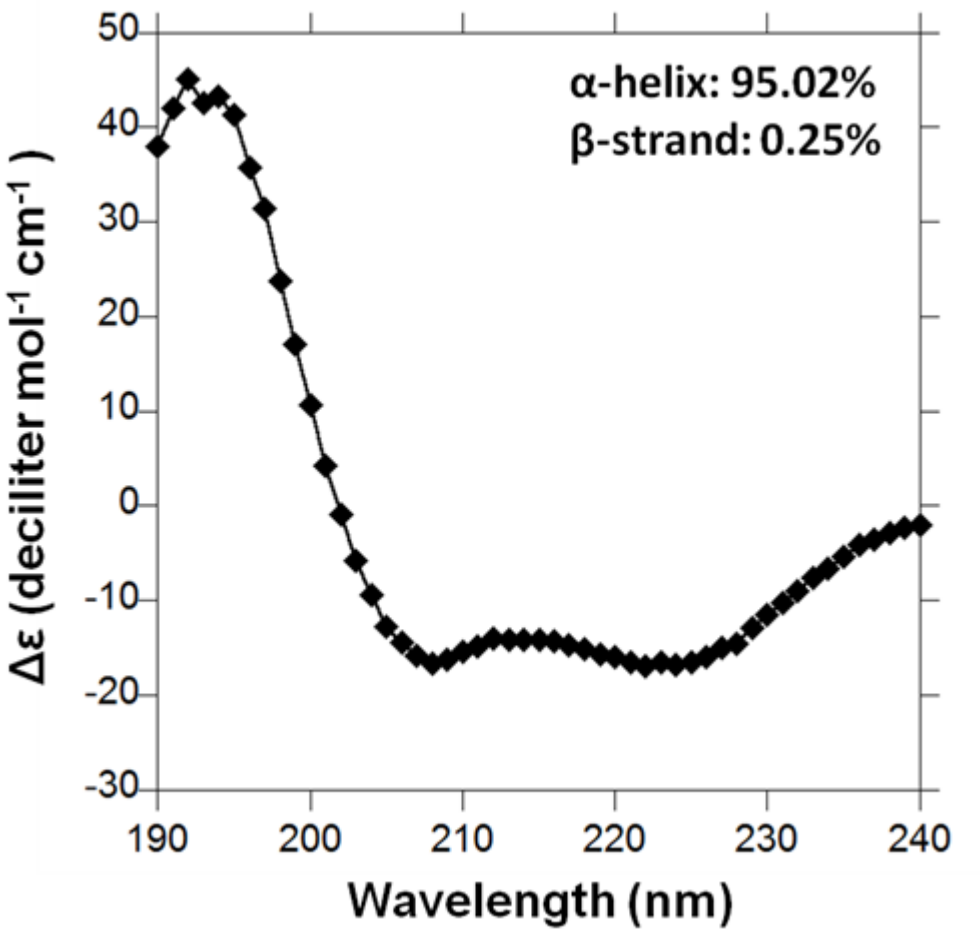

Figure 3.5 Representative circular dichroism spectra of purified RIII.
The final concentration is 5 μ*M*. The wavelength range is 190nm to 240nm. The helix and sheet percentage based on CDs data is shown.

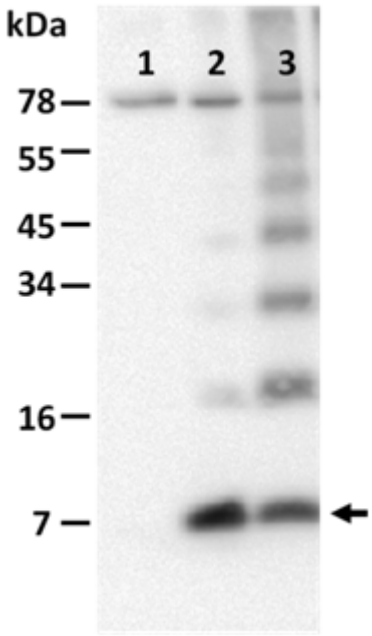

Figure 3.6 Western blotting of cross-linked RIII.
RIII forms SDS-resistant oligomers when cross-linked using DSP. *E. coli* MQλΔSR harboring plasmids pZA32RIIIs and pT4T(R-RzRz1)am were pelleted by centrifugation after the induction. 1, negative control without RIII; 2, DMSO control; 3. DSP+. The black arrow indicates the location of RIII monomer.

crosslinking; multiple RIII species other than monomer were detected under conditions where RIII expression was confirmed to delay T-mediated lysis (Figure 3.6), suggesting RIII may form transitory oligomeric species in vivo; however, no RIII-T heteromer was detected. To assess RIII-binding to T directly, we prepared inverted cytoplasmic membrane vesicles from *E. coli* cells induced for the *t* holin gene, incubated these vesicles with purified RIII protein, and then assessed whether RIII co-sedimented with the membrane material (Figure 3.7A). Western blotting showed that the amount of RIII in the soluble fraction decreased significantly after mixing with cells contain T, compared to isogenic control lacking the *t* gene (Figure 3.7B). Under these conditions, there are ~10^{14} molecules of T and ~10^{13} molecules of RIII in the reaction volume of 500 μL. These results suggest that, in the absence of RI, soluble RIII binds at least transitory to the N-terminal cytoplasmic domain of membrane-embedded T.

Quantification of RIII protein in T4 infection

It has been shown in several systems that the relative amounts of holin and antiholin protein can affect lysis timing (26, 117, 118). The stoichiometry of the RI antiholin and RIII co-antiholin proteins during T4 infection is thus of interest. We performed quantitative Western blotting using purified RIII as the protein standard. As described in Materials and Methods, protein samples were harvested from T4-infected cells by TCA precipitation, to prevent losses from cell lysis and analyzed for RIII protein content by quantitative Western blotting, using purified RIII protein as a standard. Under these conditions, RIII could be detected from 5 min after the infection, and accumulated up to ~8000 molecules per cell after 30 min of infection, which is comparable to the

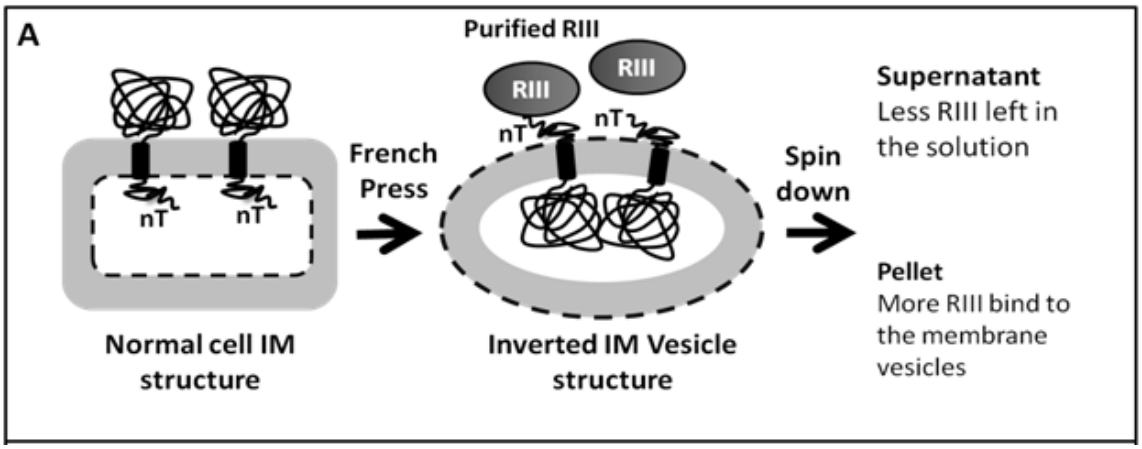

Figure 3.7 Pull-down of purified RIII protein.
(A) Schematic representatiion of the inverted membrane vesicle experiment. In inverted membrane vesicles, the N-terminal, cytoplasmic domain of T (nT) is exposed on the surface of the vesicle, allowing the direct contact of nT with externally added RIII proteins. The T protein on the vesicle then can pull RIII protein into the pellet after centrifugation. The grey area represents IM. The black dash line indicates the interface of IM originally facing the cytoplasm. PG and OM are omitted for clarity. For details, see Materials and Methods.

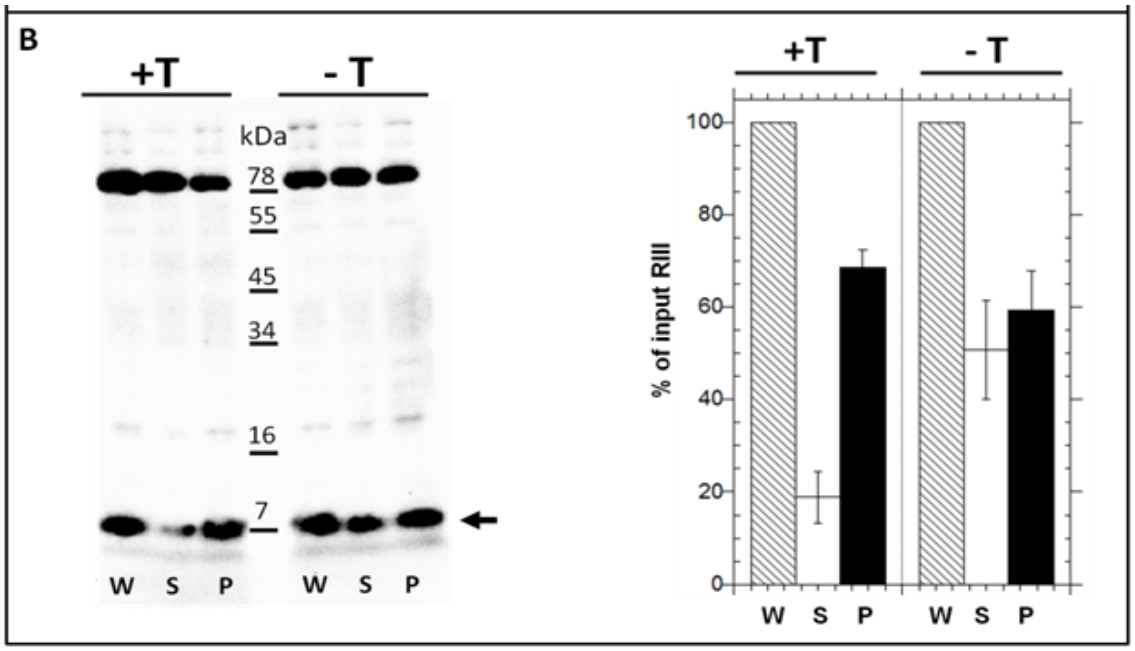

Figure 3.7 Continued.
(B) RIII binds to full-length T. After the induction of MQλΔSR cell harboring plasmid pT4T(R-Rz-Rz1)am, which express holin T but no other lysis proteins, the culture was lysed by French-press. Purified RIII was added to the lysate with a final concentration of 1 μM, incubated for 30 min at 25 °C. The membrane fraction containing the IM inverted vesicles was collected by ultracentrifugation as described. Samples containing equivalent amounts of each fraction were analyzed. Left, the Western blot result of inverted vesicles experiment. +T, with T holin expression; -T, no T control. W, whole cell sample; S, soluble fraction sample; P, pellet fraction sample. The black arrow indicates the location of RIII monomer. Right, the ratio of purified RIII in each sample was shown as the percentage of the input RIII. Patterned column, the whole cell sample ; hite column, the soluble fraction; black column, the pellet. Error bars show the standard deviation.

level of T (~4000 per cell) and RI (<10000 per cell) (Figure 3.8). The amount of RIII did not change drastically after ~30 min, which is the normal lysis time in the absence of LIN. There is no significant change in the amount of RIII expressed in T4D and T4Δ*rI*, showing that the absence of *rI* expression did not affect the production or stability of RIII. We conclude that both RI and RIII are expressed at about twice the level of T protein, consistent with the notion that inhibition of T holin triggering may involve the formation of inhibited ternary complexes containing the bulk of the holin in the infected cell, rather than the imposition of lysis inhibition by the formation of "poison" non-functional complexes.

Discussion

Although still mysterious at the molecular level, it is clear that holin triggering terminates the infection cycle and that the timing of the triggering event appears to reflect the attainment of a critical two-dimensional concentration of functional holin in the membrane (19). It is thus not surprising that in the three most intensively studied lysis systems, that of λ, phage 21 and T4, intrinsic holin timing is buttressed with an antiholin that functions to inhibit holin triggering (22, 34, 104, 119). In the λ and phage 21 cases, the holin and antiholin are produced as alternative products of cistrons with dual translational starts (28), and the antiholin products exert their inhibitory effect by heterodimerizing with the holin product and removing a proportion of the latter from the mass-action equation that leads to nucleation of the large two-dimensional aggregates, or 'rafts', within which hole formation presumably occurs (58). Importantly, there is as yet no evidence that there is any real-time or physiological regulation of the dual start

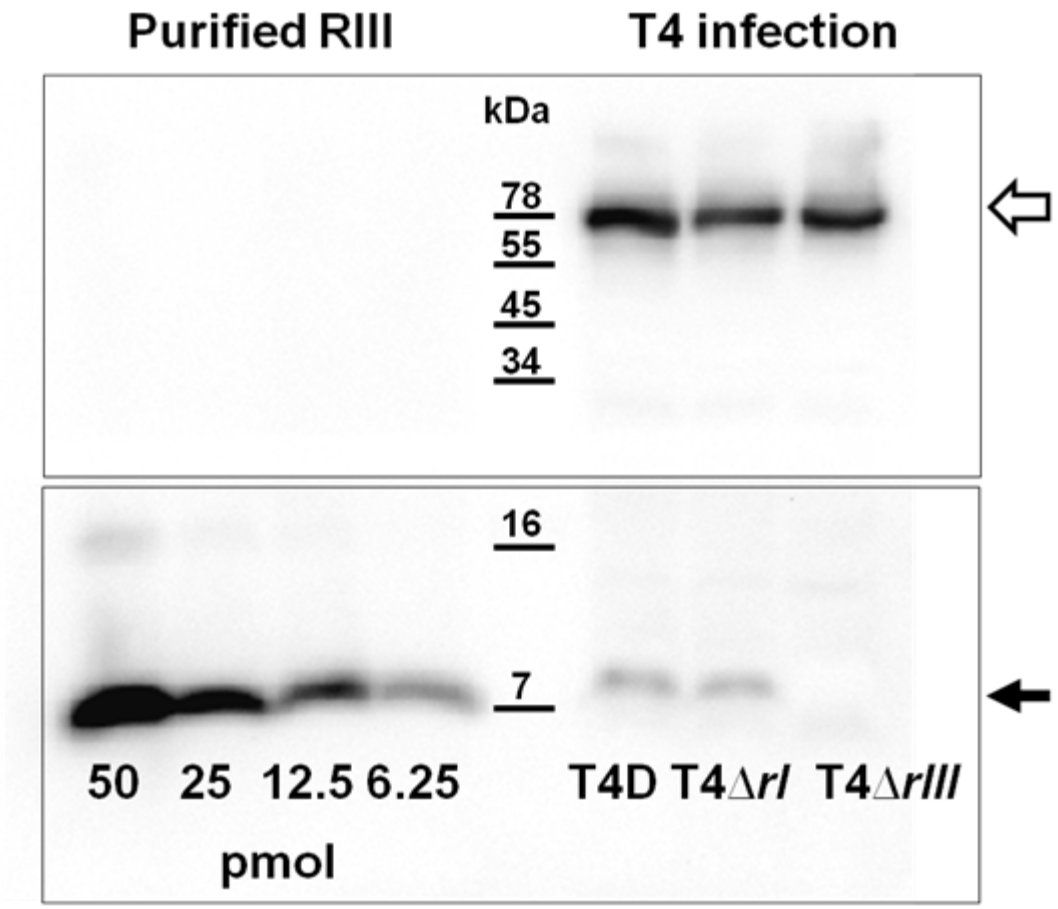

Figure 3.8 Quantification of RIII in T4 infection.
Samples were taken at 20min after infection. Black arrow indicates the location of RIII monomer (9.2 kDa) on the blot; Open arrow head shows the background bands which serve as the sample loading control. Purified RIII was used as protein standards with indicated amount.

systems, which instead seem to have evolved to fixed set points (32). Since λ and phage 21 are the best-studied class I (3 TMDs; N-out, C-in) and class II (2 TMDs, N-in, C-in) holin systems, and these are the two most common types of holins identified in phage genomes (85), it is fair to say that real-time or physiological regulation of lysis timing has not yet been shown to be important. However, the work reported here provides important details to the real-time, environmentally-controlled regulation of lysis timing of the T4 T holin, the type member of the much smaller group of class III holins (one TMD, N-in, C-out, large globular periplasmic domain).

Here we have presented results that show RIII is, as predicted from the absence of localization signals, a cytoplasmic protein and that, in T4 infections it is produced at levels that are comparable to that of the main antiholin, RI and the holin, T (Figure 3.8). RIII is thus the first soluble, cytoplasmic antiholin to be identified. In addition, CD analysis of purified RIII protein revealed that the secondary structure is dominated by alpha-helix (Figure 3.5); this matches structure predictions that posit three helical segments corresponding to N-terminal, middle and C-terminal domains (Figure 3.1). Interestingly, the RIII protein proved to be intolerant of oligo-histidine tags at either the N- or C-termini. Both N-his and C-his tagged proteins accumulated in heterogeneous oligomeric products that were resistant to SDS-PAGE (Figure 3.3C). In contrast, the wt protein has normal SDS-PAGE mobility and, in non-denaturing conditions, exhibits a tendency to dimerize, with ~5% of the total RIII involved in dimers stable enough to survive gel filtration (Figure 3.3). This tendency presumably underlies the capacity for self-interaction revealed by our B2H experiments (Figure 2.5). We speculate that the N-

terminus of RIII (nRIII) is responsible for this self-interaction, based the presence of that the N-terminal helical segment has coiled-coil character (Figure 3.1) and the finding that the G24D mutant allele abrogates the B2H signal (Figure 2.5). Similarly, based on the sensitivity of the B2H signal for RIII and nT to the distal mutants of RIII, we suggest that the C-terminal domain of RIII (cRIII) binds to the T protein. In support of this overall model, it is notable that the most conserved residues in RIII are in the C-terminus of RIII and presumably constitute the homotypic dimerization interface.

Although the over-expression of RIII can exert a mild delaying effect on T-mediated lysis, the major phenotypic effect of RIII is a robust stabilization of RI-mediated LIN (Figure 2.8). The current model for LIN is that RI, normally extremely unstable once released into the periplasm, is stabilized by superinfection by a T-even virion, by virtue of the ectopic localization of the contents of the virion head into the periplasm (35, 56). It is not known what component of the head contents, which includes ~10^3 protein molecules and the 170kb genomic DNA (53), causes the stabilization of RI. In any case, the stabilization raises the effective RI concentration and allows the formation of RI-T complexes involving the C-terminal globular domain of T (35, 57, 70). Until now, no comparable mechanism could be proposed for the activation of the RIII, since nothing from the superinfecting phage is thought to be transmitted either into the cytoplasmic membrane or into the cytoplasm. Recently, a crystal structure was determined for the complex between sRI (the RI protein with the SAR domain removed) and cT (the soluble periplasmic domain of T), in which sT:sRI:sRI:sT hetero-tetramers were the crystal forms (55). If this represents the true situation *in vivo*, then one

consequence of RI activation is to cause the formation of complexes with two nearby cytoplasmic nT domains. This leads to the simplest model for the "activation" of RIII, in the sense that, as a consequence of the initial binding of activated RI to T, a pair of nT domains appropriately arranged are made available for binding to a dimer of RIII. Thus the cytoplasmic complex would be nT:cRIII-nRIII::nRIII-cRIII:nT, where ":" represents a nT-nRIII heterotypic interaction and "::" the cRIII-cRIII homotypic interaction (Figure 2.8). A possible rationale for having a complex with determinants on both sides of the membrane is that it allows for the RI determinant to fluctuate significantly in its periplasmic concentration, as it would have if the stabilizing signal is just temporary, due to degradation of whatever the signaling molecule from the superinfecting phage. It has been reported that continued LIN requires re-superinfection at ~8 min intervals, presumably to allow the LIN block to collapse if environmental conditions (i.e., dilution of existing phages or increased availability of hosts) become more favorable (120). Moreover, LIN is also subject to collapse if the membrane potential is destroyed; loss of membrane energization might change the disposition of the TMD of protein T and thus alter the binding site of RIII. Future efforts are aimed at determining the details of the RIII-nT interaction.

CHAPTER IV

PHAGE MU: A NEW LYSIS REGULATION IN A PARADIGM PHAGE

Introduction

Mu is a *E. coli*-infecting myophage with a 36,717kb linear dsDNA genome (83). Since its discovery, phage Mu is well known mostly for its fascinating transposition ability, which has led to the deciphering of the mechanism of transposition of insertion elements in prokaryotes (76) . Mu replicates via a 'copy-and-paste' transposition mechanism involving random insertion of phage genomes into the host chromosome. As a result, more than 100 copies of the Mu genome can be found in a single host cell by the end of the infection cycle (79). Therefore Mu is considered as the most efficient transposon in nature. The DNA transposition is required for both lytic and lysogenic pathways of Mu. The Mu lysogen, unlike some lysogenic phages such as λ or P1, cannot be induced by usual prophage induction methods such as UV or DNA-damaging chemicals (77). To study the lytic phase of Mu, a series of Muc^{ts}mutants, which allow the thermal inactivation of the main repressor gp*c* protein at 42°C, were used. As a repressor, gp*c* not only maintains the lysogeny, but also immunizes the lysogen from phage homo-superinfection and thus contributes to the stability of the Mu lysogen (77) .

