BIOCHEMICAL AND GENETIC CHARACTERIZATION OF BACTERIOPHAGE HOLINS

A Dissertation

by

KAM HO TO

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,	Michael J. Benedik
	Siegfried Musser
	Matthew S. Sachs
Head of Department,	Tom McKnight

December 2013

Major Subject: Microbiology

ABSTRACT

Bacteriophages infect and kill bacterial cells. During the infection cycle, a phage attaches to the host cell surface, then ejects its DNA into the cytoplasm, where its progenies are subsequently assembled. The final step of the infection cycle is host cell lysis, which allows the progeny virions to escape into the environment. However, the timing of lysis, and thus the length of the infection cycle, is independent of endolysin biosynthesis and rather depends on the function of a second class of lysis proteins, the holins. Holins are small integral membrane proteins that accumulate harmlessly in the membrane during the infection cycle, until they suddenly form lethal lesions in the membrane at an allele-specific time. This membrane damage allows the endolysin to attack the cell wall.

This dissertation focuses on several aspects of the structural and functional aspect of holins. First, *Y* is the putative holin gene of the paradigm coliphage P2. Although Y is not related to the S holin of phage lambda according to its primary structure, its characterization might prove useful in discerning the essential traits for holin function. In this instance, physiological and genetic approaches are utilized to show that Y exhibits the essential holin functional criteria, namely, allele-specific delayed-onset lethality and sensitivity to the energization of the membrane. These results suggest that class I holins share a set of unique features that are needed for their remarkable ability to program the end of the phage infection cycle with precise timing. Nevertheless, I report studies involving phenotypic analysis of a systematic library of clustered site-directed

mutants of S105, and then conclude with experiments designed to probe the structure of the mature "S-hole" in the membrane of the cell using chemical probes. Furthermore, I address whether the Y holin and the S^{21} pinholin of phage 21 effect membrane depolarization with the same all-or-nothing fashion as S while using the same tethered-cell assay previously employed for studying S. Finally, the holin and antiholin in Mu, one of the few paradigm coliphage, were identified and characterized. The introductory chapter is intended to serve as an update to the last major review on holin function in 2000.

DEDICATION

This dissertation is dedicated to my parents, Alan and Angel, for their endless love and support. They never stop believing in me. I couldn't have done it without you.

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Ryland F. Young, and my committee members, Dr. Michael J. Benedik, Dr. Siegfried Musser, and Dr. Matthew S. Sachs for their insight, guidance, and support throughout the course of this research at Texas A&M University. I am also thankful for past and present Young lab members for their support, discussions, and critiques that only make me a better scientist and presenter.

I would like to thank my friends, colleagues, collaborators, faculty and staff in both the department of Biology and Biochemistry for making my time at Texas A&M University a great experience. I also want to extend my gratitude to all the undergraduate students and laboratory rotations graduate students, Michael Dachowski, Matthew Maddox, Yuan Yang, Chelsea Harris, and Karthik Chamakura who worked on numerous research projects with me throughout my years at Texas A&M University. A special thanks to Daisy Wilbert for your unwavering support to all the Young lab members.

Finally, thanks to my mother and father for their encouragement, patience and love. I am indebted to my family who has been extremely supportive of my education.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION TO BACTERIOPHAGE HOLINS

The lambda lysis cassette

The major system for studying lysis has been the bacteriophage lambda, primarily because the lambda lytic cycle can be initiated synchronously throughout a culture by thermal induction of a lambda lysogen (1). This allows the timing and extent of lysis to be monitored easily by following bulk features of the culture, including culture mass, usually as A_{550}, viability, production of lysis proteins and virion proteins, progeny virus particles and other parameters. The lambda lysis proteins are encoded in a "lysis cassette", with four adjacent genes: the antiholin/holin (*S*), the endolysin (*R*), and the spanins (*Rz/Rz1*), clustered downstream of the lambda late promoter, $P_{R'}$ (Fig. 1). My focus is on the holin, but I first describe briefly the endolysin and spanin genes.

The lambda R endolysin

In lambda, *R* encodes the endolysin. The term endolysin refers to the enzyme with muralytic activity that the phage uses for degrading the cell wall during lysis. In lambda, the R protein accumulates in the cytoplasm from the beginning of the late gene or morphogenesis period, that begins approximately 8 minutes after infection. R has transglycosylase activity, which means it forms a cyclic product, rather than simply hydrolyzing the glycosidic bond of the peptidoglycan (Fig. 2). It cleaves the β-1,4 glycosidic bond between MurNAc and GlcNAc moieties of the peptidoglycan, using

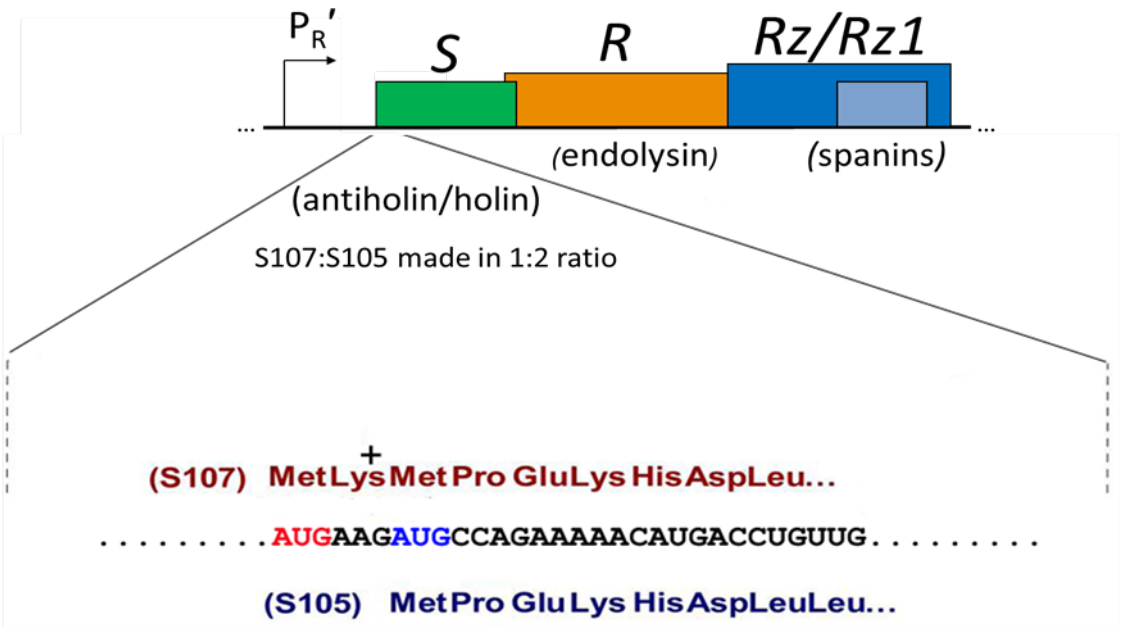

Fig. 1. The lysis cassettes of Lambda. Downstream of Pr' is the lambda "lysis cassette", which contains genes encoding the holin (*S*) shown (green), endolysin (*R*) (orange), and the spanins (*Rz*) (dark blue) and (*Rz1*) (light blue). *S* encodes two proteins: S107 (red), which is the antiholin (2), and S105 (blue), which is the holin. S107 and S105 are made in a 1:2 ratio. S107, which has two additional residues at the N-terminus, is 107 amino acids long, and its translation begins with Met1. S105 is 105 amino acids long, and its translation begins on Met3. This N-terminus differentiates between the antiholin and the holin and is called the "dual start motif".

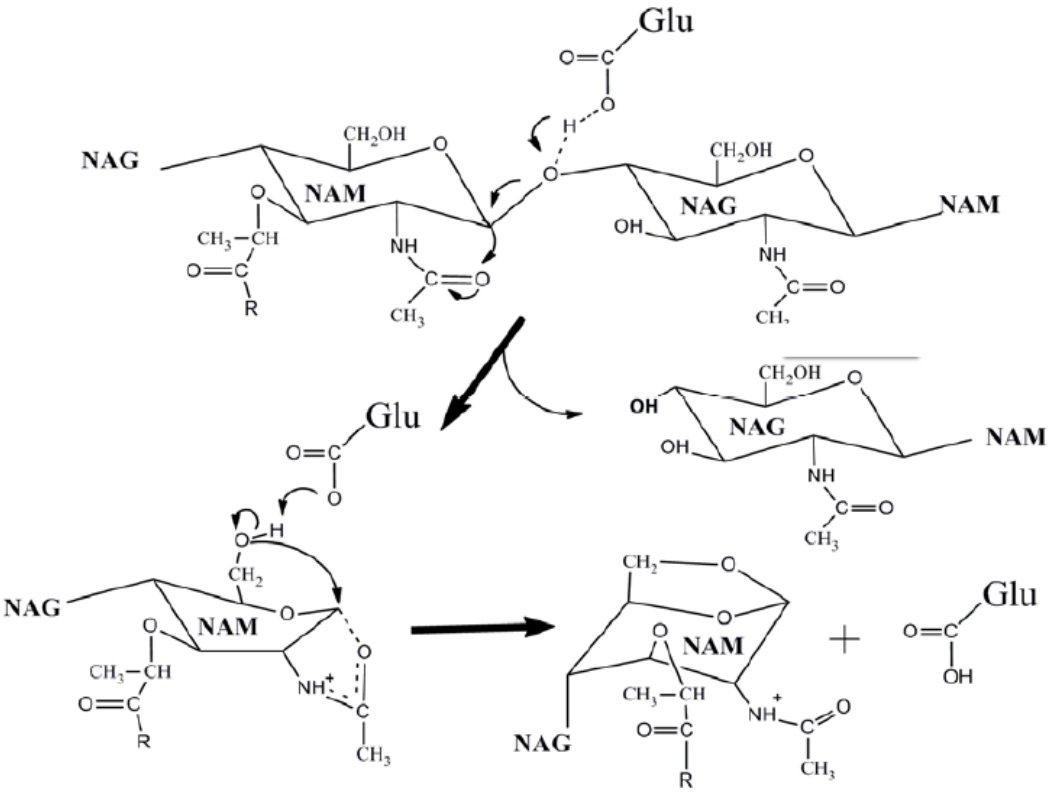

Fig. 2. Lytic transglycosylase mechanism. Transglycosylases cleave the β-1,4 glycosidic bond between MurNAc and GlcNAc moieties of PG. A catalytic glutamate acts as an acid to protonate the bond to release the GlcNAc moiety (3). The same Glu then acts as a base to remove a proton from the C6 hydroxyl of the remaining MurNAc residue forming a resonance stabilized intermediate. The intermediate collapses to form a 1, 6-anhydro-N-acetyl-muramoyl product. Adopted from Scheurwater et al. (3).

a catalytic glutamate as an acid to protonate the glycosidic linkage, releasing the GlcNAc moiety (4). The same Glu then acts as a base to remove a proton from the C6 hydroxyl of the remaining MurNAc residue forming a resonance stabilized intermediate. The intermediate collapses to form a 1, 6-anhydro-N-acetyl-muramoyl product.

Not all phages use a transglycosylase for the endolysin. In fact, four different muralytic enzyme activities are found in endolysins. Besides the transglycosylase, there are also the following enzymatic classes: glycosidase or "true lysozyme" (e.g., T4 gp*e* lysozyme and P22 gp*19* lysozyme), amidase (e.g., T7 gp*3.5* lysozyme and φ11 murein hydrolase), and endopeptidase (e.g., φ11 murein hydrolase) (5). There is no dependence on the muralytic activity for holin function. The gp*13* holin of P22 has 90% identity compared to lambda S, but has a glycosidase endolysin.

The Rz/Rz1 spanin genes

The *Rz* and *Rz1* genes encode a third class of lysis proteins, the spanins, responsible for disruption of the outer membrane (OM) and thus restricted to phages of Gram-negative hosts (1). The *Rz* and *Rz1* genes have an unusual architecture, in that the gene sequence for the latter is entirely embedded within that of the former, in the +1 reading frame (Fig. 3). Rz is an type II integral membrane protein with a C-terminal periplasmic domain, whereas Rz1 is a small outer membrane lipoprotein. During late gene expression, the Rz and Rz1 proteins accumulate in the envelope and form complexes by C-terminal interactions. The complex spans the entire periplasm, which is why the complex has been designated as the spanin complex. The function of the spanin complex is the disruption of the OM (6). The evidence for this is that if either spanin

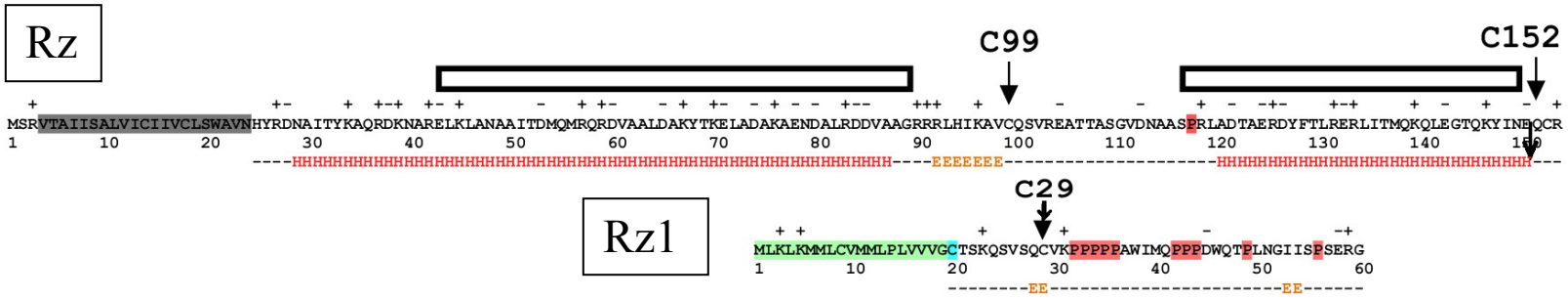

Fig. 3. The primary structures of Rz and Rz1. Aligned in the actual register of the embedded *Rz1* gene (i.e., the start codon of Rz1 begins at the second base of codon Asp73 of Rz). The Rz (Type-II signal anchor) and Rz1 (lipoprotein) signal sequences are highlighted in gray and light green, respectively. Alpha-helical (red capital H) and extended (orange capital E) predicted by JPRED (7) (http://www.compbio.dundee.ac.uk/www-jpred/) are shown below each primary structure. Coiled-coil regions in Rz predicted by PairCoil2 (8) (http://groups.csail.mit.edu/cb/paircoil2/) are shown by open rectangles above the sequence. Proline residues in the mature periplasmic domains and the modified N-terminal Cys of mature Rz1 are highlighted in red and teal, respectively. Arrows indicate unmodified periplasmic cysteine residues of Rz (C99 and C152) and Rz1 (C29) in the intermolecular disulfide bonds studies (9). The *Rz* and *Rz1* genes of lambda were originally studied because the latter gene is entirely embedded within the former in the +1 reading frame. This is the only case where two different reading frames of the DNA sequence encodes two different proteins essential for the same function. Rz1 is an outer membrane lipoprotein of a mature length of only 40 amino acids. It forms a C-terminal complex with the Rz spanin, constituting a complex that spans the periplasmic space between the inner and outer membrane. Adopted from Berry et al. (9)

subunit is inactivated, the terminal phenotype of the infected cell is a spherical cell, in which the peptidoglycan has been destroyed but the OM is intact (Fig. 4). The mechanism by which the spanin complex functions is unknown, although it has been established that destruction of the murein layer is necessary, presumably because the spanin complexes are trapped within the peptidoglycan matrix. Spanin function is independent of holin function. This was shown by inducing peptidoglycan destruction with a secretory endolysin (6). In the presence of the lambda spanins, complete lysis was observed. However, in the absence of the spanin complex, the cells were converted into spherical forms bounded by the OM.

Equivalents to the lambda *Rz-Rz1* spanin genes pair have been found in most Gram-negative hosts (10). In addition, they are found in two other architectures, besides the *embedded* arrangement seen in *Rz/Rz1*(10). In some cases, the *Rz1*-equivalent gene extends beyond the end of the *Rz* gene (*overlapped*) and in others, the genes are entirely *separated*. In all cases, the *Rz*-equivalent is always the first gene, and there are never any intervening genes between the *Rz* and *Rz1* genes. To accommodate this diversity, the general terms *i-spanin* (for inner membrane spanin) and *o-spanin* (outer membrane spanin) have been adopated for the functional equivalents of *Rz* and *Rz1*, respectively, in all these diverse loci.

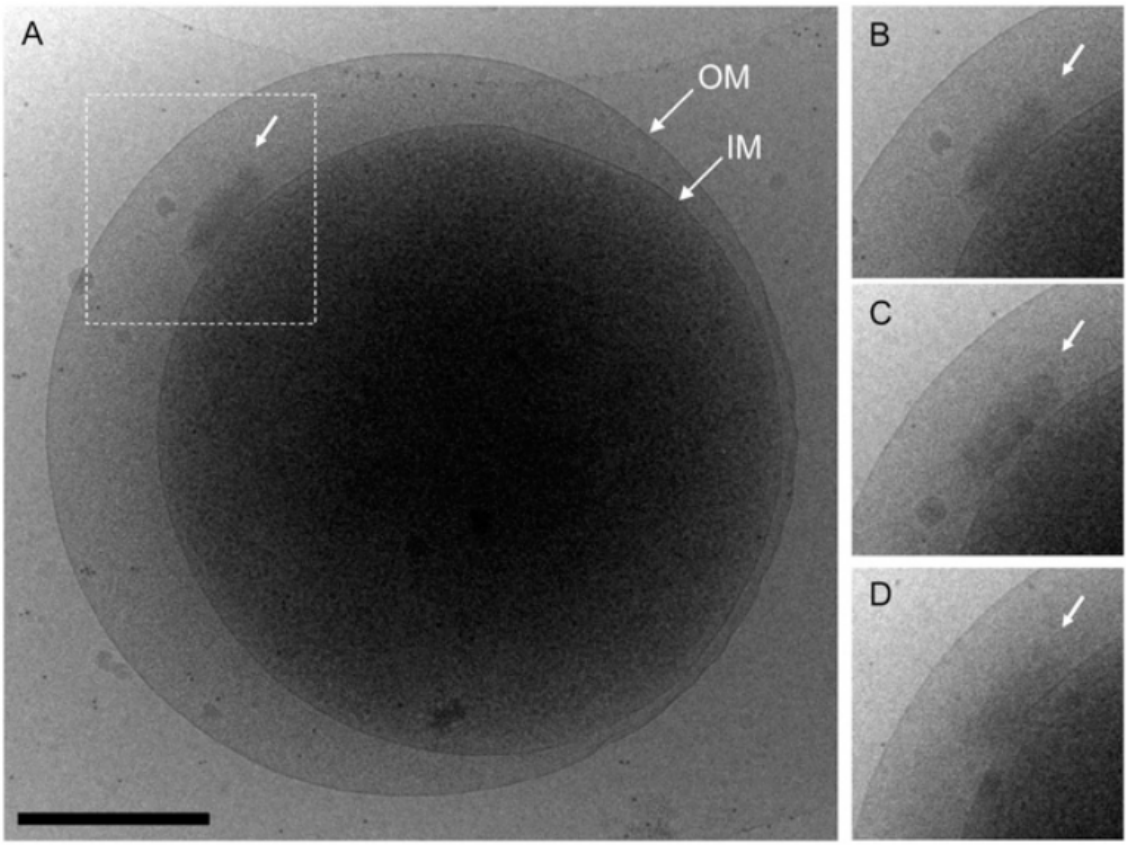

Fig. 4. Cytoplasmic leakage from a cell coexpressing S105 and R. (A and B) At a tilt of 0°, the image indicates a cloud of density escaping from the cytoplasm. White arrows point to the protein density escaping from the cytoplasm. The white dashed box indicates the area depicted in B. IM, inner membrane; OM, outer membrane. Tilting of the stage by 20° (C) and 30° (D) revealed the membrane lesion from which the leakage originated. (Scale bar: 500 nm.) Adopted from Dewey et al. (6).