Compared to its transposition, little study has been conducted regarding host lysis, the last event of the Mu infection cycle (77). For dsDNA phages, the holin-endolysin-spanin system appears to be the universal answer to an important question:

how to conduct efficient and carefully scheduled lysis event. Holins make holes in the host inner membrane (IM), thus depolarizing the membrane and opening up a non-specific channel for cytoplasmic macromolecules, including endolysins to go into the periplasm. Endolysins are muralytic enzymes that degrade the peptidoglycan layer in the periplasm, and spanins disrupt the outer membrane to complete the lysis event (19). However, for Mu there is only one lysis related gene that has been identified to date: *lys*.(77)

Mu *lys,* or, as named in the annotated Mu genome, gene *22,* is predicted to encode a 171 aa protein with a N-terminal SAR domain, which is a TMD that consists of mostly weakly hydrophobic residues, and a large C-terminal soluble catalytic domain including the classic catalytic triad sequence, E-7X-D-5X-T that was identified in homologs of the canonical true lysozyme, the T4 endolysin E (121). Therefore, the product Lys is considered to be a SAR-endolysin, which is synthesized as an inactive membrane protein that is 'anchored' in the IM, then is activated by conformational change upon escape from the IM. The most recent bioinformatic analysis suggested that Lys is a R^{21}-like SAR-endolysin (38).

lys is mapped between gene *21*, which encodes late promoter activator C, and gene 23, which encodes a 128 aa IM spanin (i-spanin) with a N-terminal TMD (47). *lys* is the first gene on a late transcription unit, that also includes phage head morphogenesis genes *D* , *E, H* and *F* (genes *27-30*) controlled by late transcription activator C protein (77). The most recent annotation effort (47) has identified a new gene, *23a*, which is partially overlapped with gene *23* and encodes a lipoprotein, s thus unambiguously

identifiable as the OM spanin (o-spanin) gene of Mu. The proximity of the spanin gene pair (g23/g23a) and the experimentally confirmed endolysin gene *(lys*) suggests that Mu has a λ -like lysis cassette, which includes holin, antiholin, endolysin and spanins (47). However, other than the endolysin and i-spanin genes, there are three other genes encoding proteins with one or more TMDs in this gene cluster, genes *19, 20*, and *25*, making it unclear which protein could serve as a holin. Here we report the identification and analysis of the holin-antiholin gene pair in the Mu lysis cassettes and the discovery of a novel cofactor for the SAR-endolysin. The results are interpreted in terms of the new model for the pinholin-SAR-endolysin lysis pathway.

Materials and methods

Bacterial growth and induction

Phages and bacteria strains used in this study were listed in Table 4.1. When plating bacterial strains, standard LB-agar plate supplemented with the appropriate antibiotics (ampicillin, 100 μg mL-1; chloramphenicol, 10 μg mL-1; kanamycin, 40 μg mL-1) were used (56, 85). Plates were incubated at 30°C for Muc^{ts} lysogens and 37°C for non-lysogenic *E. coli* strains as described. For making overnight cultures of strains harboring the pBAD plasmid, glucose was added to the LB to a final concentration of 0.4%, in addition to antibiotics. To start fresh cultures, overnight cultures inoculated from single colonies were diluted to A_{550} ~ 0.03 in LB with antibiotics and grown at 30°C or 37°C with aeration. Bacterial growth was monitored using a Gilford Stasar III sipping spectrophotometer (Gilford Instrument Inc, Oberlin, OH), as described (85). Muc^{ts} lysogens were induced by shifting incubation temperature from 30°C to 42°C. For

λ lysogen induction, cultures were aerated in 42°C for 15 min and then shifted to at 37°C. pBAD and pZA plasmids were induced by adding arabinose to a final concentration of 0.4% (v/v) (122) or 1 mM isopropyl b-D-thiogalactosidase (IPTG) (123), respectively . MC4100 araR strains were isolated from the survival colonies after plating cells on the TB-agar plates containing 1% (v/v) arabinose and re-streaked on 1% arabinose TB plates for two times (124, 125).

PCR and plasmid construction

See Table 4.1 for plasmids used and constructed in this study. Isolation of plasmid DNA, DNA amplification by polymerase chain reaction (PCR), DNA transformation, and DNA sequencing were conducted as described in previous chapters. Mu lysogens and lysates were used as templates for cloning Mu genes. Mu *19*, *22* (*lys*), and *25* were cloned in pBAD24 (122) between its unique EcoRI and HindIII sites, or in pBAD33 (119) plasmid between its unique KpnI and HindIII sites, with their native ribosome binding Shine-Dalgano (S-D) sequence (plasmids named with "osd") or with stronger S-D sequence (AGGAG) (plasmids named with "ssd"). Lys or Gp25 was tagged with the C-myc tag or His6-tag, respectively, by inserting the DNA sequence encoding theses tags in at the end of the gene. Gene *22* and *25* mutants were constructed by Quick-Change site-directed mutagenesis, as described previously, on pBAD33-g22ssd or pBAD24-g25ssd, respectively. To construct a plasmid carrying artificial TMD Gp25 (Gp25TMDswap), a synthetic dsDNA fragment containing *25* gene with its native TMD aa sequence (LISVLALWPYLLPVVAGGAVWAM) replaced by an artificial TMD sequence VLLIIVVVVVVVVIILLI (41) (See Table 4.1 for DNA sequences) was

ordered from gBlocks® Gene Fragments service provided by Integrated DNA Technologies (Coralville, CA) in a purified form. The synthetic dsDNA was then digested with EcoRI and HindIII before being ligated into the pBAD24 plasmid.

Construction and analysis of Mu knockout mutants

Construction of Mu*c*^{ts}Cam knockout mutants were performed as described (118, 124). Briefly, various flanking regions of the target genes (*19*, *20*, *25* and *lys*) were attached to the *cat* gene via PCR using primers listed in the Table 4.1. Then, 10-100 ng of purified linear dsDNA PCR products was digested with DpnI and transferred into MC4100 araR cells carrying Red helper plasmid pKD46 by electroporation. MC4100 araR cells prior to the electroporation were grown at 30°C in LB supplied with Amp and L-arabinose (Ara) and then resuspended in 10% glycerol. The transformants were plated onto LB/Cam plates to select Cam-resistant colonies. The cat insertion was confirmed by PCR and sequencing. To compare the plaque morphology of Mu wt and mutants, phage lysates were prepared from various induced Mu*c*^{ts}lysogens in a method that have previously described. Spot-titer of phages was conducted by spotting 5µl of phage lysate serial dilution onto a lawn of MG1655 and then incubating overnight at 42°C. The plaque size was analyzed by ImageJ (NIH).

Table 4.1 Strains and phages used for Chapter IV.

Phages	**Description**	**Source**
Muc^{ts}	Phage Mu that carry a ts mutant c protein that can be deactivated in 42°C .	Laboratory Stock
λ*Q*^{21} *Δ(SRRzRz1)*^{21}	λ cI^{857} hy(*Q*^{21}*(SRRzRz1)*^{21}*::CamR*) *bor::kan*	Pang Ting 2009
Bacteria Strains	**Description**	**Source**
MC4100	*E. coli* K-12 F- *araD139 □(argF-lac)U169 fhuA rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR*	Laboratory Stock
MG1655	*E. coli* F- λ - *ilvG- rfb-50 rph-1*	Laboratory Stock
MG1655 *tonA::Tn10*	MG1655 with Tn10 inserted in *tonA* gene	Laboratory Stock
MDS12 *tonA*::Tn*10 lac*I^{q1}	MG1655 with 12 deletions, totaling 376,180 nt including cryptic prophages	(35)
Plasmids	**Description**	**Source**
pS^{21}68	Phage 21 lysis cassette under its native promoter pR'^{21} on pBR322 backbone; encodes S68 only	(85)
pMu19	Isogenic to pS^{21}68; S^{21}68 replaced by *Gp19* from Mu; (nt 9440-9825 of phage Mu)	(85)
pMu20	Isogenic to pS^{21}68; S^{21}68 replaced by *Gp20* from Mu; (nt 9815-9955 of phage Mu)	(85)
pKD46	*(araD-araB)567 lacZ4787(::rrnB-3)□□*$^{-}$ *rph-1(rhaD-rhaB)568 hsdR514*	(126)
pKD3	*(araD-araB)567 lacZ4787 (::rrnB3)□□□(phoB-phoR)580* $^{-}$ galU95 *uidA3::pir+ recA1 endA9 (del-ins)::frt rph-1 (rhaD-rhaB)568 hsdR514*	(124)
pBAD24	pBR322 based plasmid with P_{BAD} promoter; araC, AmpR	(119)
pBAD33	pACYC184 based plasmid with P_{BAD} promoter; araC, CamR	(122)
pBAD24-g19-ssd	Mu *19* inserted in pBAD24 with plasmid S-D	This study
pBAD24-g22-osd	Mu *22* inserted in pBAD24 with its native S-D	This study
pBAD24-g22-ssd	Mu *22* inserted in pBAD24 with plasmid S-D	This study
pBAD33-g22-osd	Mu *22* inserted in pBAD33 with its native S-D	This study
pBAD33-g22-ssd	Mu *22* inserted in pBAD33 with plasmid S-D	This study
pBAD33-g22-K6G	K6G mutation in g22 gene of pBAD33-g22-ssd	This study
pBAD33-g22-K7G	K7G mutation in g22 gene of pBAD33-g22-ssd	This study
pBAD33-g22-K9G	K9G mutation in g22 gene of pBAD33-g22-ssd	This study
pBAD24-g25-D71V	D71V mutation in *25* gene of pBAD24-g25-ssd	This study
pBAD24-g25-E87V	E81V mutation in *25* gene of pBAD24-g25-ssd	This study

Table 4.1 Continued.

Plasmids	**Description**	**Source**
pBAD24-g25-D91V	D91V mutation in *25* gene of pBAD24-g25-ssd	This study
pBAD24-g25-TMDs	Native TMD replaced with an artificial TMD sequence in *25* of pBAD24-g25-ssd	This study
pZA32-mycR-lacZ	mycR-lacZ replacement of *luc* in pZA32-luc	(123)
Description	**DNA sequence**	**Source**
Native Gp25 TMD	TTGATTTCAGTTTTAGCGTTATGGCCTTACCT GTTGCCTGTTGTGGCCGGTGGGGCCGTCTG GGCGATG	This study
Artificial TMD	GTGCTGCTGATTATTGTGGTGGTGGTGGTGG TGGTGGTGATTATTCTGCTGATT	(41)

SDS-PAGE and Western blotting

Protein sample preparation, SDS-PAGE and Western blotting were conducted as described previously (24). Briefly, the whole-cell samples were taken at various time points before or after induction and added with 10% v/v ice-cold Trichloroacetic acid. Samples were then washed with cold acetone for 3 times before mixing with SLB. 10% Acrylamide SDS gel was used for SDS-PAGE. For Western blotting, 3% (w/v) milk-TBS buffer was used as blocking solution. His6-tagged or C-myc tagged proteins were detected by Sigma-Aldrich (Carlsbad,CA) mouse monoclonal α-his or α-Cmyc primary antibody 9E10 purchased from Thermofisher (Waltham, MA) at a dilution of 1:4000 or 1:1000, respectively. Blots development and visualization were performed as described. The band density of the blots was analyzed using ImageJ (112).

Complementation assay of Mu holin in S^{21} lysis cassette

The pS^{21}68 derivatives pMU19 and pMu20 were constructed as described previously (85). pS^{21}68 is a pBR322-based, medium copy-number plasmid carrying the phage 21 lysis cassette $(SRRzRz1)^{21}$ under its native promoter (21). The induction of pS^{21}68 thus requires the presence of phage 21 antiterminator Q^{21}. pMu19 and pMu20 have hybrid lysis cassettes where Mu *19* or *20* replaced S^{21} holin gene in phage 21 lysis cassette $(SRRzRz1)^{21}$, respectively. pMu19 or pMu20 was transformed into *E. coli* MDS12 *tonA*::Tn*10 lacI*^{q1}cells lysogenized by a chimeric phage λ $Q^{21}\Delta(SRRzRz1)^{21}$, which is a λ phage with all of its lysis genes deleted and had its late transcription antiterminator Q replaced by Q^{21} (21). The induction of this lysogen thus can induce pS^{21}68 and its derivative plasmid by transactivation. Lysis profiles were monitored after

the induction of lysogens using a Gilford Stasar III sipping spectrophotometer (Gilford Instrument Inc, Oberlin, OH), as described (118).

Results

Identification of holin and antiholin candidates in My lysis cassette

As the first step in identifying the Mu holin, we have analyzed the membrane topology of all undefined hypothetical proteins using TMHMM (DTU, Denmark) to look for proteins contain at least one TMD. We found three holin candidates: the product of gene *19*, *20*, and *25*. As shown in Figure 4.1, all these three genes are closely localized to *lys*, and all of their products are predicted to be integral membrane proteins. However, none of the candidates have a topology profile that fits our previously studied holin classes (Figure 1.4): Gp19 has four TMDs with both N and C termini in the cytoplasm; Gp20 and Gp25 are type II integral membrane proteins with a single N-terminal TMD, which is similar to T4 holin T, but Gp20 has very small soluble domains in cytoplasm and periplasm, and Gp25 has a long cytoplasmic tail instead of the large periplasmic domain that T has.

To examine the function of these three holin candidates, we have knocked out gene *19*, *20*, *lys*, or *25* separately in MC4100(Mu*cts*) lysogens and analyzed their lysis profiles upon induction. The wt Mu*cts* lysogen started to lyse at ~ 40 min after the thermal induction (Figure 4.2A, left). The induction of Mu *lys::Cam* lysogen did not result in lysis, as expected for an endolysin-defect mutant (Figure 4.2A, right). The plateau in culture growth reflects the lethal triggering of holin. It was experimentally confirmed that the holin-independent lysis caused by the spontaneous release and

activation of SAR-endolysin is usually delayed compared to the wt, holin-triggered lysis (20). This delay might led to the differences in plaque morphologies of phage 21 wt and ΔS^{21} mutant (20). The deletion of gene *19* caused about ~ 20min of delay in lysis timing in liquid culture (Figure 4.2A) and notably smaller plaques when plated on the lawn of MG1655 compared to those of Muc^{ts} wt (Figure 4.2B, middle row). The deletion of *g20* also caused only slightly delayed lysis (~5 min), but the growth of the cell was significantly retarded after the induction of the lysogen. The knock-out of *20*, in contrast to the knock-out of *19*, did not affect the plaque morphology (Figure 4.2B, bottom row). These results made Gp19 and Gp20 two primary candidates for Mu holin.

Mu Gp19-Gp20 is a pinholin-antiholin pair

To study the function of Gp19 and Gp20, we have replaced the S^{21} pinholin gene in plasmid pS^{21}68 with Mu *19* or *20*, resulted in plasmid pMu19 or pMu20, respectively. pS^{21}68 is a plasmid carrying the entire phage 21 lysis cassette under the control of phage 21 pR' late promoter. In this lysogen, the expression of the plasmid genes is activated in trans by by the induction of prophage through production of the phage 21 antiterminator Q^{21} (20). When induced by thermal shift, as described previously, the gene *19* construct (pMu19) supported lysis of the culture with a lysis profile similar to that of wt phage 21 lysis cassette construct (pS^{21}68) (Figure 4.2C).

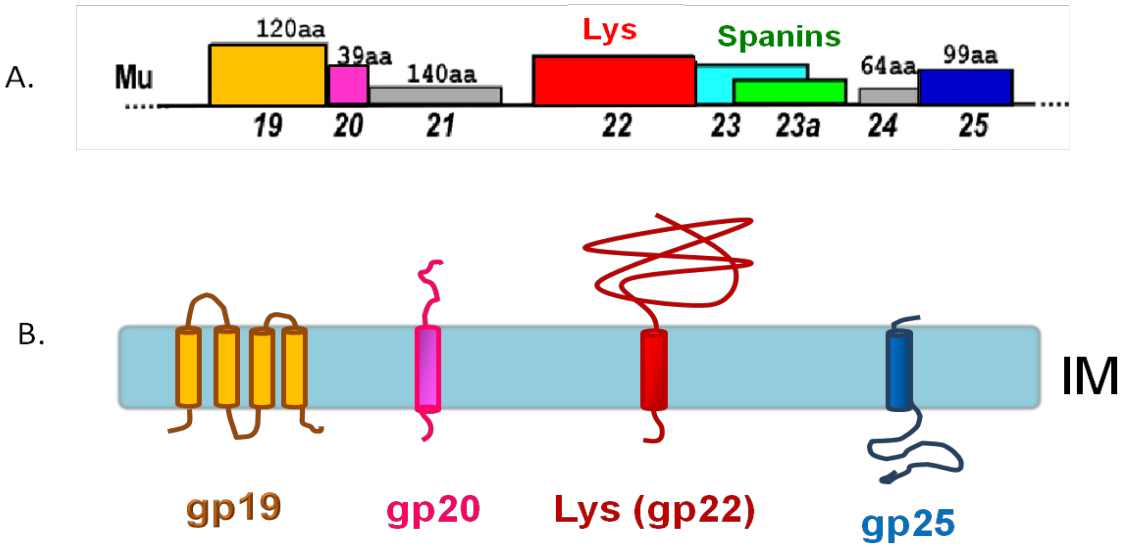

Figure 4.1 Prediction of topology and structure of Mu lysis proteins.
(A) Lysis cassette of Mu.
(B) Topologies of Lys and holin candidates.

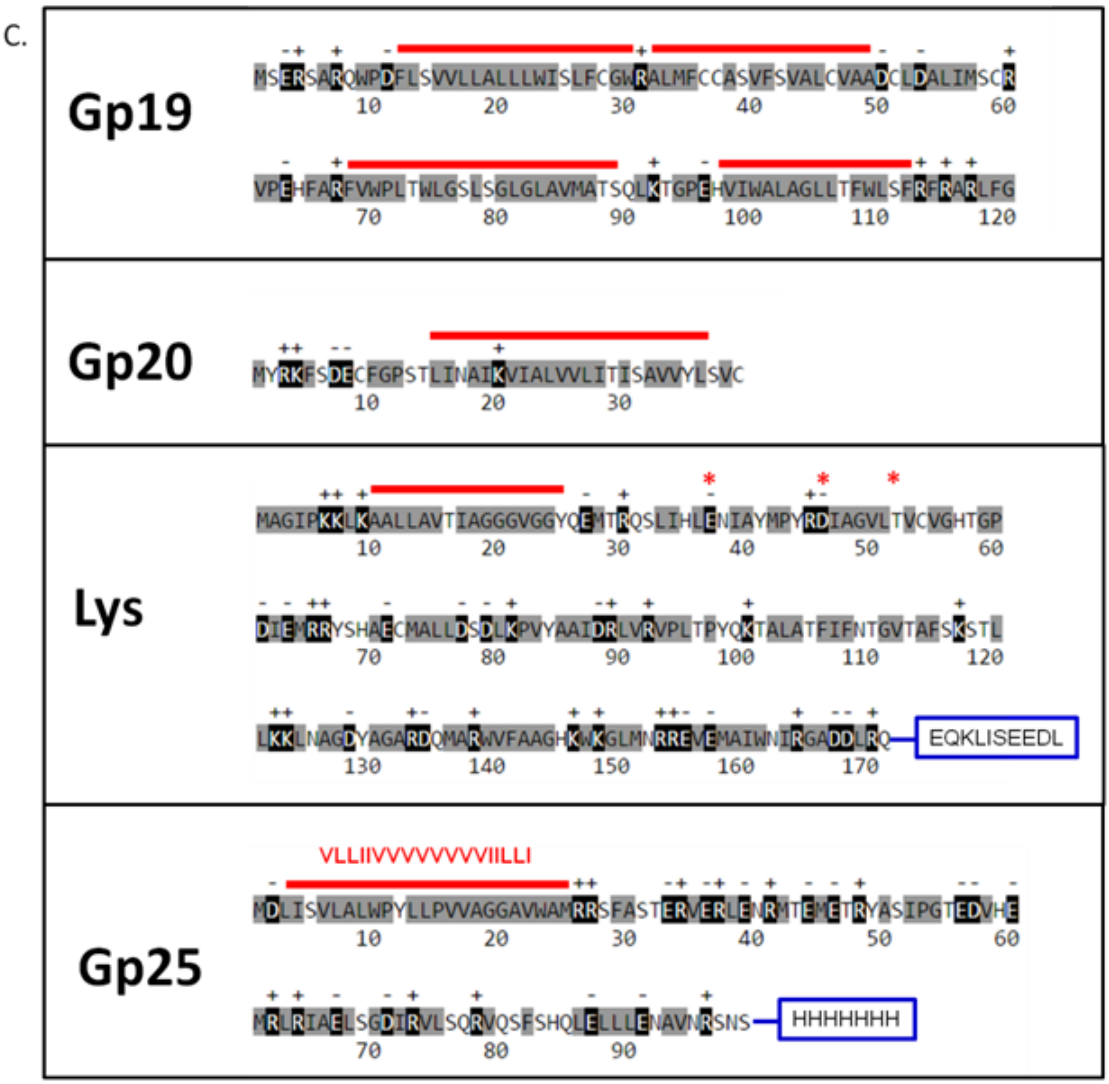

Figure 4.1 Continued.
(C) Primary structure of Lys and three holin candidates. Predicted TMDs are shown as red lines. The C-myc tag in Lys and C-his tag in Gp25 is shown in the blue box. The artificial TMD that replaced the native TMD of Gp25 is shown above the wt sequence.

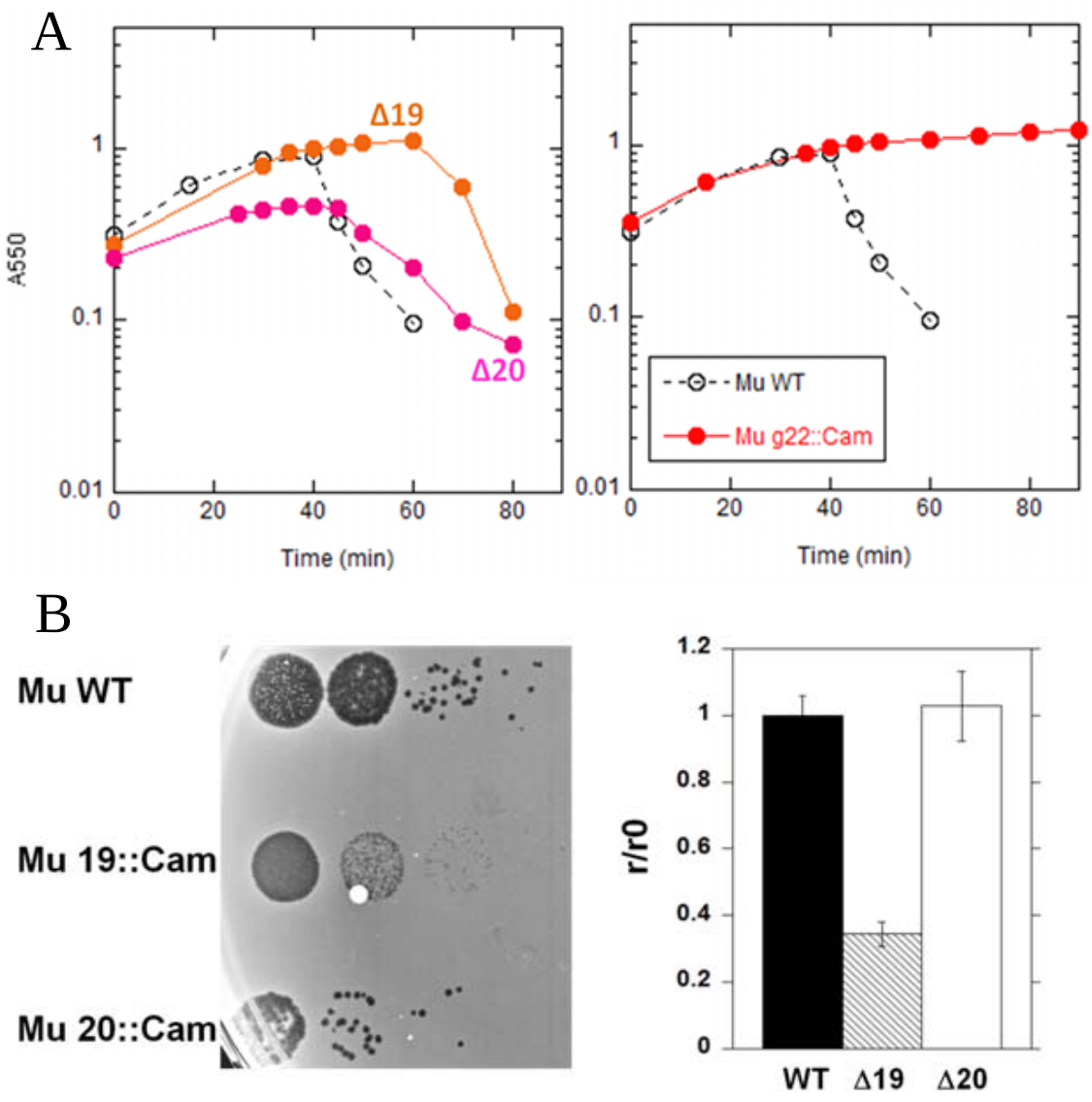

Figure 4.2 Deletion of gene *19* and *20*.
(A) (Left) Lysis profiles of Muc^{ts} wt (open circle), 19∷Cam (Orange circle), 20∷cam (pink circle), and (right) lys∷ Cam (red circle) mutants induced at 42 °C.
(B) Plaque morphology of Muc^{ts} wt and mutants. The right panel shows the average size of plaques (r) as the ratio to the average size of wt plaques (r0).

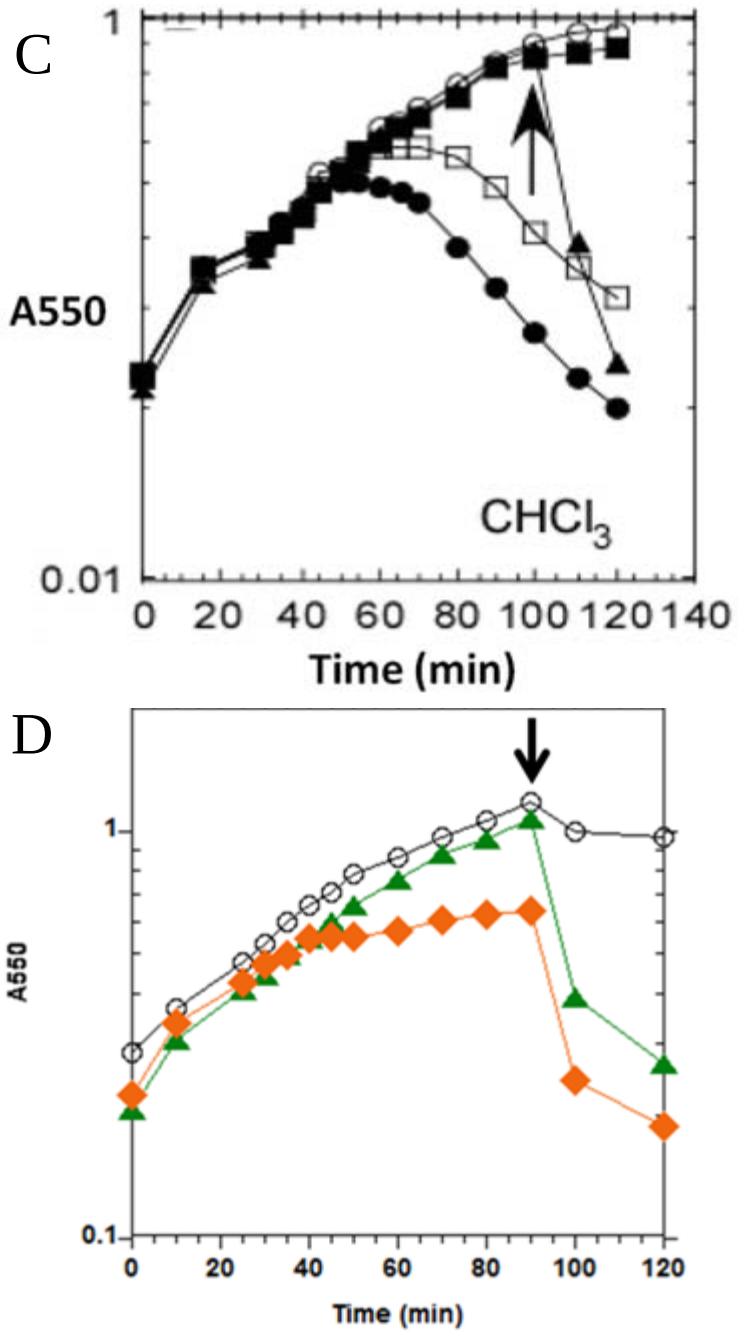

Figure 4.2 Continued.
(C) Gp19 can complement S^{21} in phage 21 lysis cassette. $pS^{21}68$ (black circle), pS 2168am (black square), pMu19 (open square), pMu20 (open circle). pMu20 with $CHCl_3$ added (black triangle) at the time indicated by black arrow (85).. **(D)** Gp19 does not function with the cytoplasmic endolysin. MG1655 (open circle) carrying pZA32-cmycR-lacZ (green triangle) alone or with pBAD24-g19-ssd were induced with IPTG or arabinose. $CHCl_3$ was added to all samples at the time indicated by black arrow.

In contrast, the induction of pMu20 did not cause lysis. The immediate lysis after adding $CHCl_3$ (Figure 4.2C, black arrowhead) confirmed that the lack of lysis is not due to the shortage of SAR-endolysin (R^{21}), but due to the lack of the functional holin in pMu20 construct. The retardation of the culture growth upon induction and the earlier lysis in Mu *g20*::Cam (Figure 4.2A) suggests that Gp20 inhibits the membrane toxication from the accumulation of holin to support the robustness of the host cells and prevent the premature lysis by blocking holin triggering. Taken together, we assigned the function of Mu holin to Gp19, and the function of antiholin to Gp20.

The presence of SAR-endolysin suggests that the Mu holin is only required for the depolarization of the IM for the fast and saltatory release of muralytic enzyme from the membrane. Therefore, it is not necessary for Mu to have a canonical holin which makes large membrane lesions. However, the phage P1 lysis cassette includes a SAR-endolysin Lyz and a canonical holin LydA, which forms IM holes that are large enough to allow passage of cytoplasmic endolysins (30).To test if Mu Gp19 forms canonical IM holes, we co-expressed Mu Gp19 and Cmyc-R-lacZ, which is a hybrid of C-myc tagged λ endolysin R and an intact β-galactosidase with a total size of ~0.5 MDa (123). As shown in Figure 4.2D, the expression of Gp19 caused culture growth to plateau at ~ 40 min, which is the wt Mu lysis time, but did not lyse the cell with the presence of Cmyc-R-LacZ. These results indicated that Gp19 is a pinholin that only forms small IM holes. The TMD prediction result using TMHMM suggested that Gp19 adopts a N-in, C-in topology with four predicted TMDs, which is distinct from all three classes of holins that have been studied experimentally. (Figure 1.4)

Mu gene 25 *is required for lysis*

As mentioned above, we have originally identified three holin gene candidates in the proximity of lys: *19*, *20* and *25*, and have constructed knock-out mutants for each of them. After we assigned the holin function to gene *19* and the antiholin function to gene *20*, the lysis system of Mu seems complete, since we have identified all protein factors that have been previously identified in a phage lysis cassette: holin Gp19, antiholin Gp20, endolysin Lys, and the spanin complex gp23 and gp23a. However, to our surprise, the induction of Muc^{ts} *25::Cam* prophage did not result in lysis (Figure 4.3A). Instead, the growth of the culture reached a plateau after ~ 40 min, indistinguishable from the induction of the Muc^{ts} *lys::Cam*. The possibility of the unexpected mutation in other lysis genes was ruled out by the sequencing result of the entire Mu lysis cassette in Muc^{ts} *25::Cam* lysogen. To examine if the lack of *25*-encoded protein is responsible for the impeded lysis, we constructed a pBAD24 plasmid carrying a clone of *25* gene with its original S-D sequence. The resulting plasmid, pBAD24-g25-osd, was able to restore the lysis phenotype when induced in MC4100 araR Muc^{ts} *25::Cam* lysogen but no in *lys::Cam* lysogen (data not shown). These results suggest that Gp25, the product of gene *25*, is required in the host lysis event of Mu, despite the presence of a functional version of the five known types of lysis proteins.

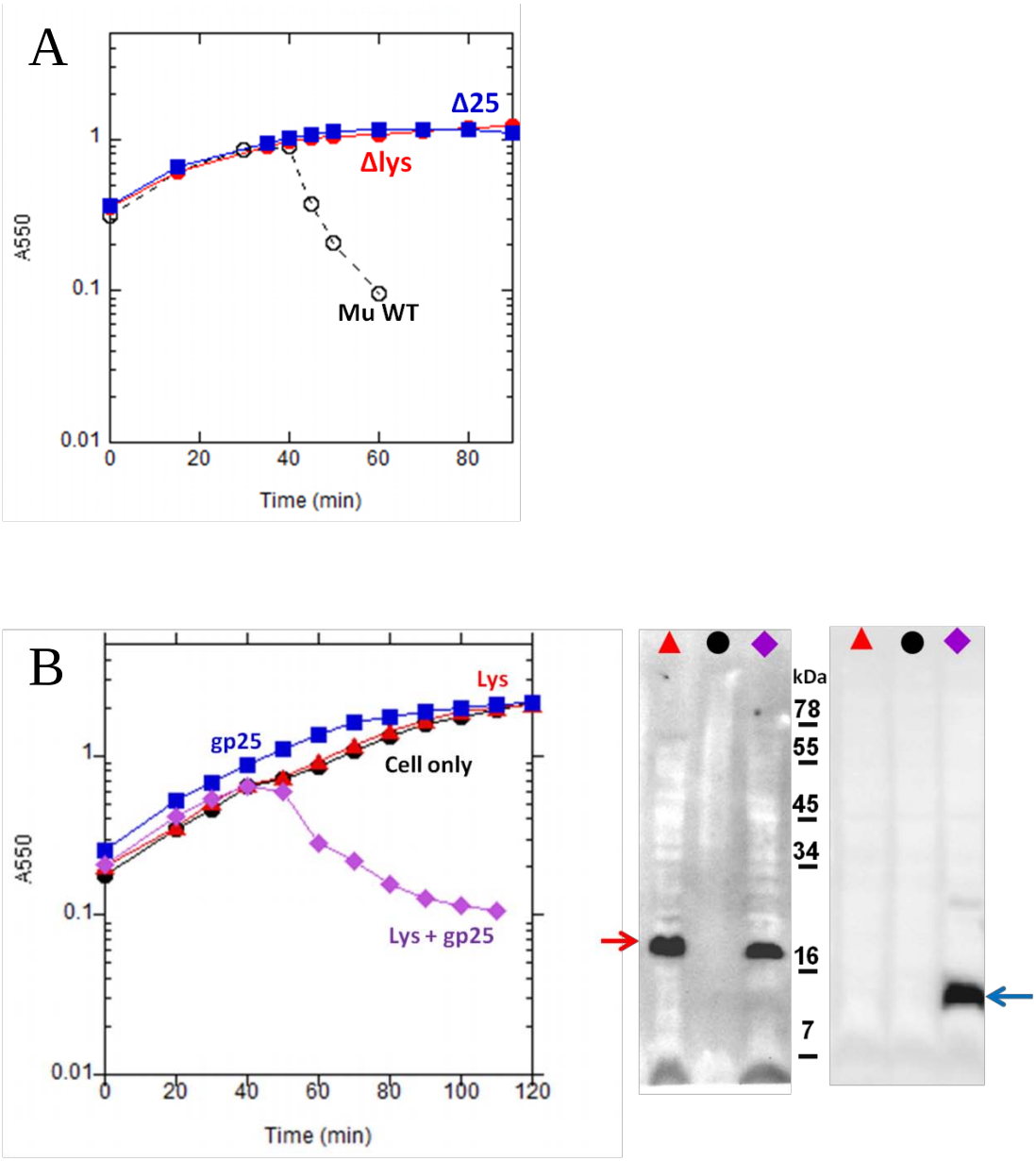

Figure 4.3 Co-expression of Gp25 and Lys.
(A) Induction of Muc^{ts}*25*∷Cam lysogen (blue square) does not cause lysis the host. The red curve indicates the induction of Muc^{ts} lys∷Cam.
(B) (Left) Induction of pBAD24-g25-ssd (blue circle), pBAD33-g22-ssd separate (red triangle) or together (purple diamond) in MG1655 *tonA::Tn10*. The black curve is the cell-only control. (Right) WB using α-Cmyc (left panel) or α-his (right panel) antibody. The red or blue arrow indicates the band corresponding to Cmyc-Gp22 or C-his-Gp25, respectively.

Gp25 assists the function of endolysin

Since the Lys is a SAR-endolysin that can cause gradual lysis without the holin, the only reasonable explanation to the loss of lysis in Mu *25::Cam* mutant is that the function of SAR-endolysin was abolished without Gp25. To test this hypothesis, we constructed two plasmids, pBAD33-lys-Cmyc-ssd and pBAD24-g25-Chis-ssd, to express C-myc tagged Lys and C-terminal his-tagged Gp25, respectively. The function of C-myc-Lys or C-his-Gp25 was tested by complementation assay with Mu knock-out mutants, and both tagged proteins retain wt function. As previously reported (40), plasmid-encoded SAR-endolysin can cause lysis without holin or spanin in our experimental condition, because once the cell wall was degraded by SAR-endolysin, the shearing forces of the shaker flasks (e.g. using of the orbital shakers for aeration) is sufficient for complementing the lysis defect associated with spanin mutantss. When induced in *E. coli* MG1655 *tonA::Tn10* cells, neither pBAD33-lys-cmyc-ssd nor pBAD24-g25-Chis-ssd alone caused lysis (Figure 4.3B, left panel), even though the expression of C-myc-Lys or C-his-Gp25 protein was confirmed by Western blotting (Figure 4.3B, right panel), but the co-expression of C-myc-Lys and C-his-Gp25 led to the lysis at ~50 min (Figure 4.3B). Taken together, Gp25 is a requiredcofactor of Lys that activates its muralytic enzyme function.

Gp25 does not affect the accumulation of Lys

Previously, SAR-endolysin was illustrated as a membrane-anchored muralytic enzyme that can be released into periplasm, where it gets activated and degrades the cell wall, in a *sec* system-dependent and membrane-potential dependent manner. Although

the holin defines the lysis timing and helps to achieve the saltatory and rapid lysis, the release and activation of SAR-endolysin is considered to be independent from any phage encoded proteins (40). However, without the presence of Gp25, Mu Lys did not respond to the membrane depolarization by the holin triggering (Figure 4.2A) or the adding of 10 mM KCN (Figure 4.4). To examine if higher expression level of the protein can lead to a Gp25-independent lysis, we have constructed four different Lys-encoding plasmids, pBAD33-lys-osd, pBAD24-lys-osd, pBAD33-lys-ssd, and pBAD24-lys-ssd, which have different expression level of Lys when induced. Compare to the lowest Lys expression level (pBAD33-lys-osd), other three plasmids (pBAD24-lys-osd, pBAD33-lys-ssd, and pBAD24-lys-ssd) have 2-fold, 10-fold, and 50-fold increased Lys expression, respectively. All of these four plasmids caused lysis when co-expressed with Gp25. However, when expressed alone, only pBAD24-lys-ssd, which has 50-fold increase in protein expression level, can cause lysis independently (Figure 4.5A). This result brought up a question if Gp25 helps Lys by up-regulating the protein expression or by stabilizing Lys. As reported by Tran et al, in the case of T4 antiholin RI protein, the SAR-domain caused instability of the protein and resulted in a rapid protein degradation in the periplasm (35). In contrast to T4 RI, which has a half-life of ~90 sec, Lys is stable *in vivo* (Figure 4.5B). Also, as shown in Figure 4.3B, the accumulation of Lys was not affected by the co-expressed Gp25 (Figure 4.3B, right panel). These results suggest that Gp25 is activating Lys in a way other than stabilizing the protein.

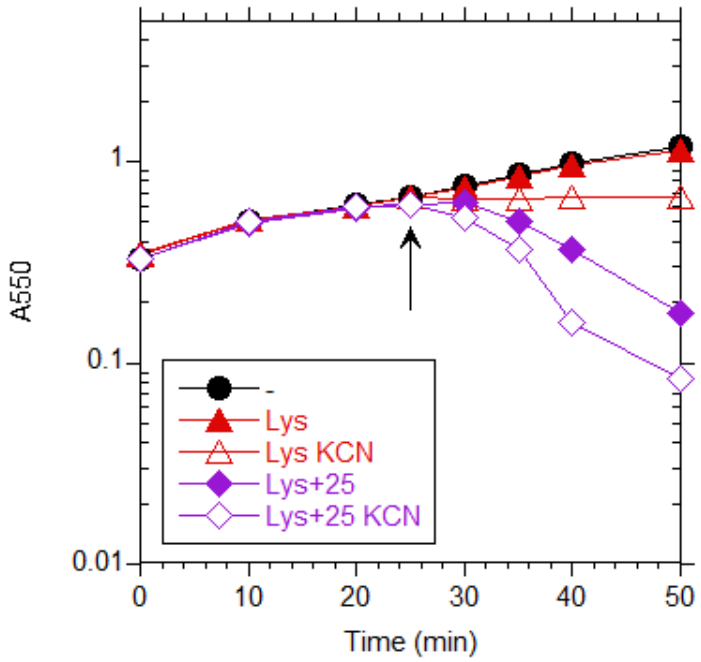

Figure 4.4 Lys did not respond to KCN.
pBAD33-g22-osd was induced by arabinose in MG1655 tonA::Tn10 alone (red triangles) or with pBAD24-g25-ssd (purple diamonds). The black circle represents uninduced control. 10mM of KCN was added at 25min after induction (indicated by a black arrow) in samples represented by open symbols.