The lambda holin gene S

The *S* gene was discovered in 1967 by mapping mutations that abolished lambda lysis (11). Over the last 46 years, the role of S has been studied in detail, summarized recently in two major reviews (5, 12) and in a book chapter (13). Here I will summarize what was known about holin structure, function and mechanism as of the time of these major reviews, and then devote the rest of this introduction to the progress made up to the beginning of my dissertation work. This material includes the understanding of the S holin function in lysis at the cellular level, as well as biochemical and topological information about the S protein in vivo and in vitro.

In general, the *S* gene was known to be required for lysis because the R endolysin had no secretory signal (1). Thus, the fully folded and active endolysin with muralytic activity was trapped within the cytosol during the late gene expression period. Independent of endolysin function, the *S* gene product was known to cause a sudden cessation in respiration and macromolecular synthesis, concomitant with a leakage of cytoplasmic ions, collapse of the membrane potential and cell death (5, 12). This is called "triggering". In the wild-type (wt) phage, this occurs at ~ 50 min after the beginning of the lytic cycle, but many *S* alleles with different triggering times, both earlier and later than the wt, were known. Allele-specific, temporally regulated lethal triggering was considered to be a general feature of holins (14).

An important feature of the *S* gene is that there are two translational start sites at positions Met1 and Met3, which serve as the beginning residues of two separate proteins, S107 and S105 (15) (Fig. 1). The two translational start sites result in the

production of two membrane proteins that differ in length by only two amino acids (aa) but have opposite biological effects. The region containing both translation start codons is referred to as a "dual-start motif". Numerous unrelated holin genes have been found with similar motifs (5). The shorter product, S105, is the holin that functions in permeabilization of the membrane, collapse of membrane potential, and lysis timing (5, 16, 17). In contrast, the S107 protein is called the antiholin because it heterodimerizes with S105 and inhibits its function; see below (18). As a result, the proportion of S107 is related to the delay in the timing of cell lysis (19). Both in vitro and in vivo experiments indicated that the ratio of S105 molecules to S107 expressed from the wt lambda *S* gene is around 2:1 (20). This ratio of expression is controlled by two mRNA stem-loops, one upstream of the start codon and the other approximately 30 nucleotides (nt) downstream, within the *S* gene (16).

The lambda S105 holin

Topology of the S protein

The structure and membrane topology of lambda S105 has been subjected to intensive studies in the last two decades (Fig. 5). Multiple approaches including circular dichroism with purified protein, protease accessibility assays, phoA-LacZα fusions, cysteine-scanning mutagenesis, and chemical modification with thiol-specific reagent were performed (5). Methods using traditional *phoA*, *lacZ*, and *bla* gene fusions to examine the membrane topology were not successful, presumably because, due to the small size of the S holin, loss of the C-terminal topology determinants caused misfolding or mislocalization. Nonetheless, a bifunctional *phoA-lacZα* reporter was successfully

constructed and acted as the connector loop between transmembrane domain (TMD) 2 and TMD3, which is located in the periplasm (21). The findings from these experiments, coupled with primary structure prediction, suggested that S105 holin had three TMD_S that span the inner membrane (IM) of the *E. coli* cell (22-24). The N-terminus is located in the periplasm and the C-terminus is located in the cytoplasm (24).

Both biochemical and genetic approaches were used to determine the topology of S105 in the energized inner membrane. A milestone experiment by Graschopf and Blăsi suggested that lambda S105 consists of three TMDs (23). The secretory signal sequence (SS) of the coat protein of M13 was fused to the N-terminus of the S105 sequence. As a result, the SS-S105 chimera maintained its original lytic function but became leader peptidase (Lep) dependent. These results suggest that the N-terminus of S105 is located in the periplasm for holin function. Cysteine-modification studies allowed the S105 topology to be probed in membranes using a large collection of *S* variants with single-cysteine substitutions (Fig. 5). Most of these mutants were stable, localized in the membrane, and maintained lytic function. These alleles were expressed in *E. coli* cells and both inverted membrane vesicles (IMVs) and spheroplasts were prepared. The various samples were treated with 4-acetamido-4′-((iodoacetyl)amino)stilbene-2,2′-disulfonic acid (IASD), a cysteine-specific modification and membrane impermeable reagent (24). Three hydrophobic regions in the S protein were found to be resistant to modification, corresponding to the three TMDs predicted by primary structure analysis and the gene-fusion experiments (Fig. 5). Moreover, analysis of Cys-labeling kinetics

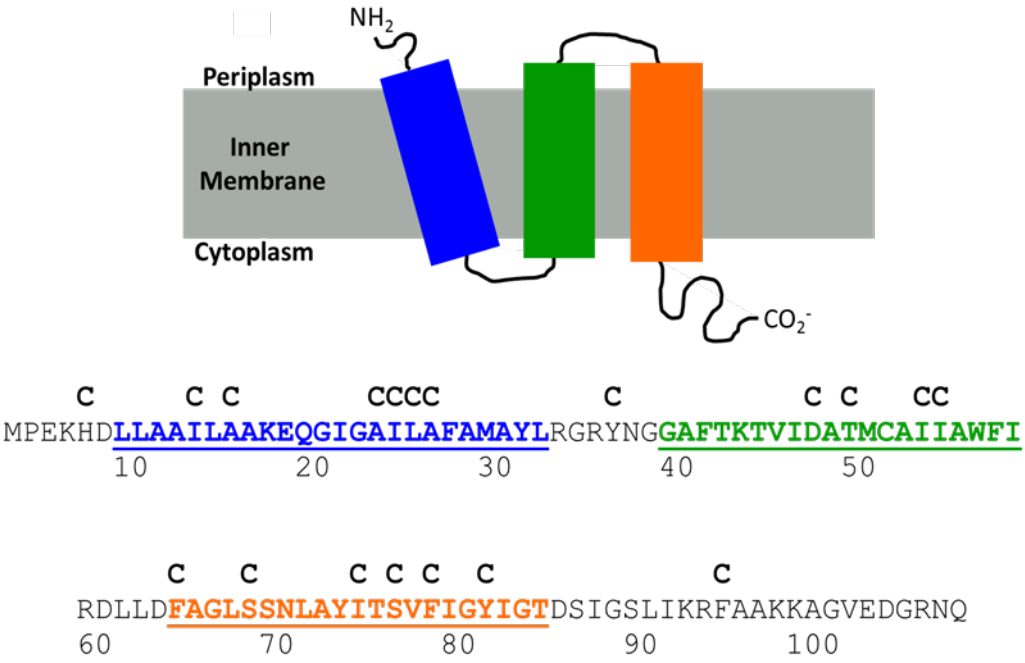

Fig. 5. Membrane topology of S105. The N-terminus is located in the periplasm, and the C-terminus is located in the cytoplasm. The three colors represent the three TMDs based on numerous experimental results mentioned in the text. The primary structure of S105 and the positions of single-cysteine substitution studies mentioned in the text are shown. (Note: the numbering system shown here are from S107)

indicated that the highly charged C-terminus was partially protected, possibly by being bound by Coulombic interactions to the inside surface of the energized inner membrane.

Genetic analysis of S

Extensive mutagenesis of *S* has been conducted, both by random and site-directed methods, generating a large library of missense alleles with single residue changes throughout the S protein (12, 25, 26) (Fig. 6). Most mutations that were isolated were obtained by simply inducing a plasmid clone of the *S* gene and selecting for survivors. In most membrane proteins, usually the TMDs are tolerant of substitutions, at least as long as the hydrophobicity of the TMD is not changed by the missense mutation. However, missense mutations in all three TMD helices of *S* were shown to have lysis-defective phenotypes, even though many were conservative changes that had no significant effect on the overall hydrophobicity of the TMD involved. In addition, these lysis-defective alleles exhibited unexpected dominance-recessiveness character, when co-expressed with the wt allele (26). Under the assumption that the S protein must undergo extensive oligomerization in order to form lethal membrane lesions, it was expected that some alleles would be dominant, presumably poisoning the oligomerization so that lethal holes could not form, whereas others would be recessive, presumably because the mutant proteins would not participate in the oligomers (25). By determining the lysis phenotype of double lysogens, where one carried the wt allele and the other a lysis-defective mutant *S* allele, each defective allele was subjected to dominance/recessiveness tests. Both dominant and recessive alleles were observed (15).

However, an unexpected phenotype was observed in several cases, in which lysis occurred earlier when the wt and the absolute-defective mutant allele were expressed together than with two wt alleles. The existence of these "early dominant" alleles suggested that lysis timing was a result of protein-protein interactions within the TMDs of the S protein (12). Thus, even if a homogeneous population bearing a certain missense change was unable to trigger, presumably because some oligomerization step was blocked, a mixed population carrying that change and also the wt residue at that position might trigger even earlier than the homogeneous wt population.

In addition, two non-plaque-forming alleles that were isolated and that mapped to a single residue, Ala52, in the middle of TMD2 were shown to have opposite lysis phenotypes (12). One allele, Ala52Val, was an absolute-lysis defective; the $S105_{A52V}$ protein was shown to accumulate in the membrane and was insensitive to exogenous triggering by energy poisons. In contrast, the Ala52Gly allele caused lysis at 19 min after induction of the prophage carrying this mutation. Thus, the non-plaque forming phenotype of Ala52Gly is due to triggering before the first virion, on average, is assembled (at ~ 20 min). These results demonstrate that the *S* gene has two essential functions: one is to permeabilize the membrane to allow the endolysin access to the peptidoglycan, and the other is to impose allele-specific timing on the triggering event.

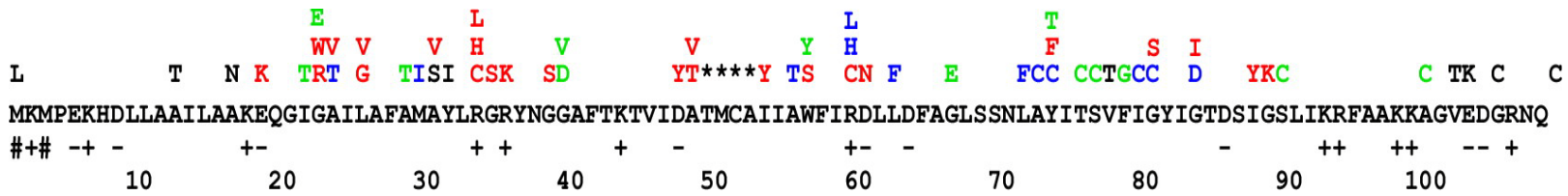

Fig. 6. Mutations in S. Missense changes relevant to the text are shown. Red represents non-lytic; Green represents early lysers; Black represents similar to wild-type lysis timing; blue represents delayed lysis timing. Translation starts for S107 (Met1) and S105 (Met3) are indicated by # signs. Four residues there were mutagenized to all possible amino acids are indicated by asterisks and the result are showed on page 99 (26).

Sensitivity to the proton motive force

The question of what constitutes the timing "clock" has been one of the primary issues about holin function. The only physiological clue was the fact that, for all phages dependent on holin-endolysin lysis, depolarizing the membrane caused premature triggering. Although why the loss of membrane energization should cause hole formation was not at all clear, this led to a general model that holin timing might simply reflect that holins caused leakage for protons or other ions (27, 28). Thus, as the holins increased in concentration in the membrane, ion leakage increased until reaching a level unsustainable by the cell. At this point, the membrane potential would collapse, which would in some way cause hole formation. Thus, the holin clock would simply reflect the titration of the membrane potential by the accumulation of holin. To test this model, Gründling et al. visualized cells induced for a clone of the lysis cassette while tethered to a surface by a flagellum (29). The rotational velocity of the flagellum is directly proportional to the proton motive force (pmf) of the membrane. The results had shown that rotation of the cells continued without change after induction of the lysis cassette until suddenly stopping at an allele-specific time. Lysis occurred within seconds after the sudden halt. This was interpreted as showing that the holin accumulated without deleterious effect on the membrane until it reached the triggering concentration, at which point hole formation occurred. The possibility that the accumulating holins were causing proton leakage but the cell was able to compensate by pumping protons out at a faster rate was ruled out by showing that, in uninduced cells, the flagellar rotation rate was reduced linearly as the concentration of an uncoupler, dinitrophenol (DNP), added

to the medium was increased. The linearity suggested that the cell was unable to compensate for the uncoupler and maintain the pmf at a constant level. It was also shown that DNP added at a level that reduced the rotation speed by ~ 40% was sufficient for premature triggering of the holin. This indicated that the lambda holin is sensitive to fractional changes in the pmf, i.e., that complete depolarization is not required.

Biochemical and structural aspects of the S holin

Despite many attempts to identify the molecular structure of the holin protein using X-ray crystallography and nuclear magnetic resonance (NMR), there is still no crystal structure of the S105 holin (or any other holin). NMR studies require purified protein in high concentration, and crystallography studies require crystal formation. Purification of holin protein has been difficult due to its toxicity and membrane localization. Fortunately, T7-based hyper-expression systems have been successful in achieving overexpression of the protein (30). When compared to parental controls, they exhibited a 50-fold increase in final holin concentration compared to the lethal level observed under physiological conditions. This overexpression developed within the first 10 minutes after induction via isopropylthio-β-galactoside (IPTG). The overexpressed S105 was localized in the membrane and extracted with the detergent. Many positions throughout the S105 primary structure were tested for retention of holin function after the insertion of a decapeptide (GGH_6GG) tag. It was found that the His-tag could be inserted between codons at positions 94 and 95 without affecting lysis phenotype. The presence of the His-tag allowed purification of 1-3 mg holin protein per liter of culture, which allowed physical and biochemical analysis of the holin (31). The results from

circular dichroism (CD) spectroscopy of purified holin were found to be detergent-dependent. When octyl-glucoside (OG) detergent was used, CD analysis indicated that the holin had ~ 40 residues in α-helical conformation, consistent with 2 TMDs. But when conducted with empigen-BB (EBB) detergent, the same analysis yielded results that described a holin with ~ 60 residues in α-helical conformation, consistent with three TMDs (32). Of the two analyses performed, only the results from the latter agree with previous topology studies (24).

Hole formation in vitro

The availability of milligram quantities of purified S protein allowed attempts at achieving membrane permeabilization in vitro. The most common system for detecting the formation of membrane pores has been the use of the dye-release assay, using unilamellar liposomes filled with calcein, which is self-quenched at the concentrations obtained in the liposome lumen (33-35). Permeabilization of the liposome results in release of the calcein and thus an increased fluorescence signal. This assay has been extensively used on cytolytic toxins (32). However, the S105 holin differs from cytolytic toxins in that it is integrated into the membrane prior to hole formation. Thus, the major barrier to studying hole formation by holins is the delivery of the holin protein to the bilayer, which in vivo is accomplished by the cellular secretion machinery (36). After diluting the S105 protein into a suspension of the liposomes, dye-release was obtained with the wt protein but not with purified A52V protein, thus mirroring the in vivo lysis phenotypes. In later experiments, dye release was observed at 30°C with purified $S105_{A55T}$ protein, but not at 42°C. Since the A55T allele is temperature-

sensitive (*ts*) for lysis in vivo, this experiment rigorously demonstrated that hole formation in these assays was S105-dependent rather than caused by a co-purifying contaminant. These assays, although clearly indicating that the holin is capable of permeabilizing membranes without the participation of other host or phage proteins, suffer from the defect that the membrane is not energized during the hole-formation event. So it is unclear whether the pathway to hole formation is the same in the liposome system as it is in vivo.

Recent advances in holin biology

Visualizing holin triggering in real time

In order to visualize the localization of S105, a functional S105-GFP chimera was constructed. Deconvolution fluorescence microscopy has shown that the S105-GFP fusion protein accumulates uniformly within the membrane. The holin remains uniformly dispersed until approximately the time (+/- 1 min) of lethal triggering, upon which it forms aggregates or rafts (37). Furthermore, fluorescence recovery after photobleaching (FRAP) data revealed that nonlethal mutant and untriggered hybrids protein are mobile in the membrane. But the rafts formed by the wild-type S105-GFP are immobile (37). The results of these studies indicated that triggering occurs when the holin reaches a certain critical concentration and nucleates to form rafts. The critical concentration of S105 molecules in the membrane at the triggering time is allele-specific. Early lysing mutants are found to have less S105 protein in the membrane at the time of lysis, while later lysing mutants have more protein in the membrane at the time of lysis. The allele-specific critical concentration model has a precedent in the

formation of the purple membrane after the induction of bacteriorhodopsin. The extensive studies of bacteriorhodopsin (BR) suggested that a single mutation could alter the concentration that is required for its oligomerization or "purple membrane" formation, and BR accumulates as a monomer until the critical concentration (38).

The nature of the S hole

Not surprisingly, previous cross-linking data of Cys51 indicated that S holin forms homo-oligomers in the membrane. Gel-filtration analysis of the purified S protein, which was extracted using OG detergent, implies that high-oligomer aggregates formed (32). Assuming that 1 nm is the approximate wall thickness of one TMD (39), the minimum number of TMDs necessary to allow the passage of the18 kDa endolysin is require approximately eight S105 molecules if all three TMDs were involved that would be 24 nm in diameter (40). This model was tested by using endolysin β-galactosidase (β-Gal) fusions, and tested for lytic competence in the context of the late gene expression system. The resultant fusion protein, R-β-Gal, (a homo-tetramer of ≥480 kDa), was fully folded and functional in S105 holin-mediated lysis (Fig. 7). This suggests that the S105 lesion, or hole, must be much larger than the minimum size proposed by (40), in order to accommodate the massive chimeric endolysin. Also, immunoblot and gel filtration analysis verified that the lytic function was not due to proteolytic release of the endolysin domain of the chimeric protein. Both early-lysis variants of S105 holin, and T holin from phage T4 were functional with the chimeric endolysin. In contrast, holes formed through premature triggering of the holin by using energy poisons were non-permissive for this massive complex, indicating that the size of the premature lesions