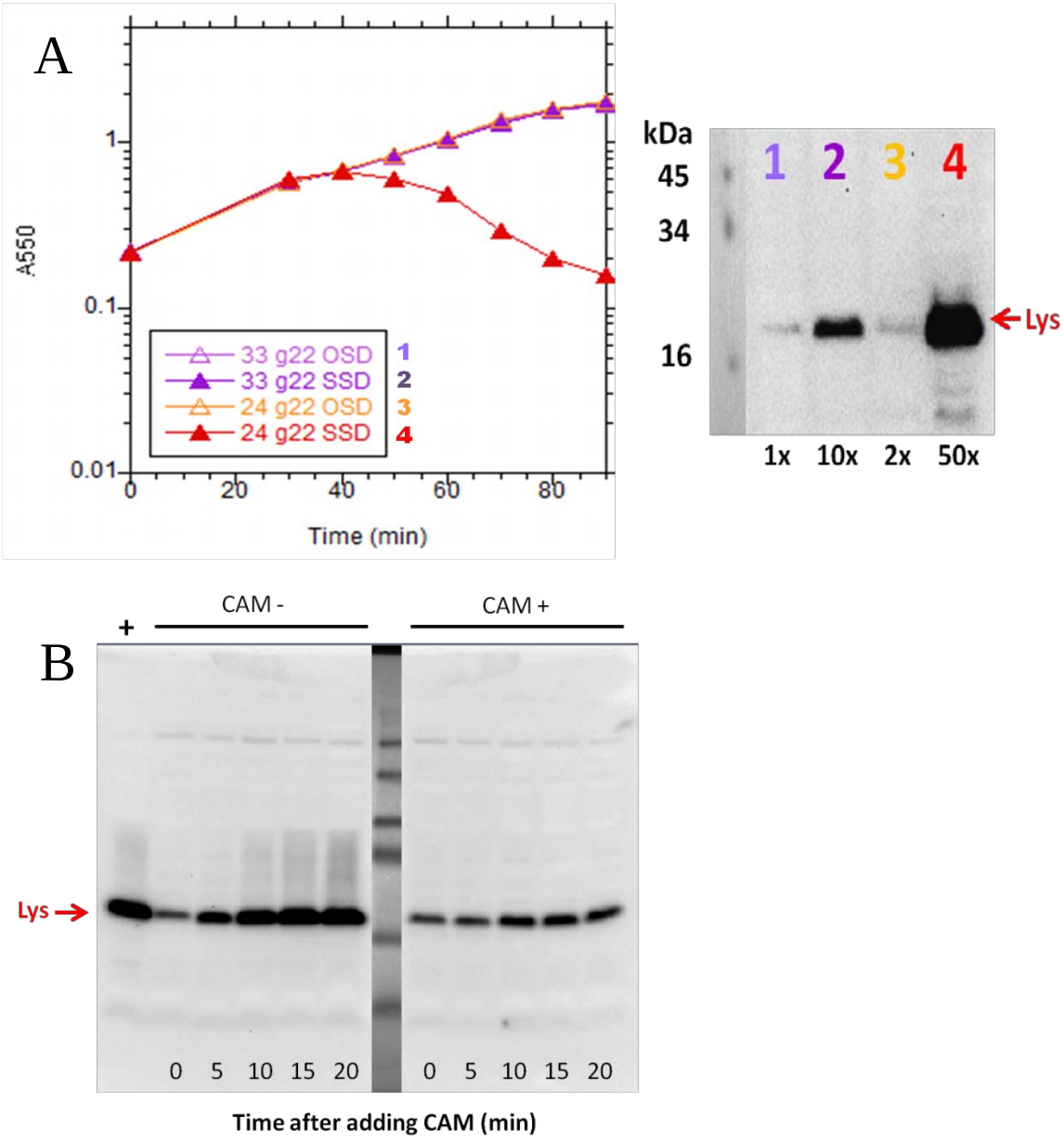

Figure 4.5 Accumulation of Lys.
(A) Induction of Lys from various pBAD vectors using the native *22* S-D (osd) or artificial S-D (ssd). 1, pBAD33-g22-osd; 2, pBAD33-g22-ssd; 3, pBAD24-g22-osd; 4, pBAD24-g22-ssd. The right panel shows the Western blotting (WB) results level of Lys accumulation from the equivalent amount of samples taken from 1-4. The expression level was mentioned as the fold to the lowest level (Lane 1).
(B) Lys is stable in the cell. +, WB positive control of Lys; CAM-, no chloramphenicol added; CAM+, added with 300ug/mL of chloramphenicol at time=0.

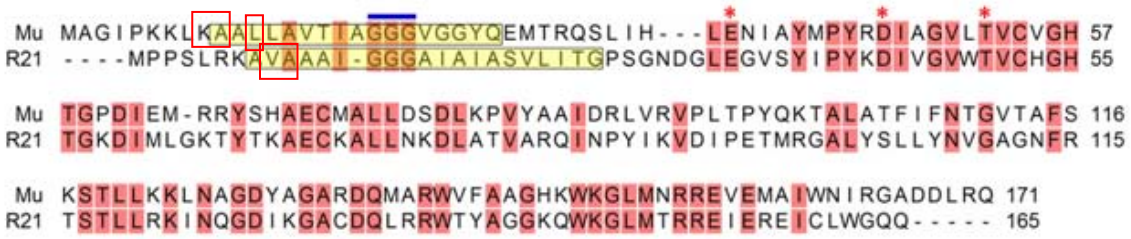

Figure 4.6 Alignment of Mu Lys and phage 21 SAR-endolysin R^{21}.
The identical residues are labeled with red rectangles. The predicted SAR-domain is highlighted in yellow. The catalytic triad (E-D-T) is indicated by red asterisks. The conserved triple Glycine (GGG) linker in the SAR-domain is indicated by a blue line. The positively charged residues (R and K) in the N-terminal cytoplasmic 'tail' are indicated with open red boxes.

The N-terminal charged residues blocks Lys function

Previously published articles reported that Lys in a R^{21}-like SAR-endolysin, which does not require the disulfide bond isomerization but the refolding of the SAR-domain to activate the enzymatic function (38). The SAR-domain of Lys contains a triple-Glycine region (Figure 4.6), which is very similar to the turn region between two parallel helices formed in the SAR-domain of R^{21} in the periplasm. This suggests that Lys might undergo the similar SAR-endolysin refolding, which results in the stabilization of the catalytic Glu (37). However, despite the homology between Lys and R^{21}, there is no evidence that R^{21} requires any phage-encoded cofactor for activation. By comparing the primary structure of Lys and R^{21}, we realized that there are more positively charged residues in the N-terminus of Lys (K6, K7, and K9) than in R^{21} (R6 and K7) (Figure 4.7A). It is known that adding positively charged residues to the N-terminus of SAR-domain abolishes its ability of escaping the IM (21, 22, 109) and thus 'locks' the protein in the membrane even when the membrane is depolarized. Therefore, we hypothesized that the lysine residues at the N-terminusof Lys are impeding the release of Lys from the IM, and thus blocking the activation of this muralytic enzyme.

To test this hypothesis, we constructed pBAD33 plasmids encoding the single missense lysine substitutions, Lys K6G, K7G, and K9G and analyzed their function with or without the presence of Gp25 (See Materials and Methods). All Lys mutants are functional as endolysin with the co-expression of Gp25 (Figure 4.7B). The K6G mutation fastened the lysis time (at ~ 40 min) compared to other mutants and wt Lys (lysis at ~ 50min). Strikingly, the expression of Lys K6G alone led to the lysis at ~ 40

min, with a lysis profile nearly identical to that of co-expression with Gp25 (Figure 4.7C). K7G mutant alone also caused lysis, but the lysis time is delayed and the decrease in the cell density was very gradual (Figure 4.7C). K9G showed identical lysis profile to the wt Lys. Lys K6G and K7G mutants are thus named as ***sel*** mutants, representing "**s**elf-dependent **e**ndo**l**ysin". The protein expression level of Lys wt and mutants were analyzed and compared by Western blotting, and no significant increase in protein accumulation was detected. In fact, K7G and K9G had slightly lower expression level compared to wt Lys. In sum, the N-terminal lysines are responsible for the dependence of Lys to Gp25, most likely by blocking the release of Lys to the periplasm.

Both TMD and the cytoplasmic domain are involved in Gp25 function

Gp25 is predicted to be a 99 aa type I integral membrane protein with a negligibly short periplasmic domain (two aa) and a large C-terminal cytoplasmic domain (74 aa). BLAST analysis indicated that Gp25 shares no homology to any protein with assigned function, although significantly related homologs were found in Mu-like prophages. The alignment of Gp25 and homologs suggested that the TMD of Gp25 is less conserved than the C-terminal cytoplasmic domain (Figure 4.8). Since we proposed that Lys is locked in the membrane by positively charged residues in its N-terminus, the highly conserved negatively charged residues in Gp25 are of interest because they could play a role in neutralizing the charges in Lys N-terminal tail to assist the release of Lys.

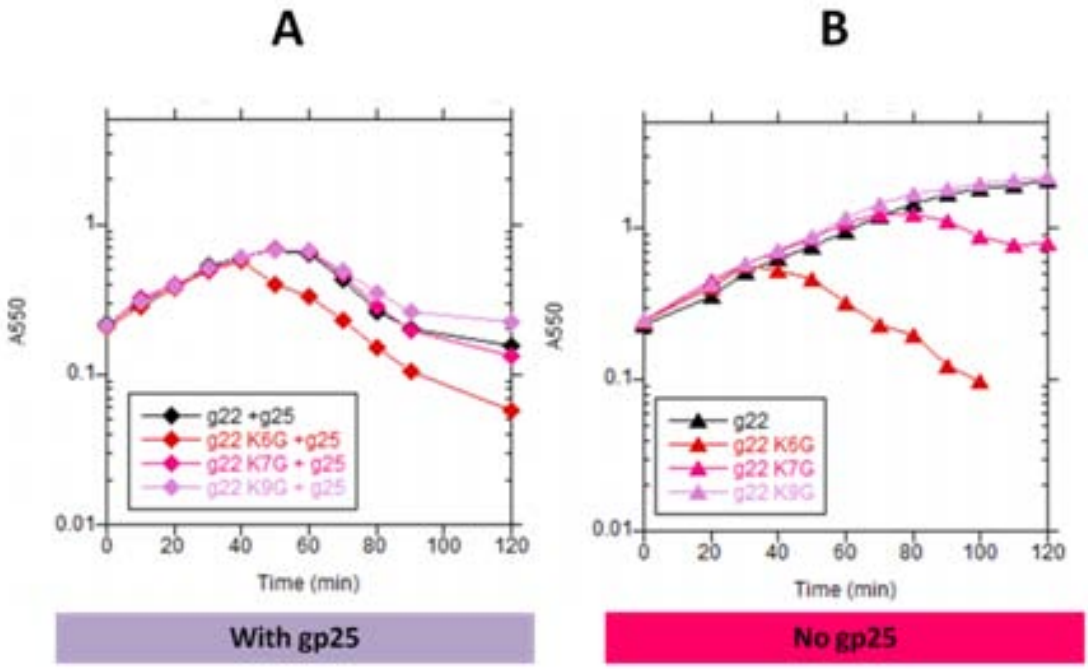

Figure 4.7 Lysis curve of Lys *sel* mutants
(A) All of the Lys mutants (K6G, K7G, and K9G) are functional as SAR-endolysin with the presence of Gp25. g22, pBAD33-g22-ssd; g25, pBAD24-g25-ssd. All samples are induced at time=0 with 0.4% arabinose. Host cell, MG1655 *tonA::Tn10*. The black line indicates the induction of wt Lys (g22) and Gp25 (g25).
(B) Lys *sel* mutants K6G (red triangle) and K7G (pink triangle) mutants can cause lysis without Gp25.

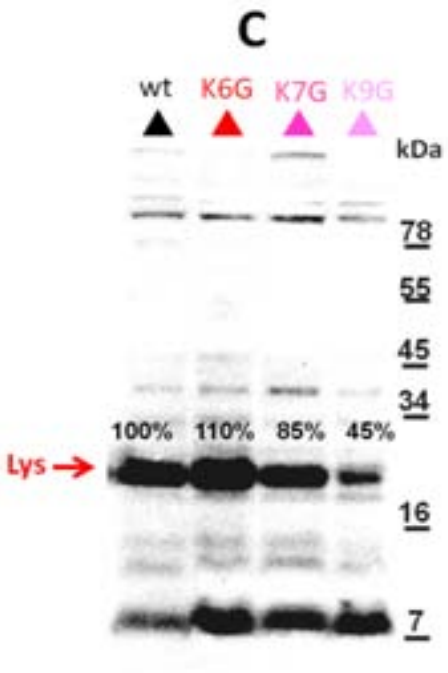

Figure 4.7 Continued
(C) Lys *sel* mutants did not change the accumulation of Lys significantly. The red arrow indicates the corresponding band of Cmyc-Lys.

In fact, mutating highly conserved negatively charged residues in the distal C-terminus of Gp25 (D71 and E91) completely abolished the activation of Lys (Figure 4.9A), suggesting the cytoplasmic domain of Gp25 is needed for its function. We also replaced the original Gp25 TMD region (residue 3-25) to an artificial TMD (VLLIIVVVVVVVVIILLI), which has been experimentally tested (41), to examine if the TMD is involved for Gp25 function. The change in TMD did not affect the membrane topology of Gp25, but blocked its function (Figure 4.9B). These results indicated that both cytoplasmic domain and TMD of Gp25 is critical for its function as a cofactor of SAR-endolysin Lys.

Discussion

Prior to this work, *lys* was the only gene that have assigned with a host lysis-related function in Mu genome. Here we have reported the identification of other important genes the Mu lysis cassette: the holin gene *19*, the antiholin gene *20*, the spanins pair *23/23a*, and gene *25*, which encodes a novel class of lysis regulatory factor.

Gp19 is a new class of holin. We have shown that the product of Mu *19* can halt the cell growth when expressed from a medium copy number plasmid and complement the holin $S^{21}68$ in the phage 21 lysis cassette (Figure 4.2C), but cannot cause lysis when expressed with a cytoplasmic endolysin (Figure 4.2D). Therefore, we concluded that protein Gp19 is a pinholin, which releases the endolysin by the depolarization of the IM rather than the formation of large membrane lesions. There are three phages have been reported to adopt the “pinholin and SAR-endolysin” lysis system so far: phage 21,

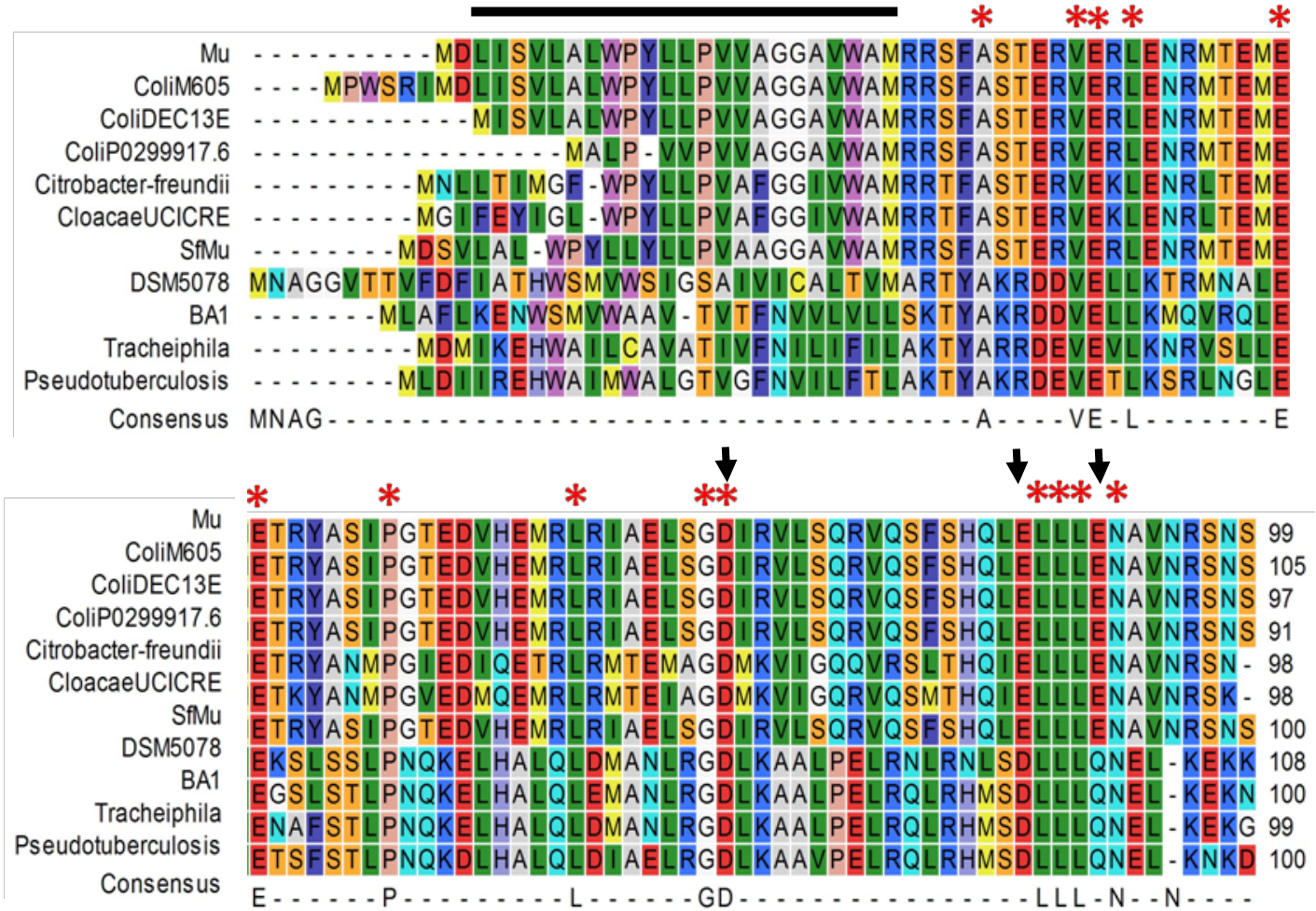

Figure 4.8 Alignment of Gp25 homologs.
The predicted TMD of Gp25 is indicated by a black line. The fully conserved residues are indicated by red asterisks. Identical homologs of Gp25 are omitted for clarity. Black arrows indicate the mutated residues.

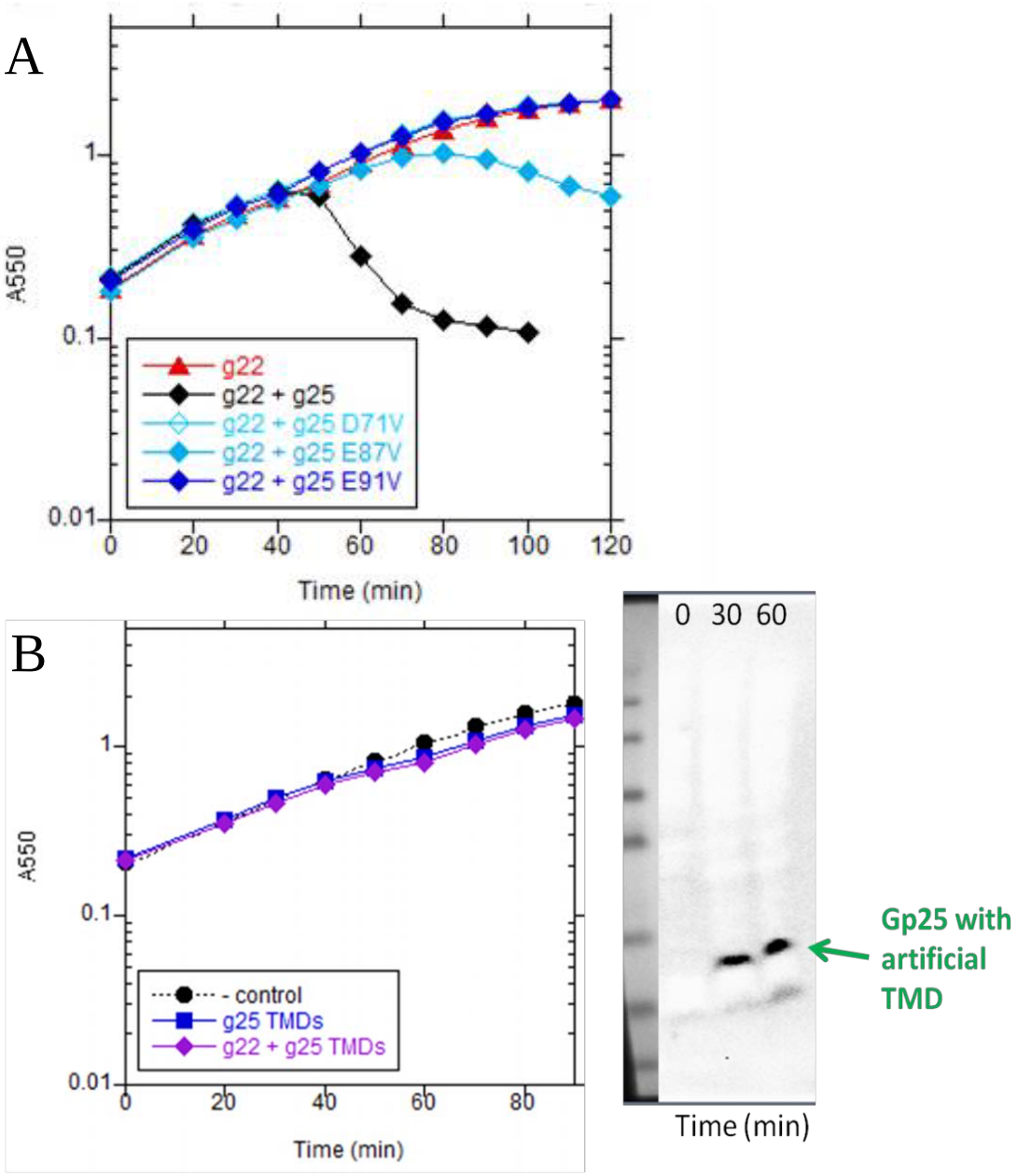

Figure 4.9 Lysis curve of Gp25 mutants.
(A) Mutation in the C-terminus abolished the function of Gp25. The location of point mutation (D71V, E87V, and E91V) in indicated by black arrows in A. The black line indicates the induction of wt Lys (g22) and Gp25 (g25). The red triangle indicates the induction of Lys (pBAD33-g22-ssd) without Gp25.
(B) Native TMD is required for Gp25 function. g22, pBAD33-g22-ssd; g25 TMDs, pBAD24-g25-TMDs, which encodes Gp25 with an artificial TMD. The WB result in the right panel shows the accumulation of Gp25-TMDs.