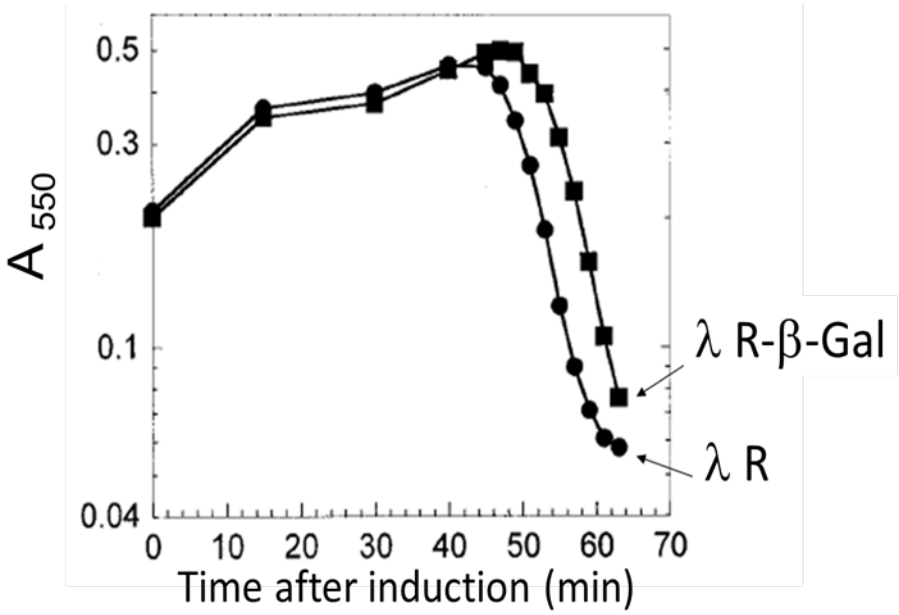

Fig. 7. Sizing the S hole with β-galactosidase fused endolysin. The circles show the cells expressing the wild-type lambda endolysin gene, and the squares show the cells expressing the endolysin fusion protein. Both strains were able to trigger hole formation and the sharp decline in A_{550} indicates that the endolysin was able to pass through the inner membrane and degrade the cell wall. These data suggested that lambda holin must form a lesion in the cytoplasmic membrane large enough to allow the passage of the 480 kDa active homo-tetramer chimeric endolysin. Adopted from Wang et al. (40).

are smaller (40). However, the structure of the hole and the identity of the TMDs lining the hole are unknown. Recent studies have used cryo-electron microscopy (cryo-EM) and tomography to visualize the hole caused by S105 *in vivo*. The lesions were found to be both heterogeneous and of unprecedented size (6) (Fig. 8). Surprisingly, the scale of the lesions was an order of magnitude greater than any previously described membrane pore. The average diameter was ~ 340 nm, with some exceeding 1 μm. The lesions were not localized to a specific region in the inner membrane but appeared randomly throughout its periphery, irrespective of size. Taking into account the geometry of viewing in vitreous ice, the observed diameters and the random positioning of the lesions, calculations suggested that there were ~ 2 holes per cell. Even though these membrane lesions are massive in size, the perimeters of the holes are still consistent with the amount of S105, which is between 1000 to 3000 molecules per cell at the time of triggering (20). Since each holin has three TMDs and an α-helix has a diameter of ~ 1 nm thickness, then 1000 S105 molecules could theoretically line a circular hole with 3 μm circumference, which would correspond to a ~ 1 μm diameter.

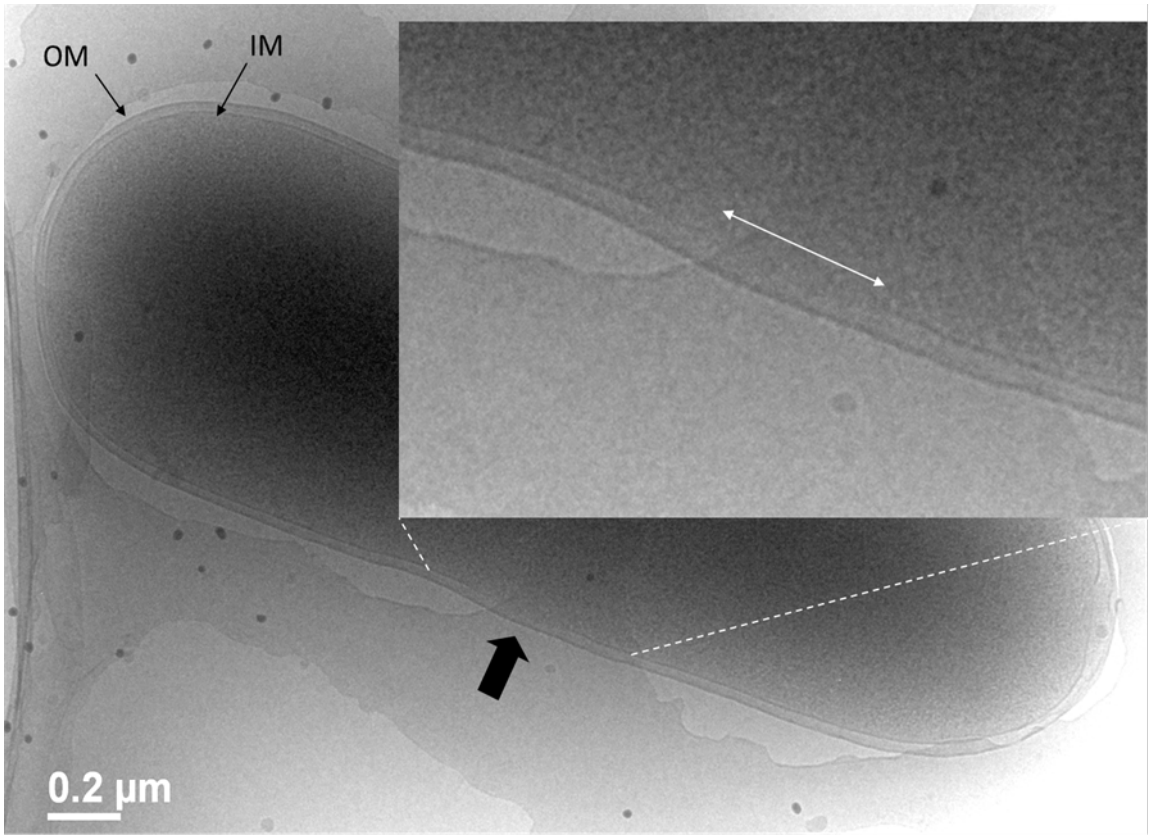

Fig. 8. Cryo-electron microscopy of an *E. coli* cell with holin induced. Cell expressing S105 is shown. Area enclosed in the white dashed line is shown enlarged. The white solid arrow indicates the location and extent of the lesion. IM, inner membrane; OM, outer membrane. Adopted from Dewey et al. (6).

Structural studies of S

The availability of purified holin protein enabled researchers to employ cryo-EM and single-particle analysis as a means to characterize lambda S105 holin *in vitro* (41). Ring structures were observed with S105 purified in the non-ionic detergent n-dodecyl beta-D-maltoside (DDM) and in the mild zwitterionic detergent EBB. The ring had a 8.5 nm inner diameter and 23 nm outer diameter, was 4 nm in height, and contained approximately ~ 75 S105 monomers, as estimated from mass-volume calculations (41). These ring structures were not found with purified protein of the lysis-defective variant $S105_{A52V.}$ Gel filtration showed $S105_{A52V}$ in micelles of a size consistent with dimers. However, these ring structures were judged not to reflect the structure of the S105 holes formed in vivo because the diameter of the lumen hole was far too small to accommodate the endolysin-β-galactosidase chimeras that were used to size the S hole in vivo previously. It was concluded that the 8.5 nm ring formation was a detergent artifact but nevertheless was indicative of the ability of S105 to form holes in a membrane bilayer, since the mutant protein, A52V, retained dimer status under the same detergent conditions.

Structural information is essential in order to understand holin function. Recent NMR efforts by Dr. Dewey provided additional information about the S105 holin structure (unpublished data from Jill Dewey and Younbok Lee). The construct examined was a truncated form of the lambda holin , $S105_{\Delta TMD1(C51S).}$ This mutant protein lacks the first TMD and also carried the C51S substitution to eliminate the possibility of oxidative damage that could occur during preparation (42). This

truncation was considered a good subject to be used for structural studies because it is small and relatively non-toxic to the membrane. The presence of a His_6 tag allowed for successful purification of 1-3 mg protein per liter culture. The purified holin protein remained soluble in DDM detergent for weeks at 4 °C.

$S105_{\Delta TMD1}$ contains only two TMDs and is thought to mimic the canonical lambda antiholin, S107, in the membrane. Sequence-specific chemical shift assignments have been obtained for two-thirds of the polypeptide sequence $S105_{\Delta TMD1}$, a variant of the holin. Secondary structure elements were identified based on Cα chemical shift deviations and backbone dynamics data (43). These observations are consistent with the previous findings regarding the locations of the two transmembrane helices TMD2 and TMD3 (21). In addition, the NMR data suggested that the helix of TMD3 was extended below the membrane. I hypothesize that the extended TMD3 helix projects into the cytoplasm from the bottom of TMD3 and could be involved in the regulation of hole formation. Both S107 and $S105_{\Delta TMD1}$ delay hole formation when expressed *in trans* to S105, and therefore are considered to be functionally relevant proteins (44). The structure of the holin-antiholin complex is expected to mimic holin-holin interactions.

A new type of holin: pinholins

Unlike holin S105, which makes large, non-specific holes and can promote the lytic function of a variety of endolysins, the S^{21} holin from phage 21 represents another type of holin. S^{21} forms tiny holes that are too small for the passage of lambda R and T4 E or periplasmic GFP (45). These holes are called “pinholes” and phage 21 holin is known as a pinholin. The small size of pinholes prevents cytoplasmic endolysins from

entering into the periplasmic space and accessing the peptidoglycan. Consequently, phages that utilize pinholins must transport their endolysin across the plasma membrane by a mechanism different from phages that use canonical holins. The phages encoding pinholins also encode SAR endolysins (12). These endolysins are exported across the plasma membrane via the host secretory system. Exported SAR endolysins accumulate in an enzymatically inactive form, tethered to the membrane by an N-terminal TMD. When the pinholin triggers, the membrane depolarizes. This allows the SAR endolysin to be released from the bilayer and refold to a catalytically active state. SAR endolysins are briefly described below.

Despite the difference in molecular function, the lysis genes of phage 21 are arranged in a lysis cassette analogous to phage lambda. Because of this architectural similarity, the lysis genes are named as they are in lambda: S^{21}, R^{21}, Rz^{21} and $Rz1^{21}$. As in lambda, the phage 21 lysis cassette is transcribed by the late gene promoter $P_R{}^{21}$. Q^{21} transactivates this late gene promoter $P_R{}^{21}$ at the onset of late gene expression (45) (Fig. 9A).

Like lambda holin *S*, S^{21} also has a dual-start motif, encoding both a holin $S^{21}68$ (68 amino acids), and an antiholin $S^{21}71$ (71 amino acids), produced at a ratio of ~ 2:1 (46). The two primary structures are different only by the tripeptide Met1-Lys2-Ser3 at the N-terminus of $S^{21}71$. Both $S^{21}68$ and $S^{21}71$ have two TMDs (45) (Fig. 9A). The topology of $S^{21}68$ is dynamic. Data from previous studies indicated that there is an absence of disulfide bond formation when $S^{21}68$ is tethered in the membrane. The current model is that the first TMD of $S^{21}68$ contains a SAR (Signal Anchor Release)

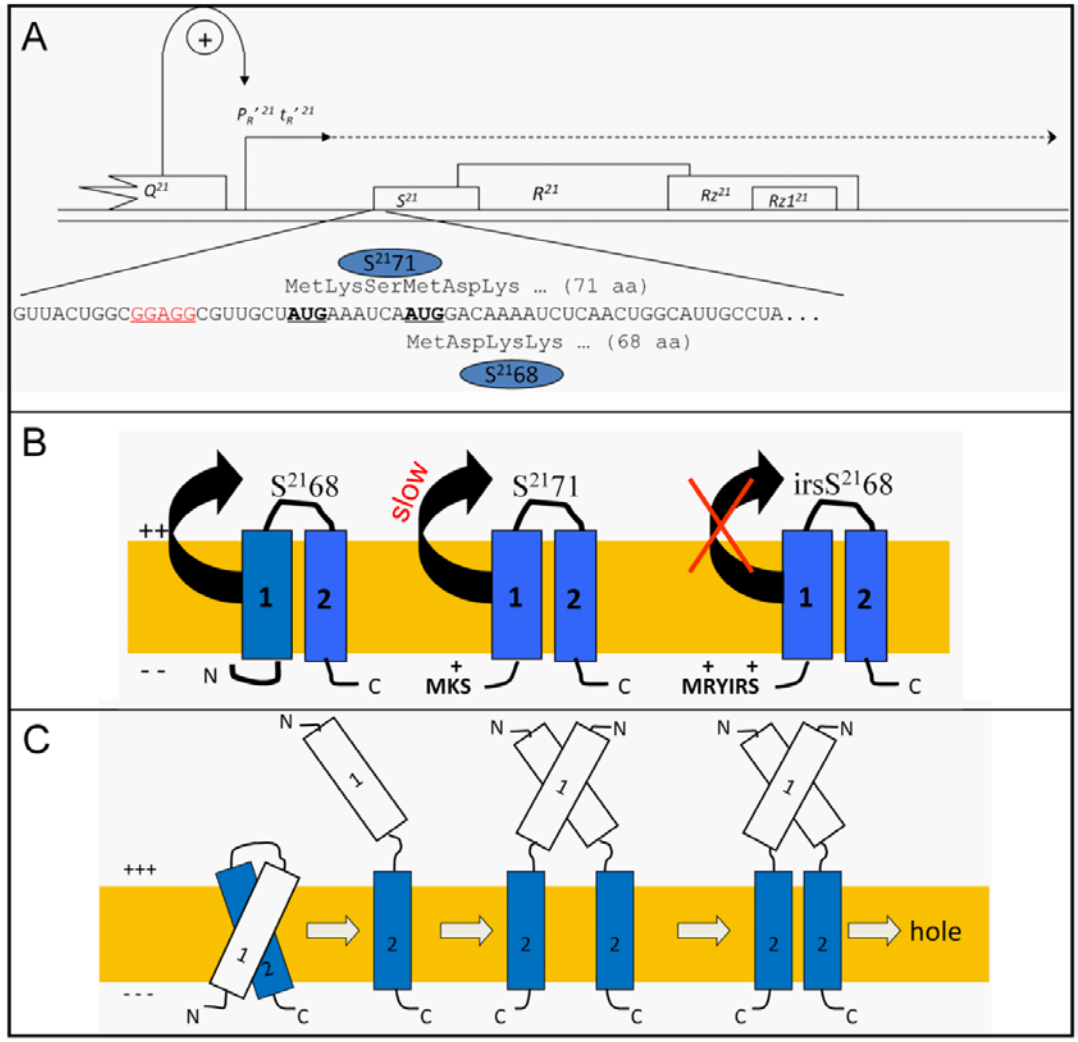

Fig. 9. Features of the phage 21 lysis system. (A) phage 21 lysis cassette. Similar to the lysis cassette of lambda, the lysis genes *S*^{21}, *R*^{21}, *Rz*^{21} and *Rz1*^{21} are transcribed from late promoter $P_R{}'^{21}$, which is transactivated by Q^{21}. The mRNA structure and corresponding amino acid at the beginning of *S*^{21} are shown in below. *S*^{21} also has a dual-start motif, encoding both a holin, S^{21}68, and an antiholin, S^{21}71. The Shine-Dalgarno sequence for both S^{21}68 and S^{21}71 is underlined and colored in red. (B) The membrane topology of holin S^{21}68, the weak antiholin S^{21}71, and the strong antiholin irsS^{21}68. Differences at the N-terminus are shown. S^{21}71 has one additional positive charge while irsS^{21}68 has two. The TMD1 is a SAR-domain. The release of TMD1 is delayed in S^{21}71, and completely blocked in irsS^{21}68. It is likely that the positive charge at N-terminus blocks its release. (C) The model of hole formation by S^{21}68. Initially, the two TMDs are inserted in the membrane, with TMD1 binding to TMD2. When TMD1 is released in the periplasm, periplasmic TMD1-TMD1 interactions occur. TMD2s are then free to oligomerize and form holes. Adopted from Park et al. (45).

domain that allows it to exit the inner membrane into the periplasm spontaneously (47). Lethal hole formation can occur only after removal of TMD1 from the inner membrane (47). An $S^{21}68$ allele that lacks the first TMD, $S^{21}68_{\Delta TMD1}$, was found to still exhibit holin function. This indicates that only the second TMD of $S^{21}68$ is essential for host lysis (47).

Furthermore, one construct of $S^{21}68$ was found to be completely lysis defective. This construct, *irs*$S^{21}68$, has an N-terminal modification in which the residues Arg-Tyr-Ile-Arg-Ser are inserted. The results suggested that the addition of the two positively charged residues abolishes holin function by preventing TMD1 from exiting the membrane (Fig. 9B). The *irs*$S^{21}68$ allele showed a dominant-negative phenotype when co-expressed *in trans* with $S^{21}68$, thus exhibiting its strong antiholin properties. The data indicated that the release of TMD1 from the membrane coincided with the triggering time of $S^{21}68$. This suggested that the membrane exit of TMD1 is the rate-limiting step of pinhole formation. The first TMD of antiholin $S^{21}71$ is capable of exiting into the periplasm; however, it does so at a rate slower than $S^{21}68$ (Fig. 9B). This allows $S^{21}71$ bind to $S^{21}68$ and delay the triggering of $S^{21}68$ (47). In terms of the pinhole formation model, Park et al. (45) proposed that two TMDs of $S^{21}68$ are initially inserted in the membrane, and then the TMD1s are released gradually into the periplasm, which allows TMD2 to oligomerize and form pinholes. However, it is not known whether all pinholins are of class II topology and/or require an N-terminal TMD that can escape from the bilayer.

Recent studies by Pang et al. have further explored the details of the pinhole formation mechanism of S^{21} and defined the final pinhole structure using both biochemical and computational approaches (48). The results suggest a structural model of the pinhole as a symmetric heptamer, with the hydrophilic side of TMD2 lining the central channel of ~ 1.5 nm in diameter. The hydrophilic face of the TMD2 appears to be facing the lumen based on the results from the cysteine-accessibility assay. This structural model also identified two helix-helix interaction surfaces for each TMD2. The authors proposed that the core of one of the surfaces is formed by a glycine zipper motif, which forms a sterically unhindered pocket that can accommodate the side chains of the bulky residues (48).

An extensive collection of S^{21} mutants was constructed to identify residues and domains essential to the function of the pinholin (49). A wide range of phenotypes, from absolute lysis-defective to early lysis triggering were observed for mutations mapping to each topological domain. The results suggest that TMD1 of $S^{21}68$ acts *in trans* as an inhibitor of the pinholin function within the inactive dimer stage (Fig. 9C). Furthermore, Pang et al. proposed a model for the structure of the inactive dimer, which identifies the faces of the two transmembrane domains involved in intramolecular and intermolecular interactions, as well as those interacting with the lipid (49).

In addition, Pang et al. observed the localization of GFP fusion constructs of $S^{21}68$ and its variants was determined using fluorescence microscopy (50). Functional pinholin mutants including the wild-type pinholin GFP fusion formed smaller rafts and more numerous small aggregates in the membrane at the time of triggering when

compared to canonical *S* holin. In contrast, GFP fusions of non-lethal mutants, as well as irsS^{21}68, were distributed uniformly in the membrane. IrsS^{21}68 was capable of preventing S^{21}68::GFP from triggering to a punctuate distribution. These results demonstrate that formation of the rafts is associated with pinhole formation (50). Combining both crosslinking and single cysteine accessibility modification results suggest that most of the S^{21}68 are involved in pinholes. If there are >6000 S^{21}68 molecules present in the cell at the time of triggering, there must be ~ 900 pinholes suddenly forming once triggering occurs (48).