A

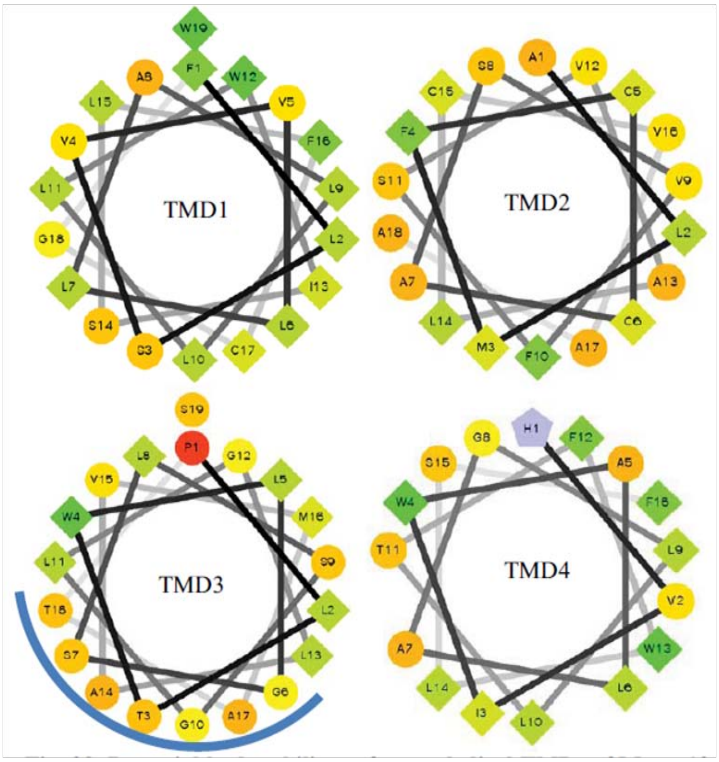

Figure 4.10 New class of holin.
(A) Potential hydrophilic surface on helical TMDs of Mu pinholin Gp19. The hydrophilic residues are shown as yellow circles, the hydrophobic residues are shown as green diamonds, the potentially negatively charged residues are shown as triangles, and the potentially positively charged residues are shown as pentagons. The blue arches indicate the predicted lumen surface (85).

B

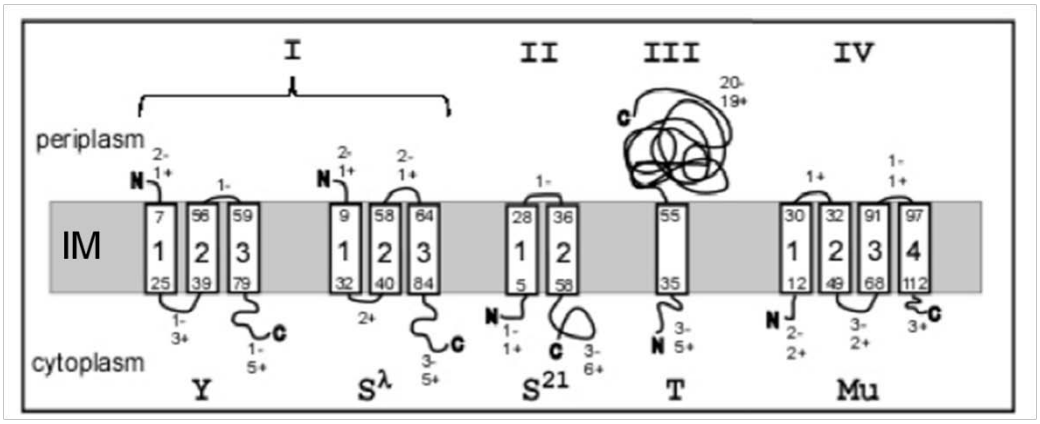

Figure 4.10 Continued.
(B) Model of the membrane topology of holins. The prototypes of each holin class were shown: Y (phage P2) and S105 (phage λ), class I; S^{21} (phage 21), class II; T (phage T4), class III; and Mu Gp19, class IV. Topology and boundary residues for TMD1, -2 , -3, and -4 are predicted by TMHMM. The first and the last aa of each TMD is indicated by residue numbers. The number of positively charged (+) or negatively charged (-) residues in the cytoplasmic or periplasmic domain are indicated (85).

phage ϕKMV, and phage TP712 (127). Their pinholins (S^{21}68, KMV44, and holTP712) are all classified as the class II holin, which has one TMD and a N-terminal SAR-domain and forms pinholes. Class II holin is synthesized as an integral membrane protein with two TMDs, then the N-terminal TMD, which is a SAR-domain, needs to exit the IM to activate the hole-forming function upon the holin triggering (22). However, Gp19 is predicted to have four TMDs (Figure 4.1C and Figure 4.10A) and does not have N-terminal SAR-domain. This topology does not fit match any experimentally-confirmed holin (Figure 4.10B). Also, as indicated by To *et al*, holins with more than two TMDs usually form large holes and have two amphipathic TMDs lining up the membrane lumens (23). In contrast, the pinholin S^{21}68, which has only one TMD remaining in the IM after the triggering, has only one TMD surface exposed to the lumen of the pinhole. The helical wheel projection of Gp19 suggested that only TMD3 has a relatively hydrophilic face (Figure 4.10A). None of other TMDs has the nature of either amphipathic helix or SAR-domain (Figrue 4.10A). Taken together, we designated Gp19 as the prototype of the class IV holin: holin that has four TMDs with an N-in and C-in topology and forms pinhole.

Gp25 is a novel factor in the dsDNA phage lysis paradigm. The data presented here showed that without functional Gp25, Mu gene *25*-mediatedlysis was blocked even with the presence of holin and SAR-endolysin (Figure 4.2A). Gp25, a 99 aa membrane protein, is predicted to contain a N-terminal TMD and a large cytoplasmic domain (Figure 4.1) and has no homology to any protein except gp*25* from other Mu-like phages.At first, Gp25 was one of the Mu holin candidates we have proposed. However,

unexpectedly, the thermal-induction of Muc^{ts} *25::Cam* lysogen resulted in a halt in culture growth and stasis, but no lysis was observed. The similarity of the lysis profiles between Muc^{ts}*25::Cam* and Muc^{ts}*lys::Cam* mutants suggested that the loss of Gp25 impeded the function of Lys. Since Lys is a SAR-endolysin, it is considered to be capable of escaping the IM in a host *sec* system-dependent nature. No phage-encoded protein was reported to be critical for the release and activation of the SAR-endolysin. However, our results showed that without Gp25, Lys did not respond to the membrane depolarization caused by holin triggering or by the adding of energy poison such as KCN. Analysis of Lys *sel* mutants revealed that the basic residues in the N-terminal tail were responsible for the crippled SAR-endolysin function of Lys (Figure 4.7). Taken together, the role of Gp25 is most likely to neutralize the effect of positively charged residues in the N-terminus of Lys and thus assist the escaping of Lys from the IM. Due to the lack of the periplasmic domain, it is very unlikely for Gp25 to interact with Lys after Lys was released from the IM. Therefore, either TMD or C-terminal cytoplasmic domain of Gp25 must be involved in the interaction with Lys, with a mechanism that is yet to be understood. The simplest idea is that the C-terminal domain of Gp25 forms a micro-environment around the extreme N-terminus of Gp22. In this microenvironment, the pK of one or both of Lys 6 and Lys7 is lowered to ~7-8, resulting in deprotonation of the Lys residue(s). Once the Lys residue is no longer charged, the SAR domain can exit the bilayer and refold to active conformation.

Our model for Mu lysis paradigm is summarized in Figure 4.11. Before the lysis time, the pinholin Gp19 is inhibited by the antiholin Gp20, and the crippled SAR-

endolysin Lys is secured in the inactive form in the IM. When both Gp19 and Gp25 are accumulated to the critical level, the pinholin Gp19 triggers and opens up pinholes in the IM, and then Lys is released from the depolarized IM with the help of cofactor Gp25. In sum, Mu is the first phage that was identified with two different regulators for two protein targets in the lysis paradigm: the antiholin Gp20 inhibits pinholin Gp19, and the endolysin cofactor Gp25 assists the activation of SAR-endolysin Lys. The most reasonable explanation for this double-regulator system is to prevent premature lysis. The future direction includes the mutational analysis of holin Gp19, the confirmation of class IV holin topology, the visualization of the Mu-hole, isolation of spontaneous Lys *sel* mutants, and the protein-protein interaction between targets and regulators in the lysis paradigm.

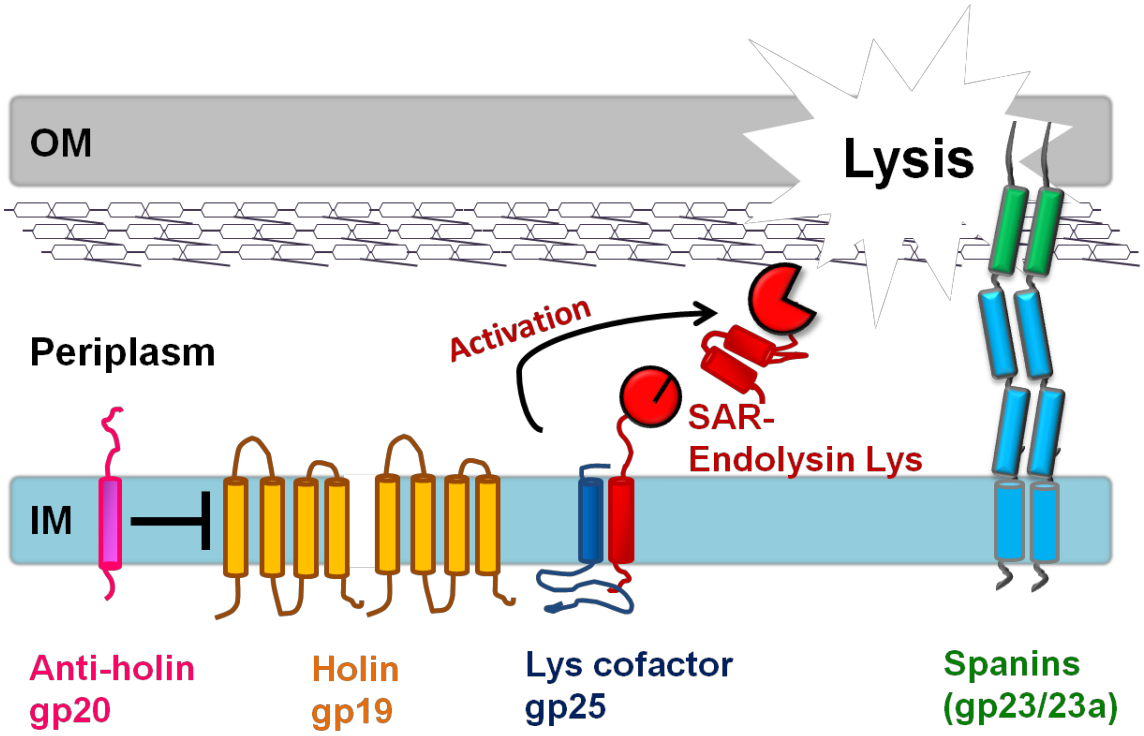

Figure 4.11 Current model of Mu lysis paradigm.
Holin Gp19 (yellow) is inhibited by antiholin Gp20 (pink) until the designated lysis time. The triggering of holin de-energizes the IM, makes proton motive force (PMF) drop to zero. Then SAR-endolysin Lys (red) is released from the IM with the help of cofactor Gp25 (blue) and degrades the PG. Finally, OM is disrupted by spanins.

CHAPTER V

N15-LIKE PHAGES DEPEND ON *MANYZ* FOR INFECTION

Introduction

Despite the re-emergence of phage as potential antibacterials (2) and as a modality for manipulation of the microbiota of the human body (128), the molecular pathways by which phage achieve specific adsorption and efficient transfer of the viral DNA into the host cytoplasm are poorly understood features of the phage infection cycle (7). Among the best studied systems for study of this pathway is the canonical siphophage λ. Phage λ has two virion features involved in specific adsorption: the central tail spike, formed by three copies of the J protein, and the four side tail fibers (STFs), each formed by three copies of the Stf and Tfa proteins (129). In the adsorption pathway, the tail spike binds specifically to the trimeric maltoporin LamB in the outer membrane (OM) of *Escherichia coli.* Indeed, only the C-terminal domain of the J protein is required for the specific LamB binding (129). The STFs interact with the major *E. coli* OM porin OmpC, also trimeric; although this interaction is dispensable, the presence of the wt side tail fibers results in a ~10-fold increase in the λ adsorption rate (130).

The steps subsequent to the final, energy-dependent, irreversible LamB binding remain obscure, in part because of decades-old genetics studies. In 1974, in the course of a selection for bacterial mutants with new host-specific restriction properties, Scandella and Arber isolated *E. coli pel* (for penetration of λ) mutants which allow normal λ adsorption but progress to a productive burst or lysogenization event only rarely (2-10%

of infections). Inductions of *pel* strains lysogenized by λ yielded lytic bursts at the normal time and with normal yield of virions, indicating that the *pel* defect is at the level of DNA entry into the cytoplasm (9). Support for this notion was obtained from experiments showing that λ DNA was not subject to host restriction in *pel* infections (9) and that, in negative-stain electron micrographs, phages adsorbed to *pel* hosts had full heads, whereas phages adsorbed to wt hosts had empty heads (9). Nevertheless, phages adsorbed to a *pel* host could not be eluted in either active or inactive form.

These early findings suggested that Pel played an important role in the penetration of the cytoplasmic membrane by the λ DNA. Genetic mapping in *E. coli* indicated that the *pel* locus coincided with *ptsM*, the locus encoding the mannose phosphotransferase system (131). Fine structure analysis using cloned *ptsM* restriction fragments indicated that the *pel* phenotype was due to mutations in *manY* and *manZ*, encoding the PTS enzyme II subunits II-A^{Man} and II-B^{Man}, respectively (131). Interestingly, the *pel* phenotype was 100% associated with defects in mannose import, but many *manYZ* import-defective mutants exhibited no Pel defect, and physiological experiments indicated that inhibition of ManYZ import function did not affect λ infectivity. These results indicated that although ManYZ was required for λ DNA penetration, its sugar import function was not (131).

Overall, these findings suggest that λ DNA penetration requires host protein complexes at each membrane: the LamB maltoporin at the OM and the ManYZ PTS complex at the IM. However, no coherent model has been proposed for the pathway of the 30 MDal λ DNA across either membrane, much less for how the two host proteins

are involved. Moreover, λ genetics, normally a powerful tool for clarifying molecular mechanisms, has in this case proved somewhat confusing. The original *pel* mutants had been isolated in multiple selections using λ*vir* and λ*cI*, and exhibited a plating deficiency of ~10^{-6} for those phages (9). The severe plating defect allowed Scandella and Arber to isolate λ suppressors able to plate on the *pel* hosts: *hp* mutants (for host-range of pel) (9). These *hp* mutants were subjected to deletion and recombinational mapping and found to map to two tail genes, *V* and *H* (10). V is the major tail protein and H, at the time known only as a minor tail component, is now known to be the tape-measure protein, which, as a trimer, occupies the central cavity of the tail tube (132). Some of the λ*hp* mutants exhibited a temperature-sensitive (*ts*) lethal plating phenotype; all of the λhp*ts* alleles mapped to *V* (10). Later, Roessner *et al* (133) reported that the H protein was ejected from the tail along with the DNA upon interaction with solubilized LamB receptor or LamB mounted in liposomes, leading to the notion that H served as a DNA pilot protein, as had been shown for ssDNA phages (134), deploying from the tail after its irreversible attachment to LamB and interacting directly and specifically with the ManYZ in the IM.

Here we report results from initial experiments aimed at resolving these issues and addressing generally the molecular mechanism of siphophage DNA penetration. The results, which include the isolation and characterization of novel N15-like phages dependent on *manYZ*, are discussed in terms of the current model for this paradigm phenomenon.

Materials and methods

Bacterial growth and induction

Phages and bacteria strains used in this research were listed in Table 5.1. Standard LB-agar plates supplemented with antibiotics (ampicillin, 100 μg mL-1; chloramphenicol, 10 μg mL-1; Kanamycin, 40 μg mL-1) were used for plating bacteria. Plates were incubated at 37°C unless otherwise indicated. Overnight cultures and fresh cultures were prepared as described in previous chapters. The cell mass was estimated by measuring absorption at 550 nm wavelength using a Gilford Stasar III sipping spectrophotometer (Gilford Instrument Inc, Oberlin, OH).

PCR, plasmid construction, and bacterial genome modification

See Table 5.1 for plasmids used and constructed in this study. Standard DNA manipulations and sequencing were conducted as described in previous chapters. Plasmid pMLB113-*manY* and pMLB113-*manZ* were constructed by inserting gene *manY* and *manZ* cloned from the genome of *E. coli* MG1655 into pBR322 based, IPTG-inducible medium copy-number plasmid pMLB113 between its EcoRI and HindIII sites. The insertion was confirmed by plasmid DNA sequencing service provided by Eton Bioscience (San Diego, CA). To delete gene *manX, manY, manZ, lamB,* or *tonB* from the MG1655 genome, P1 transduction using the phage P1 virulent strain, P1vir, was performed as described. The Keio strains (135) with desired genes replaced with a KanR marker were used as the donor. The deletion of the gene in MG1655 was confirmed by PCR.

Table 5.1 Phages, strains, and plasmids used in Chapter V.

Phages	**Description**	**Source**
λ105	λ *cI857*	Laboratory Stock
λ20	λ *b221 cI-*; contains ~10kb deletion in the chromosome	Laboratory Stock
λ128	λ *b221 cI-*; contains ~5kb deletion in the chromosome	Laboratory Stock
N15	Wild type N15	Laboratory Stock
phi80	Wild type phi80	Laboratory Stock
Scandella	Coliphage Isolated from sewage sample	This study
SirWerner	Coliphage Isolated from sewage sample	This study
Bacteria Strains	**Description**	**Source**
BW25113	*E. coli* Keio correction parental strain	(135)
BW *manY::Kan*	*E. coli* Keio manY::Kan mutant	(135)
MG1655	*E. coli* F- λ- *ilvG- rfb-50 rph-1*	Laboratory Stock
MG1655 *tonA*::Tn10	*tonA* knock-out mutant by insertion of Tn10	Laboratory Stock
MG1655 *tonB*::Kan	tonB replaced by KanR gene; made by P1 transduction using Keio correction as the donor	This study
MG1655 *lamB*::Kan	lamB gene replaced by KanR gene	This study
MG1655 *manX*::Kan	manX gene replaced by KanR gene	This study
MG1655 *manY*::Kan	manY gene replaced by KanR gene	This study
MG1655 *manZ*::Kan	manZ gene replaced by KanR gene	This study
Plasmids	**Description**	**Source**
pMLB113	pBR322 origin, IPTG inducible plasmid. AmpR.	(49)
pMLB113-ManY	pMLB113 carrying manY gene from MG1655	This study
pMLB113-ManZ	pMLB113 carrying manZ gene from MG1655	This study

Phage infection, mutagenesis, and sequencing

Phage preparations were made by the "plate lysate" method, as described previously. Briefly, 100 μL of original phage lysates were mixed with 100 μL of the permissive strain cultures, and then added to the LB top agar (0.75% agar) and immediately poured onto the standard LB plates. After incubated overnight at 37°C, the top agar layer was disrupted and collected into a 15 mL Falcon tube and mixed with 5 mL LB medium containing 1% $CHCl_3$ (v/v). The mixture was then cleared by centrifugation at 5,000 x g and the supernatant was filtered through a 0.22 mm syringe filter. To observe plaque morphology, phage infection experiments were carried out as described (19, 21). Five μL of serial-diluted phage lysates were spotted onto the lawn of host cells to determine the PFU (plaque forming unit). The efficiency of plating (EOP) was calculated as described. Phage particles were concentrated and purified by CsCl-gradient ultracentrifugation following the standard protocol provided by the Center for Phage Technology (College Station, Texas, https://cpt.tamu.edu/).