SAR endolysins

The first SAR endolysin discovered was Lyz from *E. coli* phage P1 (51). Like all members of this group of endolysins, Lyz has an N-terminal TMD domain. This TMD contains a signal sequence that allows the protein to be secreted using the host's *sec* dependent pathway. Once transported across the plasma membrane, Lyz is not cleaved from its signal sequence and consequently remains tethered to the embedded TMD (45, 47) (Fig. 10). To highlight the unprecedented capacity for escaping the bilayer, the signal-anchor TMD of Lyz was designated as a SAR (signal anchor release) domain. The ability of a SAR domain to exit from the membrane is dependent on a high content of small, weakly hydrophobic and polar residues. P1 Lyz contains seven cysteine residues, six of which are in the periplasmic domain and one in the SAR domain. The first cysteine, Cys13, is embedded in the inner membrane and retains its sulfhydryl form, but all other six cysteines are incorporated in disulfide bonds (52). Among them, the third cysteine Cys51, which forms the catalytic triad together with

Glu42 and Thr57 in the active enzyme, forms a disulfide bond with the second cysteine Cys44 (Fig. 10). The resultant membrane inserted Lyz is inactive. The release of the SAR domain from the membrane activates Lyz upon disulfide isomerization. Then it leaves the catalytic Cys51 free while Cys13 and Cys44 form a disulfide (Fig. 10). In the meantime, the catalytic domain undergoes a conformational change and forms a catalytic cleft similar to lysozymes (52).

Recent bioinformatic surveys have shown that the P1 Lyz-like SAR endolysins are not the most prevalent class of SAR endolysins. The phage 21 endolysin R^{21}, which does not have either a cysteine in the SAR domain or a catalytic cysteine, is more representative of this protein class as a whole (53). Similar to Lyz, the release of the SAR domain from the membrane is an essential step in R^{21} endolysin activation (53). Both Lyz and R^{21} are capable of degrading the host peptidoglycan without the presence of holin because they are spontaneously released into the periplasmic space. However, this form of lysis is delayed and gradual, because spontaneous release is very slow (45, 51). The addition of energy poison (KCN) or an uncoupler (DNP) is shown to induce pinholin function immediately (47, 54). These results suggest that when the pinholin triggers and depolarizes the inner membrane, the SAR endolysin escapes from the bilayer and folds into its enzymatically active form, resulting in degradation of the peptidoglycan (47, 54) (Fig. 10).

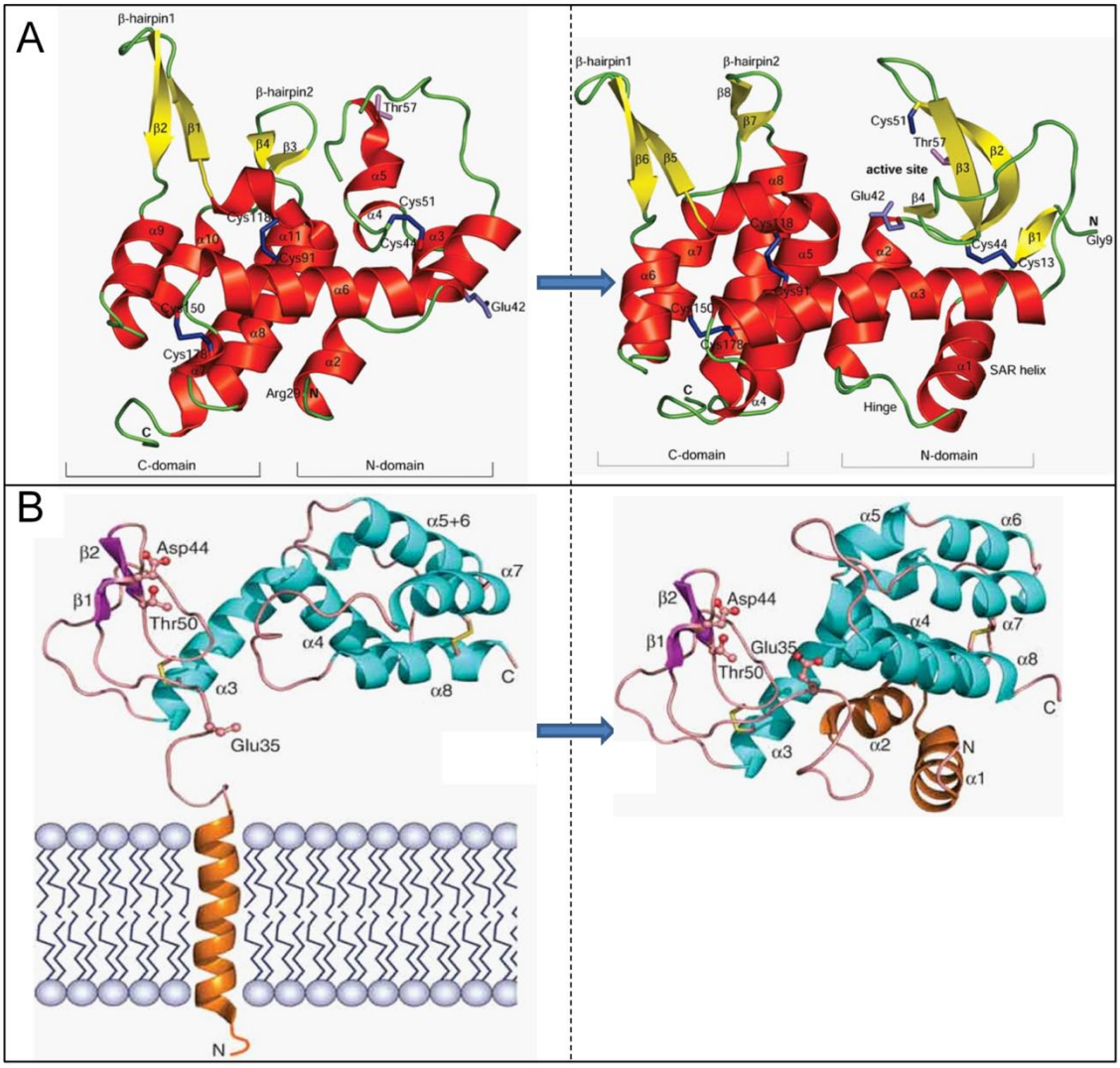

Fig. 10. Topological and conformational changes upon SAR-endolysin activation. (A, P1 Lyz; B, R^{21}). In both A and B, the crystal structures of inactive SAR-endolysins (^{i}Lyz, $^{i}R^{21}$) are shown on the left, while the active ones (^{a}Lyz, $^{a}R^{21}$) are on the right. **A.** Colors: α helices, red; β strands, yellow; loops, green; disulfides, blue. The SAR-domain in ^{i}Lyz (residues 1 to 28) is truncated. ^{a}Lyz has additional strands β1 to β4, but lacks helices α3 to α5 of ^{i}Lyz. In addition, ^{a}Lyz has disulfide bond formed between Cys13 and Cys44, instead of Cys44 and Cys51 in ^{i}Lyz. Adopted from Xu et al. (52). **B.** Colors: α helices, cyan; β strands, magenta; coils, salmon. The SAR domain is shown in orange and is depicted as a membrane-spanning helix in $^{i}R^{21}$. The catalytic triads (Glu35, Asp44, Thr50) are represented in stick-and-ball form. Disulfide linkages are represented in yellow stick. Adopted from Sun et al. (53).

The T4 lysis paradigm

The study of phage lysis started with phage T4 in the 1940s, in the pre-molecular era of phage biology. In fact, the term "endolysin" was created for the lysozyme of phage T4, the protein product E of gene *e* (55), more than 50 years ago. Despite this long history, the T4 holin was not identified until 1992 and not characterized until 2001 (56, 57). The T4 holin, the protein T encoded by gene *t*, is a 218 amino acid polypeptide with no sequence similarity to lambda holin S105, and has only one identifiable TMD (57). The T polypeptide can be divided into three domains by primary structure analysis: an N-terminal cytoplasmic domain of 32 aa, a single TMD, and a large periplasmic domain. Replacing *S* with *t* in the context of the lambda lysis cassette results in normal, sharply defined lysis, except that the timing of lysis is much earlier at about 20 minutes, compared with 50 minutes for lambda S. The ability to function with the lambda R endolysin suggests that T also forms non-specific holes in the membrane. As with *S*, single amino acid changes in *t* can alter the lysis timing significantly, both advancing or retarding the onset of lysis (5). This further reinforces the idea that the malleability of the lysis timer built into the holin gene enable phages to rapidly evolve to altered lysis times, to increase fitness when the environment changes.