Isolation of manY*-dependent phages*

To isolate *manY*-dependent coliphages, liquid sewage samples collected from College Station wastewater treatment plant (College Station, TX) were centrifuged and filter-sterilized before applying on the lawn of BW25113. Each single plaque was picked, purified and the resulting lysate was used to spot onto BW 25113 *manY::Kan,* which is a *manY* deletion mutant from the Keio collection (135) . The lysates that grew on the parental Keio collection strain (Table 5.1) but failed to form plaques on the lawn of BW25113*manY::Kan* were collected, purified, propagated, and visualized by

negative-stain electron microscopy at the Microscopy & Imaging Center at Texas A&M University (College Station, TX). DNA sequencing was conducted as described (136). Phage *hp* mutants were isolated either as the spontaneous plaque-forming suppressors appearing on the lawn of the *manYZ* deletion strain, as described (9) or isolated after UV irradiation (137). To obtain the genomic sequence of newly isolated phages, the phage DNA was extracted using Wizard® Genomic DNA Purification Kit (Promega, Madison, WI) and sequenced using the high-throughput sequencing method provided by 454 Life Sciences (Branfold, CT). Quality-controlled trimmed reads were assembled to a single contig at 15.2-fold coverage using Velvet version 1.2.10 (138). Phage genes were predicted using GeneMarkS (139) and corrected using software tools available on the Center for Phage Technology (CPT) Galaxy instance (https://cpt.tamu.edu/galaxy-pub/) .

Complementing the deletion of ManYZ from proteins expressed in trans

The complementation assay was carried out as described (140) with modification. The MG1655 cells harboring pMLB113-manY or pMLB113-manZ were mixed with top agar and poured onto the LB plates supplied with appropriate antibiotics. Even though pMLB is an IPTG-inducible plasmid, no inducer was used in those LB plates because the basal level of ManYZ expression due to the leaky promoter P_{lac} from the plasmid is sufficient for the phenotypic analysis (140). The *manYZ*-dependent or independent phages were then spotted on the top agar containing the host bacteria. All plates were incubated at 37°C degree overnight.

Results

Isolation of novel ManY/ManZ dependent coliphages

A filtered sewage sample from a local source was plated onto lawns of *E. coli* BW25113, which is the parental strain of the Keio Collection (131), and plaques picked and tested against BW25113*manY::Kan*. Two isolates that formed plaques on the lawn of BW25113 but not BW25113manY::Kan were propagated on the permissive host, purified by CsCl gradient ultracentrifugation, and visualized by negative-stain electron microscopy (EM). EM images revealed that both phages were λ -like siphophages with long, flexible tails of ~100 nm long, and isometric capsids of a size indicative of T=7 quasi-equivalent symmetry (Figure 5.1). Even though these two phages have nearly identical morphologies, they formed distinct plaques on the lawn of BW25113. One phage formed clear plaques with ~1 mm diameter, the other one formed turbid and smaller plaques (Figure 5.2A). We named the clear-plaque forming phage Scandella, and the turbid plaque forming phage SirWerner, from the authors of the original *pel* report: Dorothea Scandella and Werner Arber (9)

The genome of Scandella and SirWerner were extracted and sequenced. The results of genomic analysis suggested that Scandella was most closely related to the lambdoid coliphage N15 (97% identity over 86% of the genome) (Figure 5.3). SirWerner was highly similar to N15 as well, especially the head and tail morphogenesis genes, but also shared exhibited extensive mosaic similarities with coliphage λ and lambdoid coliphage phi80 (89% identity to phi80 over 58% of the genome), notably in the DNA

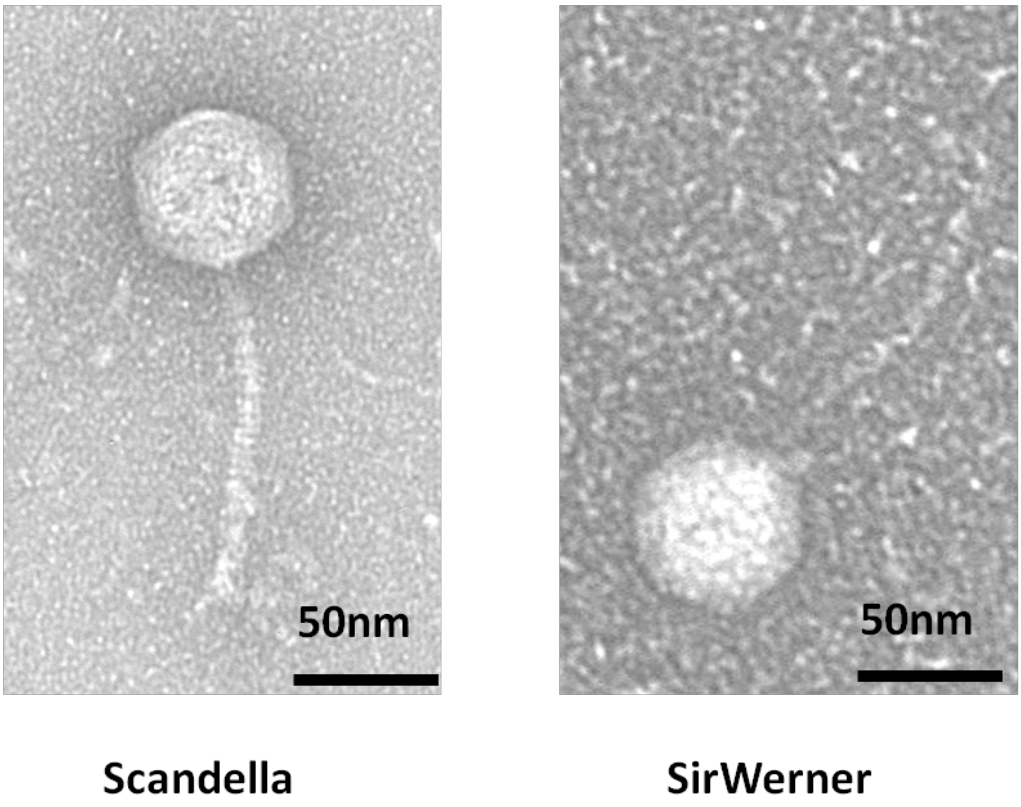

Figure 5.1 Electron microscopy image of phages.
Scandella (left) and SirWerner (right).

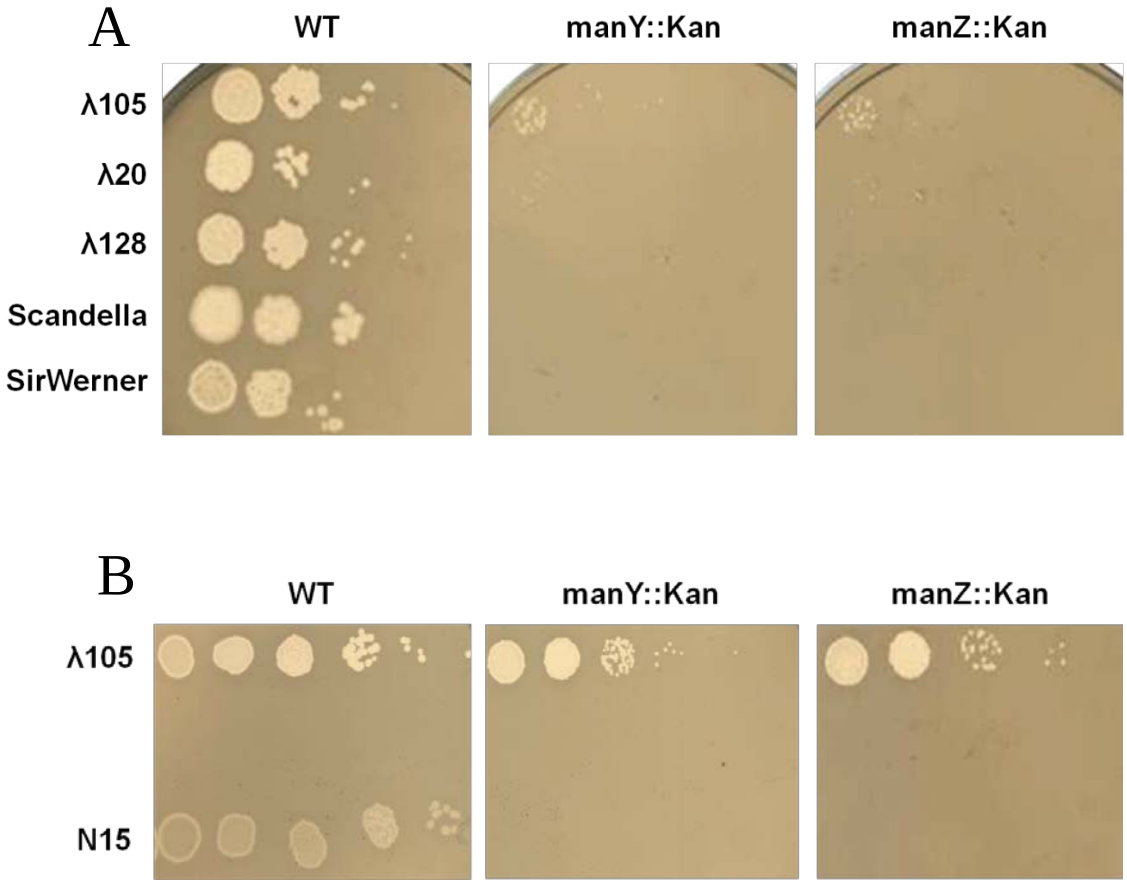

Figure 5.2 Plating morphology of various lambdoid phages.
Phages are plated on MG1655 and MG1655 Δ*manYZ* strains.
(A) λ 105, 20, 128, Scandella and SirWerner.
(B) λ 105 and N15.

replication, late gene regulation and lysis gene regions (Figure 5.3). Interestingly, in SirWerner the similarity showed a dramatic breakpoint at the C-terminal end of the J tail-spike protein (Figure 5.3, red arrow), after which it became highly similar to the J protein of phi80 (data not shown).

ManYZ complex is required for the infection of N15-like phages

To provide a more reliable host background for testing plating efficiency or Scandella and SirWerner, we deleted *manX*, *manY*, and *manZ* genes from the *E. coli* MG1655 genome using P1 transduction. Using these MG1655 mutant strains, we observed 7 to 8-log reduction in the EOP of Scandella and SirWerner on *manY* and *manZ* strains, but not on *manX* strain (Figure 5.2 and Table 5.2). These results confirmed that both Scandella and SirWerner require *manY* and *manZ* for forming plaques on *E. coli*, but do not need *manX*. In WT *E. coli*, ManY and ManZ form a complex in the IM as part of the mannose permease (141). Although all three proteins encoded from *manXYZ* genes are required for the sugar transport and phosphorylation, a previous study (141) suggested that the cytoplasmic protein ManX is not required for the DNA penetration of phage λ, probably because ManX does not affect the structure of the ManYZ complex in the IM. To test if the expression of ManYZ protein *in trans* can restore the plating efficiency, we have cloned *manY* and *manZ* separately in a medium copy number plasmid pMLB113. When pMLB113-*manY* or pMLB113-*manZ* was transformed into *manY* or *manZ* strains, respectively, the basal expression level from the P_{lac} promoter is sufficient to restore the EOP for Scandella and SirWerner (Table 5.2). However, the expression of ManY in trans did not rescue the deletion of

Table 5.2 Plating efficiency of phages on ManYZ deletion strains.

	Scandella	SirWerner	λ105	λ20	λ128	N15	phi80
Outer Membrane Receotor	?	*fhuA*	*lamB*	*lamB*	*lamB*	*tonA*	?
IMR	*manYZ*	*tonB, manYZ*	*manYZ*	*manYZ*	*manYZ*	*manYZ*	?
E.O.P (compare to MG1655 WT)							
MG1655	1	1	1	1	1	1	1
MG1655 manX::kan	0.8	1.7	1.3	1.0	0.9	N/A	N/A
MG1655 manY::kan	$4.2x10^{-7}$	0	0.1	$6.3x10^{-5}$	$8.1x10^{-6}$	0	1.01
MG1655 manZ::kan	$2.5x10^{-8}$	0	0.21	$2.1x10^{-5}$	$5.7x10^{-5}$	0	0.97
MG1655 lamB ::kan	2.2	2.1	0	0	N/A	N/A	N/A
MG1655 tonB::kan	1.4	0	N/A	N/A	N/A	N/A	N/A
MG1655 manY::kan pMLB-manY	2.0	1.7	N/A	N/A	N/A	N/A	N/A
MG1655 manZ::kan pMLB-manZ	1.8	1.0	N/A	N/A	N/A	N/A	N/A

manZ (data shown). These results suggested the complex of ManYZ is required for the infection of Scandella and SirWerner.

Since the genomes of both Scandella and SirWerner share similarity to N15 and phi80 genome, we also plated those phages, neither of which have been reported to rely on any IM receptor for infection, on *manYZ* deletion strains. N15 did not plate on the *manY* host, as expected (Figure 5.2). However, the EOP of phi80 was not affected by the loss of *manY* (Table 5.2). We also plated three different λ phages with various genome sizes on the *manYZ* deletion hosts. λ105, which has a genome size nearly identical to the parental strain λpapa, showed only ~10 fold reduction in EOP. This result is replicable with other λ strains that have genome size similar to λ papa (data not shown). However, λ20 and λ128 showed much more significant (~10^{-5}) EOP reduction, which is very similar to what was reported for full-length λ in the original *pel* report (9). The linkage between the dependence on *manYZ* and the size of the phage genome remains mysterious.

Scandella and SirWerner hp mutants carry mutation in their H genes

To understand which gene is responsible for the dependence of ManYZ complex, we isolated spontaneous plaque-forming *hp*suppressors (*hp* = host range of pel) on *manY* or *manZ* strains. The mutant frequency was ~10^{-9} plaque-forming mutants per input phage for Scandella plated on *manY*- or *manZ*- strains, and was ~10^{-10} for SirWerner plated on *manZ*- strains. The Scandella *hp* mutants isolated on MG1655 *manY::Kan* or MG1655 *manZ::kan* were named ScanYhp or ScanZhp, respectively. The SirWerner hp mutants isolated on MG1655 *manZ::kan* were

named WerZhp. We were not able to isolate any spontaneous SirWerner hp mutants on *manY* hosts. However, WerYhp was isolated from a SirWerner lysate which had been subjected to UV mutagenesis and plated on MG1655 manY::Kan. All *hp* mutants were capable of forming plaques on *manYZ* deletion strains with very high efficiency (Table 5.3). Interestingly, *hp* mutants isolated on the *manY*- strain can infect *manZ*- strains, or vice versa. This indicates that *hp* mutants can completely bypass the requirement for ManYZ complex for infection.

The genome of four different *hp* mutants were sequenced and annotated. In contrast to the original *pel* article that reported *hp* mutants carried mutations in both tape measure gene protein gene *H* and the tail tube gene *V*, all of our spontaneous hp mutants only carry mutations in their H genes (Figure 5.4 and Table 5.4). WerYhp, which is mutated by UV irritation, carries a mutation in both the *H* gene and the tail cap gene *U* (Table 5.4). Surprisingly, all of four mutations in *H* gene were mapped from codon 298 to 313, which is inside the coding region for a predicted alpha-helix (Residue 297 to 315) in H protein. The concentration of *hp* mutations in this region strongly suggested that this alpha-helical structure plays an important role in the dependence of IM receptor for phages. The 297-315 is fully conserved between Scandella, SirWerner and N15, but diverse in λ and phi80 (Figure 5.4). Although this might explain why phi80 does not require *manYZ* for the infection, it cannot explain the variation of *manYZ* dependency in the λ genome- size variants (λ105, λ20, and λ128), since there is no difference in the *H* genes based on our sequencing results (data not shown).

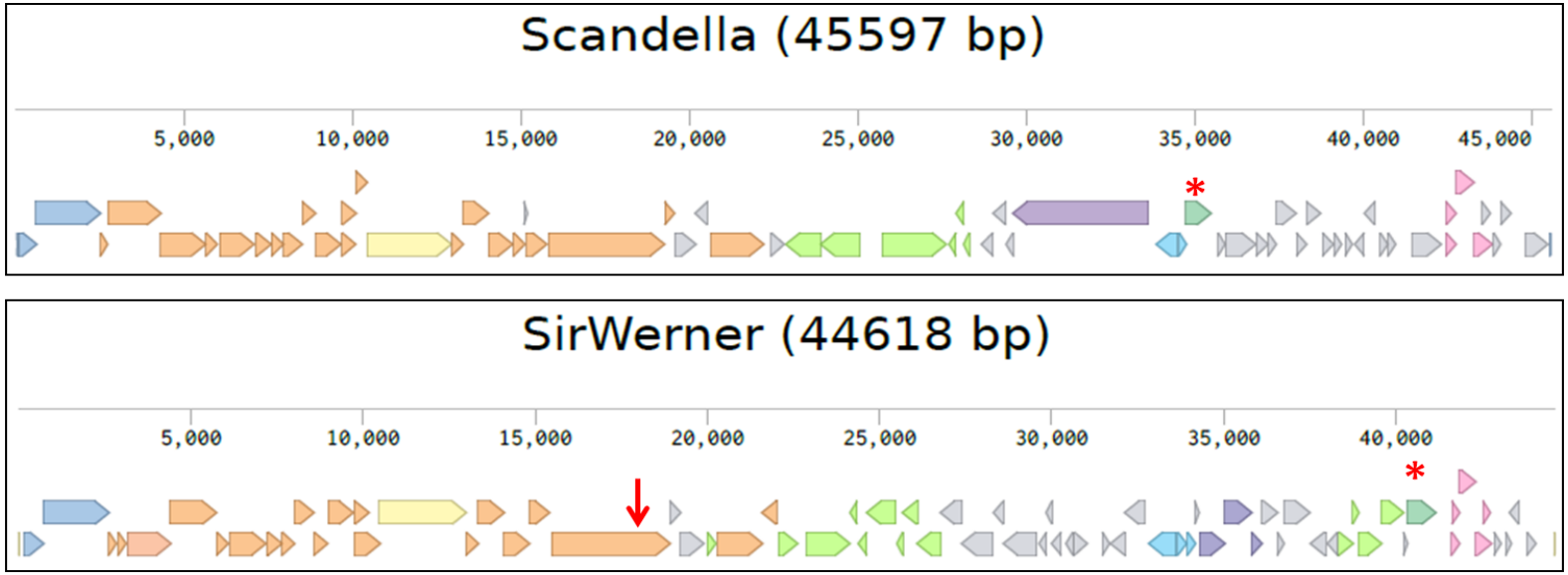

Figure 5.3 Schematic representation of the Scandella, SirWerner, λ , and N15 genomes.
Each open reading frame is represented by a pentagon with the following color schemes: Orange, phage head and tail; Yellow, the tape measure protein H; Purple, DNA replication; Cyan, repressors; Blue, terminases; Pink, lysis; Light Green: DNA integration, recombination, modification, and plasmid partitioning. Anti-terminator Q homologs are labeled in dark green and marked with red asterisks. The red arrow indicates the C-terminus of SirWerner J protein.

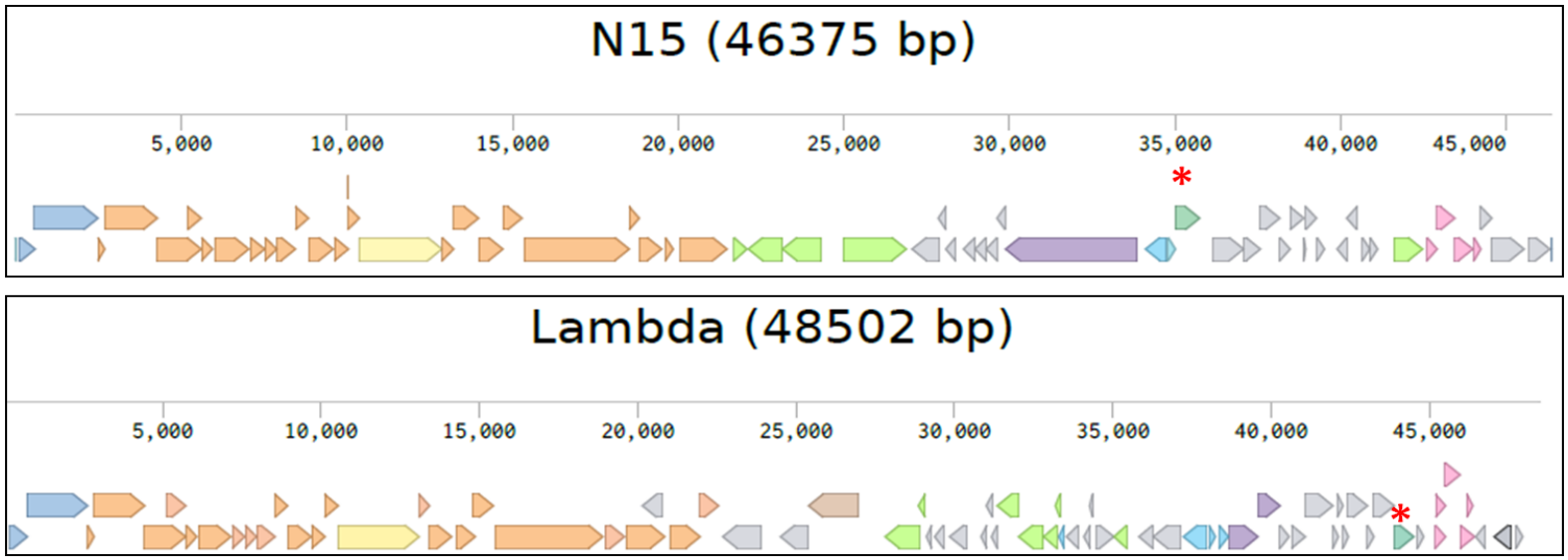

Figure 5.3 Continued

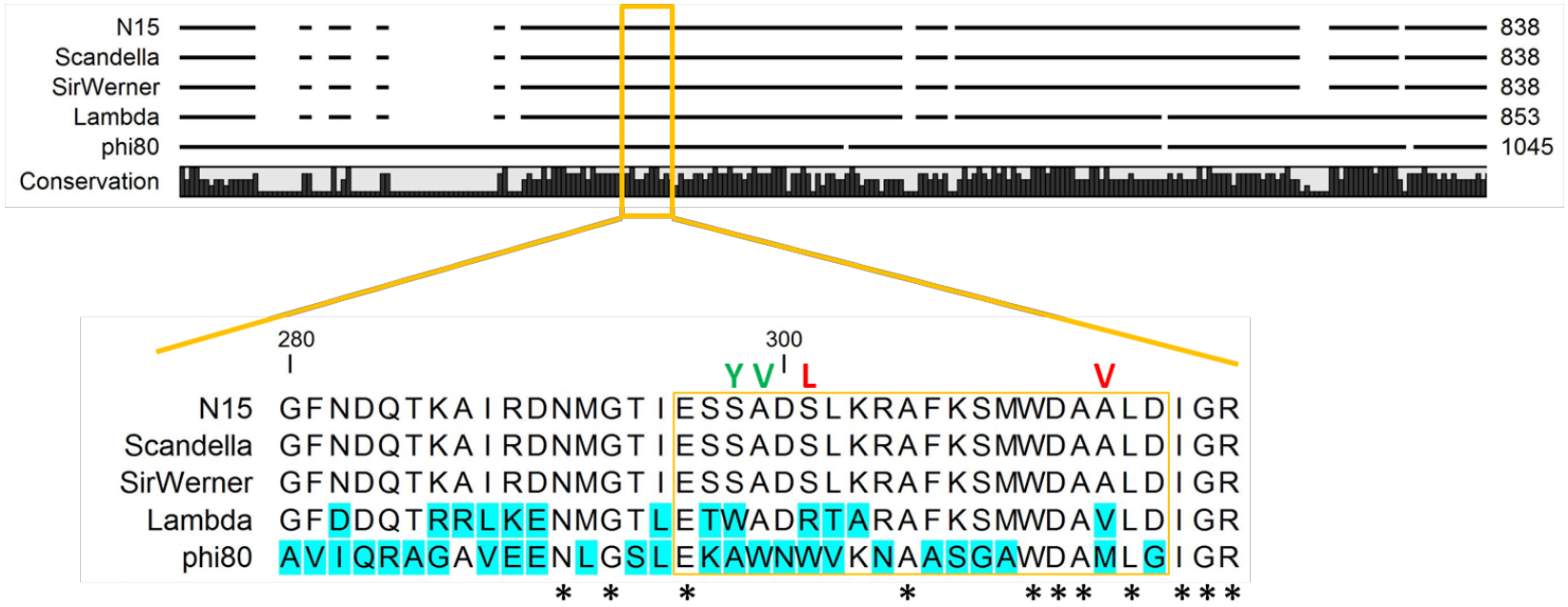

Figure 5.4 Alignment of λ, Scandella, SirWerner, N15, and phi80 H protein.
The length of each protein is indicated. Residues 280-319 of H proteins in N15-like phages are enlarged, and the predicted alpha-helix region is labeled with a yellow box. The fully conserved aa residues are indicated by black asterisks. The diverse residues were highlighted in cyan. The mutated residues in *hp* mutants are indicated above the alignment. Green, Scandella *hp* mutants; Red, SirWerner *hp* mutants.

Table 5.3 Plating efficiency of *hp* mutants on ManYZ deletion strains.

Host	Scandella WT	ScanYhp1	ScanZhp1	SirWerner WT	WerYhp1	WerZhp1
E.O.P (compare to MG1655 WT)						
WT	1	1.0	0.8	1	1.3	0.8
manY::kan	$4.2x10^{-7}$	0.9	1.3	0	0.7	0.9
manZ::kan	$2.5x10^{-8}$	0.9	0.9	0	0.8	1.0

Table 5.4 Plating efficiency of *hp* mutants on ManYZ deletion strains.

	ScanYhp1	ScanZhp1	WerYhp1	WerZhp1
Parental	Scandella	Scandella	SirWerner	SirWerner
Selection Host	*manY::kan*	*manZ::kan*	*manY::kan*	*manZ::kan*
Origin	Spontaneous mutation	Spontaneous mutation	UV mutagenesis	Spontaneous mutation
Mutation	*H* gene TCC (Ser) to TAC (Tyr) codon 301	*H* gene GCT(Ala) to GTT(Val) codon 313	*H* gene TCA (Ser) to TTA (Leu) codon 298 U gene ATG (Met) to ATA (Ile), codon 103	*H* gene GCG (Ala) to GTG (Val), codon 299

Discussion

The dependence of λ DNA injection on the host IM receptor ManYZ was first reported in 1970s (9). However, contemporaneously with the work of Scandella and Arber, a *pel* mutant designated as GL1 was accidentally isolated by the group of S. Emmons while working with λ variants with different size chromosomes. On the GL1 host, the λ plating defect was dependent on the size of the phage chromosome; so full-length λcI plated normally, but the plating efficiency dropped logarithmically with the decreasing size of the chromosome of λ deletion mutants, reaching 10^{-6} host for deletions of >20%. Subsequently, these findings on the chromosome-length dependency were confirmed and generalized for other *pel* alleles by Elliott and Arber (142) despite the earlier report of 10^{-6} plating defects (for full-length λ) that allowed the selection and mapping of the *hp* mutants. In fact, new *pel* selections were done successfully, using λ*b221*, carrying a 22% deletion (i.e., a 36 kb chromosome versus 45.6 kb for full length λ). How the original *pel* isolations and the subsequent selections for the λ hp *V* and *H* suppressor mutants were achieved using full-length λ phage has not been resolved. Our results suggested that λ20 and λ128 (deletion for 20% and 10% of the genome, respectively) showed significant decrease in EOP (Figure 5.2). The *H* gene from all λ strains used in this study was sequenced but no mutation was found (data not shown). Complete genome sequencing is needed to compare the differences between λ 105, λ 20, and λ 128.

λ, N15, Scandella and SirWerner all require ManYZ as their IM receptor, but their OM receptors vary. The interaction between λ and its OM receptor LamB (MalL)

has been extensively studied with single-molecule and single-cell methods based on fluorescent tags (141). It was found that LamB is positioned in unexplained spiral patterns along the long axis of the cell and is also found at the poles (143) Moreover, a substantial portion of LamB labeled in this way exhibited two-dimensional mobility within the spiral pattern. Real-time fluorescent video microscopy revealed that both λ phage tails and complete virions move about the surface of the host in a LamB-dependent fashion before usually localizing at the cell pole and, in the case of the complete virions, effecting the transfer of the phage DNA into the cytoplasm. It should be noted that these studies were done with the normal "laboratory" λ, which is now known to be λ*stf* , rather than with true wt λ, so it is unknown what the effect of the Stf tail fibers would be on the path, extent and velocity of the two-dimensional walk observed. In contrast to λ, N15 and its close relative SirWerner utilize the OM ferroporin TonA (FhuA) for adsorption (Table 5.2). As part of the outer membrane ferrichrome transport system, TonA forms the OM transporter for ferrichrome-iron with the IM protein TonB (TonA 2009). TonB is also required for the TonA-dependent phage adsorption, and is thought to provide free energy for the opening and closing of the TonA porin. The OM receptor for Scandella is unknown, but we have shown that Scandella does not require any of the common phage OM receptors: LamB, TonA, OmpA, OmpC, OmpF, or PhoE, since Scandella grew normally on Keio *lamB, tonA, ompA, ompC, ompF,* or *phoE* mutants (data not shown). Therefore, the requirement of ManYZ complex seems to be independent of the choice for the OM receptor for the lambdoid phages we have examined in this study.

A detailed model for the dependence of λ on ManY has been elaborated by Edgar *et al* (143) who addressed the λ infection pathway using a variety of single-molecule and single-phage fluorescence methods. These authors confirmed that λ infections occurred nearly always at the cell poles and that λ DNA could be localized to the polar cytoplasm immediately after the onset of infection. Importantly, they reported that a GFP-ManY chimera also localized to the poles of *E. coli*. Based on these findings, a model was proposed in which λ binds to the LamB maltoporin, after which the λ -LamB complex migrates to the cell pole where it interacts with other host proteins, leading to ManY-dependent DNA ejection into the polar cytoplasm. While this model is attractive, it does not address how the phage-LamB complex finds its polar destination, how it "senses" the presence of ManY in the polar cytoplasmic membrane, how H and V are involved with the DNA penetration function of ManY despite the obvious topological disparities, or how decreasing the size of the λ gDNA makes ManY essential. Thus the penetration of the DNA from the canonical siphophage l still lacks a clear conceptual framework, despite the paradigm status of the phage and the well-known facility of its genetics.

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTIONS

Host lysis and the regulation of the lysis are of interests to researchers for decades, almost ever since the discovery of bacteriophage (17). In this dissertation, different regulatory strategies targeting different factors in the lysis paradigm are discussed in two classic myophages: T4 and Mu.

New finding in T4 Lysis Inhibition

Lysis Inhibition, or LIN, the unique feature of T4 for controlling lysis timing, is considered as one of the fundamental questions of modern molecular biology (53). Despite being first reported by future Nobel Laureat A. D. Hershey in 1946, the function of *r* genes and the molecular basis of LIN were surprisingly poorly understood until very recently. Works from this laboratory by Ramanculov, Tran, and Moussa have elucidated that the binding of antiholin RI to holin T is crucial for the inhibition of T hole-formation (34, 57, 70). The most up-to-date model of LIN shows that RI, which is synthesized as an integral membrane protein with an N-terminal SAR-domain that allows RI to undergo spontaneous integration into and subsequent escape from from the inner membrane. The SAR-domain also contributes to the instability of RI in the periplasm, as RI is quickly degraded by *E. coli* periplasmic protease DegP after being released from the IM. However, when RI receives the ‘signal’ of superinfection sent from the secondary infecting T4 or T-even phages, it is stabilized and binds to the periplasmic domain of T. If the structures observed in crystals reported by Kuznetsov et

al. reflect biologically meaningful complexes, the LIN complex involves T-RI-RI-T hetero tetramers (55). This model gives RI a very important role in LIN, not only as an inhibitor of holin T, but also as a detector of superinfection signal to decide the length of the infection cycle in response to the amount of available host cells in the environment.

Even though the proposed RI-T LIN model is attractive, it has been challenged by the fact that another *r*-gene, *rIII*, seems to be crucial for LIN in addition to RI (64, 67). The work described in this dissertation has provided molecular evidence for the following new findings: 1) The product of gene *rIII* is a stable, highly alpha-helical soluble protein which is mostly monomeric but also forms homo-dimers *in vivo* and *in vitro*; 2) Together with the main holin RI, RIII can significantly extend the LIN period in a T-specific fashion; 3) RIII binds to the cytoplasmic N-terminus of T in an allele-specific manner. These findings led to a conclusion that RIII is an antiholin that inhibits the holin by direct protein-protein interaction.

RIII, the cytoplasmic antiholin

Two types of antiholins distinguished by the purpose of the holin inhibition have been studied previously: The first type, represented by λ S107, blocks holin function to prevent premature lysis. The other type, which used to have RI as the only member, regulates the lysis time to maximize the yield of phage progenies in a shortage of host cells. Nevertheless, two common features shared by all of these antiholins are 1) Direct protein-protein interaction with the holin; and 2) having at least one TMD. It is not surprising that antiholins require a TMD, since their interaction partners, holins, are integral membrane proteins (17). However, results presented in Chapter II and III of this

dissertation showed that RIII does not have any membrane-associated domain. These results suggested that RIII is the first soluble antiholin.

The presence of the 34 aa cytoplasmic domain of holin T (nT), which is the longest cytoplasmic domain among all of experimentally studied holins, is what makes the unique cytoplasmic antiholin-holin interaction possible. The expression of his-sumo-tagged nT from the plasmid led to the formation of *rIII*-type plaques in wt T4 infection, suggesting that the excess amount of nT titered out the RIII protein produced in the T4-infected host cells (Figure 2.7). The mutational analysis of T revealed that nT is required for the holin function (6), and the prediction of amphipathic alpha-helix in nT indicated that nT is involved in the membrane lesion formation (54). Further experiments, possibly including work with the nT peptide in vitro, will be necessary to evaluate this model. The interaction surface between RIII and T is also unclear. Out of 11 RIII non-functional mutants, as summarized in Bruch *et al* in 2011, six of them are in the last 10 aa of the protein. Interestingly, mutation (aa change or deletion) of the very last residue, Arg82, abolished the function of RIII (Figure 2.4). Bacterial two hybrid assays suggested that G24D mutation, which is the only LIN-defective allele in the N-terminal region, not only abolished the interaction with nT but also completely blocked the RIII self-interaction (Figure 2.5). These data indicate that RIII has at least two different functional domains for dimerization / oligomerization or holin interaction. For nT, there is no evidence that any *rV* mutant, isolated as alleles that retain holin function but does not respond to antiholin) was identified to be RIII insensitive (56, 59, 60). Screening for nT mutants that do not interact with RIII, or RIII mutants that does not bind to nT but still

retains self-interaction ability using bacterial two hybrid will permit the identification of the interaction sites.

A new factor in the lysis paradigm

Chapter VI has presented the preliminary results in the course of a study for identifying protein factors in the phage Mu lysis cassette. Besides the SAR-endolysin Lys (40), this study revealed that Mu uses a pinholin Gp19 and an antiholin Gp20 for controlling the timing of membrane depolarization, just like other dsDNA coliphages except that the pinholin has an unprecedented membrane topology with four TMDs (Figure 4.8). However, these studies also revealed that the Mu SAR endolysin Lys (Gp22) cannot be activated without the expression of an integral membrane protein Gp25. Gp25 does not share homology to any protein with known function. The knock-out of gene *25* in Muc^{ts} lysogen led to the blockage of the lysis with a lysis profile very similar to that of Mu *lys* mutant (Figure 4.3). Unlike other experimentally tested SAR-endolysins, the expression of Lys alone does not lyse the cell unless being co-expressed with Gp25 or being expressed in a ~50x higher level. The expression of Gp25 with Lys can lyse the cell without changing the accumulation level of Lys. These data strongly suggested that Gp25 plays a role in assisting the activation of Lys other than regulating the expression level of Lys. Therefore, Gp25 is considered as the novel factor in the classic lysis paradigm that regulates SAR-endolysin.

The question regarding to the nature of Gp25 function is however, not solved yet. As shown in Figure 4.5, positively charged residues in the N-terminal cytoplasmic ‘tail’ crippled the SAR-endolysin function of Lys, likely by blocking the release of Lys from

the IM. Since Gp25 does not have any periplasmic domain, it is very unlikely that Gp25 is capable of activating Lys via affecting its conformational change in the periplasm. Taken together, the most reasonable and straightforward hypothesis is that Gp25 activates SAR-endolysin by assisting the release of Lys from the membrane. Based on this hypothesis, it is important to know whether Gp25 binds to Lys directly or not. Interestingly, as a protein with a highly charged cytoplasmic domain, Gp25 has no lysines among its 24 charged aa (11 positively charged, 13 negatively charged). In fact, arginine is the most abundant aa (11%) in Gp25. This feature makes it difficult for demonstrating the interaction between Gp25 and Lys using the most widely-used crosslinking methods, which require Lys residues for covalent bond formation. Other methods such as yeast or bacterial two hybrid assays, immunoprecipitation or other pull-down type assays may have to be used.

Another question yet to be answered is why Mu needs to have an additional protein factor in the lysis paradigm to control Lys. Neither Muc^{ts} *Lys::Cam* nor Muc^{ts} *25::Cam* formed plaques on the lawn of *E. coli* host, as expected. According to the data presented in Figure 4.5, a single aa change in the Lys N-terminal tail (*sel* mutation) can bypass the dependence of Gp25. However, attempts to isolate spontaneous revertants as plaque formersfrom Muc^{ts} *25::Cam* was not successful (data not shown). It is possible that in the context of the phage genome, , the *sel* mutation may cause premature activation of Lys before holin triggering, thus leading to a reduction in the yield of progeny phage particles. Mutational analysis of both Gp25 and Lys under the conditions

that lysis can be constantly monitored may reveal more details about the functional domain of Gp25 and the significance of Lys-Gp25 interaction.

REFERENCES

1. **Carlton RM.** 1999. Phage therapy: past history and future prospects. Arch Immunol Ther Exp (Warsz) **47:**267-274.

2. **Wittebole X, De Roock S, Opal SM.** 2014. A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. Virulence **5:**226-235.

3. **Abedon ST, Thomas-Abedon C, Thomas A, Mazure H.** 2011. Bacteriophage prehistory: Is or is not Hankin, 1896, a phage reference? Bacteriophage **1:**174-178.

4. **Clokie MRJ.** 2009. Practical methods for determining phage growth parameters. p. 175-202. *In* Kropinski AM (ed.), Bacteriophages: methods and protocol. Volume 1: Isolation, characterization, and interactions. Humana Press, New York City, NY.

5. **Jardine PJ.** 2006. DNA packageing in double-stranded DNA phages, p. 49-65. *In* Calendar R (ed.), The Bacteriophages, 2nd edition ed. Oxford University Press, New York City, NY.

6. **Tran TA.** 2007. Bacteriophage T4 lysis and lysis inhibition: molecular basis of an ancient story.Ph.D. Dissertation. Texas A&M University, College Station, TX.

7. **Berrier C, Bonhivers M, Letellier L, Ghazi A.** 2000. High-conductance channel induced by the interaction of phage lambda with its receptor maltoporin. FEBS Lett **476:**129-133.

8. **Roa M, Scandella D.** 1976. Multiple steps during the interaction between coliphage lambda and its receptor protein in vitro. Virology **72:**182-194.

9. **Scandella D, Arber W.** 1974. An *Escherichia coli* mutant which inhibits the injection of phage lambda DNA. Virology **58:**504-513.

10. **Scandella D, Arber W.** 1976. Phage lambda DNA injection into *Escherichia coli* pel- mutants is restored by mutations in phage genes V or H. Virology **69:**206-215.

11. **Szczepankowska A.** 2012. Role of CRISPR/cas system in the development of bacteriophage resistance. Adv Virus Res **82:**289-338.

12. **Little JW.** 2006. General aspects of lysogeny, p. 66-73. *In* Calendar R (ed.), The bacteriophages, 2nd edition ed. Oxford University Press, New York.

13. **Zeng L, Skinner SO, Zong C, Sippy J, Feiss M, Golding I.** 2010. Decision making at a subcellular level determines the outcome of bacteriophage infection. Cell **141:**682-691.

14. **Zeng L, Golding I.** 2011. Following cell-fate in *E. coli* after infection by phage lambda. J Vis Exp **56**:e3363.

15. **Jacob F, Fuerst CR.** 1958. The mechanism of lysis by phage studied with defective lysogenic bacteria. J Gen Microbiol **18:**518-526.

16. **Silhavy TJ, Kahne D, Walker S.** 2010. The bacterial cell envelope. Cold Spring Harb Perspect Biol **2:**a000414.

17. **Young R.** 1992. Bacteriophage lysis: mechanism and regulation. Microbiol Rev **56:**430-481.

18. **Savva CG, Dewey JS, Moussa SH, To KH, Holzenburg A, Young R.** 2014. Stable micron-scale holes are a general feature of canonical holins. Mol Microbiol **91:**57-65.

19. **Young R.** 2013. Phage lysis: do we have the hole story yet? Curr Opin Microbiol **16:**790-797.

20. **Park T, Struck DK, Dankenbring CA, Young R.** 2007. The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. J Bacteriol **189:**9135-9139.

21. **Pang T, Savva CG, Fleming KG, Struck DK, Young R.** 2009. Structure of the lethal phage pinhole. Proc Natl Acad Sci U S A **106:**18966-18971.

22. **Park T, Struck DK, Deaton JF, Young R.** 2006. Topological dynamics of holins in programmed bacterial lysis. Proc Natl Acad Sci U S A **103:**19713-19718.

23. **To KH, Young R.** 2014. Probing the structure of the S105 hole. J Bacteriol **196:**3683-3689.

24. **To KH, Dewey J, Weaver J, Park T, Young R.** 2013. Functional analysis of a class I holin, P2 Y. J Bacteriol **195:**1346-1355.

25. **Wang IN, Smith DL, Young R.** 2000. Holins: the protein clocks of bacteriophage infections. Annu Rev Microbiol **54:**799-825.

26. **Johnson-Boaz R, Chang CY, Young R.** 1994. A dominant mutation in the bacteriophage lambda S gene causes premature lysis and an absolute defective plating phenotype. Mol Microbiol **13:**495-504.

27. **Grundling A, Manson MD, Young R.** 2001. Holins kill without warning. Proc Natl Acad Sci U S A **98:**9348-9352.

28. **Blasi U, Nam K, Hartz D, Gold L, Young R.** 1989. Dual translational initiation sites control function of the lambda S gene. EMBO J **8:**3501-3510.

29. **Pang T, Park T, Young R.** 2010. Mapping the pinhole formation pathway of S^{21}. Mol Microbiol **78:**710-719.

30. **Xu M.** 2004. Bacteriophage P1: a new paradigm for control of phage lysis. Ph.D. Dissertation. Texas A&M University, College Station, TX.

31. **Miller ES, Kutter E, Mosig G, Arisaka F, Kunisawa T, Ruger W.** 2003. Bacteriophage T4 genome. Microbiol Mol Biol Rev **67:**86-156, table of contents.

32. **Young R.** 2014. Phage lysis: three steps, three choices, one outcome. J Microbiol **52:**243-258.

33. **Blasi U, Chang CY, Zagotta MT, Nam KB, Young R.** 1990. The lethal lambda S gene encodes its own inhibitor. EMBO J **9:**981-989.

34. **Ramanculov E, Young R.** 2001. An ancient player unmasked: T4 rI encodes a t-specific antiholin. Mol Microbiol **41:**575-583.

35. **Tran TA, Struck DK, Young R.** 2007. The T4 RI antiholin has an N-terminal signal anchor release domain that targets it for degradation by DegP. J Bacteriol **189:**7618-7625.

36. **Matthews BW.** 1996. Structural and genetic analysis of the folding and function of T4 lysozyme. FASEB J **10:**35-41.

37. **Issinger OG, Hausmann R.** 1973. Synthesis of bacteriophage-coded gene products during infection of *Escherichia coli* with amber mutants of T3 and T7 defective in gene 1. J Virol **11:**465-472.

38. **Kuty GF.** 2011. SAR endolysin regulation in dsdna phage lysis of gram-negative hosts. Ph.D. Dissertation. Texas A&M University, College Station, TX.

39. **Pang T.** 2010. Exploring the pinhole: biochemical and genetic studies on the prototype pinholin, S^{21}. Ph.D. dissertation. Texas A&M University, College Station, TX.

40. **Xu M, Struck DK, Deaton J, Wang IN, Young R.** 2004. A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. Proc Natl Acad Sci U S A **101:**6415-6420.

41. **Sun Q, Kuty GF, Arockiasamy A, Xu M, Young R, Sacchettini JC.** 2009. Regulation of a muralytic enzyme by dynamic membrane topology. Nat Struct Mol Biol **16:**1192-1194.

42. **Kuty GF, Xu M, Struck DK, Summer EJ, Young R.** 2010. Regulation of a phage endolysin by disulfide caging. J Bacteriol **192:**5682-5687.

43. **Xu M, Arulandu A, Struck DK, Swanson S, Sacchettini JC, Young R.** 2005. Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. Science **307:**113-117.

44. **Young R, Way J, Way S, Yin J, Syvanen M.** 1979. Transposition mutagenesis of bacteriophage lambda: a new gene affecting cell lysis. J Mol Biol **132:**307-322.

45. **Hanych B, Kedzierska S, Walderich B, Uznanski B, Taylor A.** 1993. Expression of the Rz gene and the overlapping Rz1 reading frame present at the right end of the bacteriophage lambda genome. Gene **129:**1-8.

46. **Zhang N, Young R.** 1999. Complementation and characterization of the nested Rz and Rz1 reading frames in the genome of bacteriophage lambda. Mol Gen Genet **262:**659-667.

47. **Summer EJ, Berry J, Tran TA, Niu L, Struck DK, Young R.** 2007. Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts. J Mol Biol **373:**1098-1112.

48. **Markov D, Christie GE, Sauer B, Calendar R, Park T, Young R, Severinov K.** 2004. P2 growth restriction on an rpoC mutant is suppressed by alleles of the Rz1 homolog lysC. J Bacteriol **186:**4628-4637.

49. **Berry J, Summer EJ, Struck DK, Young R.** 2008. The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. Mol Microbiol **70:**341-351.

50. **Berry J, Savva C, Holzenburg A, Young R.** 2010. The lambda spanin components Rz and Rz1 undergo tertiary and quaternary rearrangements upon complex formation. Protein Sci **19:**1967-1977.

51. **Rajaure M, Berry J, Kongari R, Cahill J, Young R.** 2015. Membrane fusion during phage lysis. Proc Natl Acad Sci U S A **112:**5497-5502.

52. **Doermann AH.** 1948. Lysis and lysis inhibition with *Escherichia coli* bacteriophage. J Bacteriol **55:**257-276.

53. **Moussa SH.** 2012. Deciphering lysis and its regulation in bacteriophage T4. Ph.D. dissertation. Texas A&M University, College Station, TX.

54. **Josslin R.** 1970. The lysis mechanism of phage T4: mutants affecting lysis. Virology **40:**719-726.

55. **Kuznetsov VB.** 2011. Structural studies of phage lysis proteins and their targets. Ph.D. dissertation, Texas A&M University, College Station.

56. **Moussa SH, Lawler JL, Young R.** 2014. Genetic dissection of T4 lysis. J Bacteriol **196:**2201-2209.

57. **Moussa SH, Kuznetsov V, Tran TA, Sacchettini JC, Young R.** 2012. Protein determinants of phage T4 lysis inhibition. Protein Sci **21:**571-582.

58. **Pang T, Fleming TC, Pogliano K, Young R.** 2013. Visualization of pinholin lesions in vivo. Proc Natl Acad Sci U S A **110:**E2054-2063.

59. **Ramanculov E, Young R.** 2001. Functional analysis of the phage T4 holin in a lambda context. Mol Genet Genomics **265:**345-353.

60. **Ramanculov E, Young R.** 2001. Genetic analysis of the T4 holin: timing and topology. Gene **265:**25-36.

61. **Cornett JB.** 1974. Spackle and immunity functions of bacteriophage T4. J Virol **13:**312-321.

62. **Emrich J.** 1968. Lysis of T4-infected bacteria in the absence of lysozyme. Virology **35:**158-165.

63. **Lu MJ, Stierhof YD, Henning U.** 1993. Location and unusual membrane topology of the immunity protein of the *Escherichia coli* phage T4. J Virol **67:**4905-4913.

64. **Paddison P, Abedon ST, Dressman HK, Gailbreath K, Tracy J, Mosser E, Neitzel J, Guttman B, Kutter E.** 1998. The roles of the bacteriophage T4 r genes in lysis inhibition and fine-structure genetics: a new perspective. Genetics **148:**1539-1550.

65. **Hershey AD.** 1946. Mutation of bacteriophage with respect to type of plaque. Genetics **31:**620-640.

66. **Hershey AD.** 1946. Spontaneous mutations in bacterial viruses, Cold Spring Harb Symp Quant Biol, vol. 11: 67-77. Cold Spring Harbor Laboratory Press.

67. **Burch LH, Zhang L, Chao FG, Xu H, Drake JW.** 2011. The bacteriophage T4 rapid-lysis genes and their mutational proclivities. J Bacteriol **193:**3537-3545.

68. **Kai T, Ueno H, Otsuka Y, Morimoto W, Yonesaki T.** 1999. Gene 61.3 of bacteriophage T4 is the spackle gene. Virology **260:**254-259.

69. **Dressman HK, Drake JW.** 1999. Lysis and lysis inhibition in bacteriophage T4: rV mutations reside in the holin t gene. J Bacteriol **181:**4391-4396.

70. **Tran TA, Struck DK, Young R.** 2005. Periplasmic domains define holin-antiholin interactions in t4 lysis inhibition. J Bacteriol **187:**6631-6640.

71. **Luke K, Radek A, Liu X, Campbell J, Uzan M, Haselkorn R, Kogan Y.** 2002. Microarray analysis of gene expression during bacteriophage T4 infection. Virology **299:**182-191.

72. **Rutberg B, Rutberg L.** 1965. Role of superinfecting phage in lysis inhibition with phage T4 in *Escherichia coli*. J Bacteriol **90:**891-894.

73. **Anderson CW, Eigner J.** 1971. Breakdown and exclusion of superinfecting T-even bacteriophage in *Escherichia coli*. J Virol **8:**869-886.

74. **Stock JB, Rauch B, Roseman S.** 1977. Periplasmic space in Salmonella typhimurium and *Escherichia coli*. J Biol Chem **252:**7850-7861.

75. **Neidhardt FC.** 1996. *Escherichia coli* and Salmonella: cellular and molecular biology. ASM Press **1:**15.

76. **Harshey RM.** 2012. The Mu story: how a maverick phage moved the field forward. Mob DNA **3:**21.

77. **Symonds T, Van de Putte, Howe.** 1987. Phage Mu. p. 63-74. *In* Symonds N (ed.) Cold Spring Harbor Laboratory, New York, New York.

78. **Mizuuchi K.** 1983. In vitro transposition of bacteriophage Mu: a biochemical approach to a novel replication reaction. Cell **35:**785-794.

79. **Montano SP, Pigli YZ, Rice PA.** 2012. The Mu transpososome structure sheds light on DDE recombinase evolution. Nature **491:**413-417.

80. **Craigie R, Mizuuchi K.** 1985. Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate. Cell **41:**867-876.

81. **Mizuuchi K, Craigie R.** 1986. Mechanism of bacteriophage Mu transposition. Annu Rev Genet **20:**385-429.

82. **Grundy FJ, Howe MM.** 1985. Morphogenetic structures present in lysates of amber mutants of bacteriophage Mu. Virology **143:**485-504.

83. **Morgan GJ, Hatfull GF, Casjens S, Hendrix RW.** 2002. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in Haemophilus, Neisseria and Deinococcus. J Mol Biol **317:**337-359.

84. **Faelen M, Toussaint A.** 1973. Isolation of conditional defective mutants of temperate phage Mu-1 and deletion mapping of the Mu-1 prophage. Virology **54:**117-124.

85. **To KH.** 2013. Biochemical and genetic characterization of bacteriophage holins. Ph.D. dissertation Texas A&M University, College Station.

86. **Hershey AD, Chase M.** 1951. Genetic recombination and heterozygosis in bacteriophage. Cold Spring Harb Symp Quant Biol **16:**471-479.

87. **Benzer S.** 1955. Fine structure of a genetic region in bacteriophage. Proc Natl Acad Sci U S A **41:**344-354.

88. **Crick FH, Barnett L, Brenner S, Watts-Tobin RJ.** 1961. General nature of the genetic code for proteins. Nature **192:**1227-1232.

89. **Benzer S.** 1957. The elementary units of heredity. p. 340-367. *In* William d. Mcelroy bg (ed.), The Chemical Basis of Heredity. Johns Hopkins University Press, Baltimore.

90. **Hershey AD, Rotman R.** 1948. Linkage among genes controlling inhibition of lysis in a bacterial virus. Proc Natl Acad Sci U S A **34:**89-96.

91. **Krylov VN, Yankovsky NK.** 1975. Mutations in the new gene stIII of bacteriophage T4B suppressing the lysis defect of gene stII and a gene e mutant. J Virol **15:**22-26.

92. **Bode W.** 1967. Lysis inhibition in *Escherichia coli* infected with bacteriophage T4. J Virol **1:**948-955.

93. **White R, Tran TA, Dankenbring CA, Deaton J, Young R.** 2010. The N-terminal transmembrane domain of lambda S is required for holin but not antiholin function. J Bacteriol **192:**725-733.

94. **Barenboim M, Chang CY, dib Hajj F, Young R.** 1999. Characterization of the dual start motif of a class II holin gene. Mol Microbiol **32:**715-727.

95. **Golec P, Karczewska-Golec J, Voigt B, Albrecht D, Schweder T, Hecker M, Wegrzyn G, Los M.** 2013. Proteomic profiles and kinetics of development of bacteriophage T4 and its rI and rIII mutants in slowly growing *Escherichia coli*. J Gen Virol **94:**896-905.

96. **Miller AK, Brown EE, Mercado BT, Herman JK.** 2016. A DNA-binding protein defines the precise region of chromosome capture during Bacillus sporulation. Mol Microbiol **99:**111-122.

97. **Lutz R, Bujard H.** 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res **25:**1203-1210.

98. **Bendezu FO, Hale CA, Bernhardt TG, de Boer PA.** 2009. RodZ (YfgA) is required for proper assembly of the MreB actin cytoskeleton and cell shape in E. coli. EMBO J **28:**193-204.

99. **Battesti A, Bouveret E.** 2012. The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*. Methods **58:**325-334.

100. **Hansson MD, Rzeznicka K, Rosenback M, Hansson M, Sirijovski N.** 2008. PCR-mediated deletion of plasmid DNA. Anal Biochem **375:**373-375.

101. **Karimova G, Pidoux J, Ullmann A, Ladant D.** 1998. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc Natl Acad Sci U S A **95:**5752-5756.

102. **Mathews CK, Mosig G, Berget P.** 1983. Bacteriophage T4. p. 157-163 *In* Mathews C (ed). American Society of Microbiology, Wachington, D.C.

103. **Raudonikiene A, Nivinskas R.** 1992. Gene rIII is the nearest downstream neighbour of bacteriophage T4 gene 31. Gene **114:**85-90.

104. **Pang T, Park T, Young R.** 2010. Mutational analysis of the S^{21} pinholin. Mol Microbiol **76:**68-77.

105. **Raab R, Neal G, Garrett J, Grimaila R, Fusselman R, Young R.** 1986. Mutational analysis of bacteriophage lambda lysis gene S. J Bacteriol **167:**1035-1042.

106. **Obringer JW.** 1988. The functions of the phage T4 immunity and spackle genes in genetic exclusion. Genet Res **52:**81-90.

107. **Lu MJ, Henning U.** 1994. Superinfection exclusion by T-even-type coliphages. Trends Microbiol **2:**137-139.

108. **Drozdetskiy A, Cole C, Procter J, Barton GJ.** 2015. JPred4: a protein secondary structure prediction server. Nucleic Acids Res **43:**W389-394.

109. **Puck TT, Lee HH.** 1955. Mechanism of cell wall penetration by viruses. II. Demonstration of cyclic permeability change accompanying virus infection of *Escherichia coli* B cells. J Exp Med **101:**151-175.

110. **Blasi U, Nam K, Lubitz W, Young R.** 1990. Translational efficiency of phi X174 lysis gene E is unaffected by upstream translation of the overlapping gene D reading frame. J Bacteriol **172:**5617-5623.

111. **Zuker M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res **31:**3406-3415.

112. **Nanes BA.** 2015. Slide Set: Reproducible image analysis and batch processing with ImageJ. Biotechniques **59:**269-278.

113. **Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C.** 2012. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins **80:**374-381.

114. **Morein S, Henricson D, Rilfors L.** 1994. Separation of inner and outer membrane vesicles from *Escherichia coli* in self-generating Percoll gradients. Anal Biochem **216:**47-51.

115. **Peter Prevelige GDF.** 1989. Chou-Fasman prediction of the secondary structure of proteins, p. 391-416 *In* Fasman GD (ed.), Prediction of protein structure and the principles of protein conformation. Plenum press, New York.

116. **Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y.** 2015. The I-TASSER Suite: protein structure and function prediction. Nat Methods **12:**7-8.

117. **Chang CY, Nam K, Young R.** 1995. S gene expression and the timing of lysis by bacteriophage lambda. J Bacteriol **177:**3283-3294.

118. **Grundling A, Smith DL, Blasi U, Young R.** 2000. Dimerization between the holin and holin inhibitor of phage lambda. J Bacteriol **182:**6075-6081.

119. **Blasi U, Young R.** 1996. Two beginnings for a single purpose: the dual-start holins in the regulation of phage lysis. Mol Microbiol **21:**675-682.

120. **Abedon ST.** 1992. Lysis of lysis-inhibited bacteriophage T4-infected cells. J Bacteriol **174:**8073-8080.

121. **Weaver LH, Matthews BW.** 1987. Structure of bacteriophage T4 lysozyme refined at 1.7 A resolution. J Mol Biol **193:**189-199.

122. **Guzman LM, Belin D, Carson MJ, Beckwith J.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol **177:**4121-4130.

123. **Wang IN, Deaton J, Young R.** 2003. Sizing the holin lesion with an endolysin-beta-galactosidase fusion. J Bacteriol **185:**779-787.

124. **Bolhuis A, Bogsch EG, Robinson C.** 2000. Subunit interactions in the twin-arginine translocase complex of *Escherichia coli*. FEBS Lett **472:**88-92.

125. **Englesberg E, Anderson RL, Weinberg R, Lee N, Hoffee P, Huttenhauer G, Boyer H.** 1962. L-Arabinose-sensitive, L-ribulose 5-phosphate 4-epimerase-deficient mutants of *Escherichia coli*. J Bacteriol **84:**137-146.

126. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640-6645.

127. **Roces C, Campelo AB, Escobedo S, Wegmann U, Garcia P, Rodriguez A, Martinez B.** 2016. Reduced binding of the endolysin lystp712 to lactococcus lactis deltaftsh contributes to phage resistance. Front Microbiol **7:**138.

128. **Levin BR, Bull JJ.** 2004. Population and evolutionary dynamics of phage therapy. Nat Rev Microbiol **2:**166-173.

129. **Wang J, Hofnung M, Charbit A.** 2000. The C-terminal portion of the tail fiber protein of bacteriophage lambda is responsible for binding to LamB, its receptor at the surface of *Escherichia coli* K-12. J Bacteriol **182:**508-512.

130. **Hendrix RW, Duda RL.** 1992. Bacteriophage lambda PaPa: not the mother of all lambda phages. Science **258:**1145-1148.

131. **Williams N, Fox DK, Shea C, Roseman S.** 1986. Pel, the protein that permits lambda DNA penetration of *Escherichia coli*, is encoded by a gene in ptsM and is required for mannose utilization by the phosphotransferase system. Proc Natl Acad Sci U S A **83:**8934-8938.

132. **Katsura I, Hendrix RW.** 1984. Length determination in bacteriophage lambda tails. Cell **39:**691-698.

133. **Roessner CA, Ihler GM.** 1984. Proteinase sensitivity of bacteriophage lambda tail proteins gpJ and pH in complexes with the lambda receptor. J Bacteriol **157:**165-170.

134. **Arai K, McMacken R, Yasuda S, Kornberg A.** 1981. Purification and properties of *Escherichia coli* protein i, a prepriming protein in phi X174 DNA replication. J Biol Chem **256:**5281-5286.

135. **Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H.** 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol **2:**2006 0008.

136. **Shaw JP, Aviles Medina CA, Chen Y, Luna AJ, Hernandez AC, Kuty Everett GF.** 2015. Complete Genome of *Citrobacter freundii* Siphophage Stevie. Genome Announc **3**.

137. **Drake JW.** 1966. Ultraviolet mutagenesis in bacteriophage T4. II. Photoreversal of mutational lesions. J Bacteriol **92:**144-147.

138. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res **18:**821-829.

139. **Besemer J, Borodovsky M.** 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res **33:**W451-454.

140. **Esquinas-Rychen M, Erni B.** 2001. Facilitation of bacteriophage lambda DNA injection by inner membrane proteins of the bacterial phosphoenol-pyruvate:

carbohydrate phosphotransferase system (PTS). J Mol Microbiol Biotechnol **3:**361-370.

141. **Erni B, Zanolari B, Kocher HP.** 1987. The mannose permease of *Escherichia coli* consists of three different proteins. Amino acid sequence and function in sugar transport, sugar phosphorylation, and penetration of phage lambda DNA. J Biol Chem **262:**5238-5247.

142. **Elliott J, Arber W.** 1978. E. coli K-12 pel mutants, which block phage lambda DNA injection, coincide with ptsM, which determines a component of a sugar transport system. Mol Gen Genet **161:**1-8.

143. **Edgar R, Rokney A, Feeney M, Semsey S, Kessel M, Goldberg MB, Adhya S, Oppenheim AB.** 2008. Bacteriophage infection is targeted to cellular poles. Mol Microbiol **68:**1107-1116.