Lysis inhibition

The unique structure of T holin, specifically the periplasmic domain, is thought to serve the purpose of supporting a phenomenon called Lysis Inhibition (LIN) (58). In LIN, the normal holin-dependent lysis triggering is blocked if, ~ 5 min or more after the initial infection, the infected cell is subjected to a secondary infection by another T-even

phage. In these secondary infections, the normal infection process is blocked by several early gene proteins produced by the original infecting phage, and as a result, the contents of the head of the secondary infecting phage are ejected into the periplasm, rather than the cytoplasm (16) (Fig. 11). By an as yet undetermined mechanism, this ectopic release of the virion contents causes activation of another T4 protein, RI. RI is a SAR protein and thus is synthesized as an integral inner membrane protein with an N-terminal SAR domain. In the absence of superinfection, RI is rapidly released into the periplasm and degraded, with a half-life of about 1.5 min. However, somehow the ectopic superinfection generates a signal that activates RI, which then binds to the periplasmic domain of T and blocks its hole-formation function. Operationally, RI is thus an inducible antiholin of phage T4. If the SAR domain of RI is replaced by a cleavable signal sequence, induction of RI leads to stable LIN, irrespective of the superinfection signal. This suggests that the SAR domain of RI not only confers the ability of RI to be localized and released into the periplasm but also that it confers proteolytic instability on RI. Activation may thus simply reflect a stabilization of RI, although how that would be affected is unknown. Besides the 170 kb dsDNA molecule of T4, approximately 1000 protein molecules are also ejected into the periplasm during the superinfection. If T4 is converted to empty "ghosts" by osmotic shock, the empty phage particles still go through normal adsorption and tail contraction. However, the ghosts do not induce LIN, suggesting that the head contents are responsible for the LIN signal. Whether the DNA or the proteins constitute the signal is unknown.

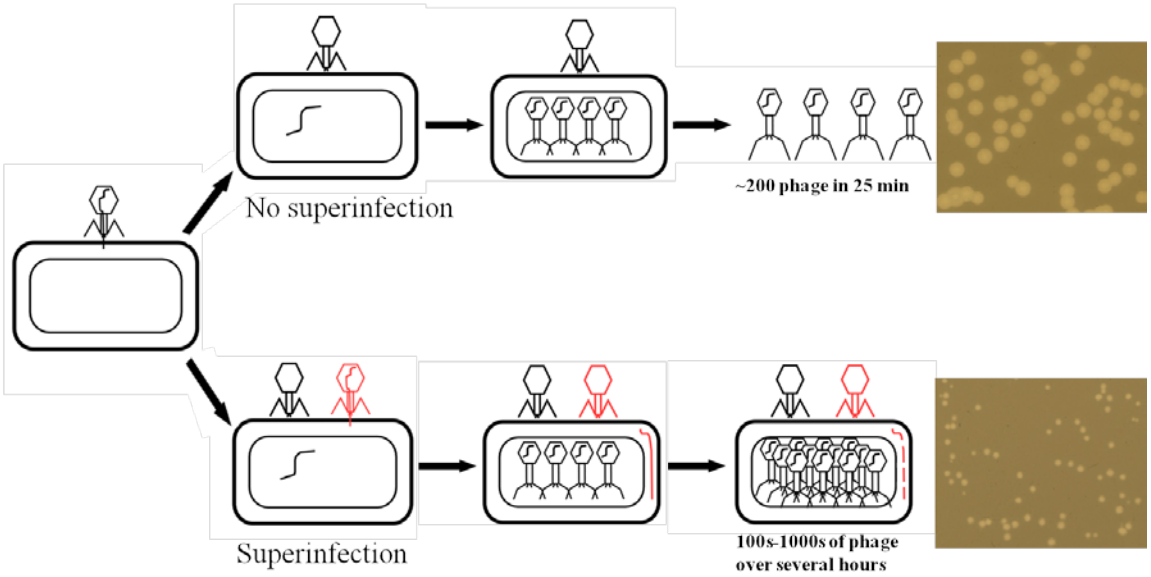

Fig. 11. LIN extends the vegetative cycle and increase burst sizes. Cartoon of T4 infection with no superinfection and lysis (top) and superinfection and LIN (bottom). Adopted from Dr. Samir Moussa.

Although it eventually collapses, LIN allows phage T4 to delay host cell lysis for hours (57, 59), which permits a 10-fold increase in intracellular accumulation of progeny virions. How does LIN improve phage fitness? From an evolutionary fitness point of view, the presence of free T-even phage particles in the medium is a signal that the uninfected host cell population has been reduced. Thus, instead of releasing progeny virions into the environment and have them adsorb to the already infected cells, the successful infection cycles are prolonged, allowing continued, linear increase in progeny yield.

Classification of holin

Only a few holins have been experimentally characterized. In the last broad survey nearly 12 years ago (12), holins were classified into three topological groups (Fig. 12). Class I holins, (e.g., S105 holin of phage lambda) contain three TMDs with the N-terminus located in the periplasm and the C-terminus located in the cytoplasm (5, 12, 24) (Fig. 12). Members within Class I consist of 95-130 amino acids, which belong to 7 unrelated families in Gram-negative hosts. Some other class I holin prototypes with no sequence similarity are Y from phage P2, LydA from phage P1, and gp13 from phage P22. Class II holins (e.g. the $S^{21}68$ holin of phage 21) contain two TMDs with both the N-terminus and C-terminus located in the cytoplasm. Members within Class II consist of 65-95 amino acids, which belong to 6 unrelated families in Gram-negative hosts (1, 60) (Fig. 12). Finally, class III holins, (e.g., the T holin of phage T4 and all T4 and T5 related phages) contain only one TMD with the N-terminus located in the cytoplasm and the extended C-terminus located in the periplasm (12) (Fig. 12).

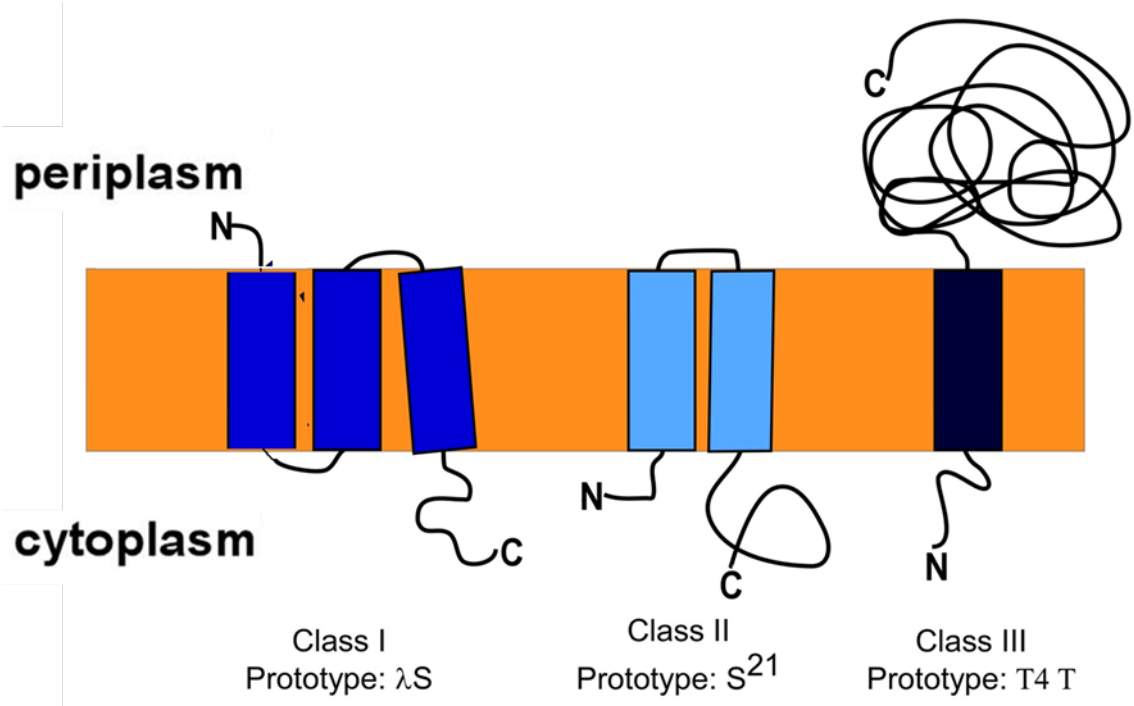

Fig. 12. Three topological classes of holin proteins. The Class I holin contains three TMDs with the N-terminus out and C-terminus in topology. The example shown here is lambda S105. Class II holin contains two TMDs with an N-terminus in and C-terminus in topology. The example shown here is $S^{21}68$ from phage 21. The Class III holin contains one TMD with an N-terminus in and C-terminus out topology. The example shown here is T from phage T4.

It should be emphasized that the TMDs have been experimentally verified for only a few holins, whereas the majority of membrane-spanning regions and their orientations are predicted by manual inspection of hydrophobic domains without charged amino acids and bioinformatic tools such as TM Hidden Markov Model (TMHMM) and TM prediction (TMpred).

In summary, holins play an important biological role in controlling lysis timing in bacteriophages. Although lambda S105 is the paradigm holin, little is known about the final structure of the hole in the membrane, or if the S105 holin features can be applied to other phages with no primary structure similarity to S105. Although more than 1000 complete phage genomes have been contributed to the public database, no further comprehensive analyses of other holins have been done. The purpose of this dissertation is to use both genetic and biochemical methods, as well as bioinformatics tools, to further study different classes of holin.

Specific aims of the dissertation research

Chapter II: I chose to examine a putative class I holin Y from the coliphage P2, which has a different primary structure but the same predicted topology as S105, in order to determine if the lessons learned from lambda S105 can be generalized. This study is pertinent because only one class I holin, S105, has been studied in detail so far. In particular, *Y* was subjected to random and site-directed mutagenesis to obtain a library of absolute and conditional mutants. These mutants with defects in hole-formation and lysis timing were characterized in terms of protein accumulation, localization, and oligomeric state. These efforts examined whether Y also uses a dimer-oligomer-raft

pathway and offer important controls for biochemical, physiological, structural and cell-biological studies.

Chapter III: I interrogated the primary structure of S105 for features that are critical for the timing of hole formation. To this end, I combined genetic, phenotypical and biochemical analysis methods, including cysteine-scanning-accessibility, to identify which TMD residues line the S105 canonical hole. The results from these experiments were used to formulate the first comprehensive model of the S105 hole formation pathway.

Chapter IV: Virtually all previous experiments monitoring membrane integrity as holin accumulates have been focused only on lambda S105. I determined whether the temporal and "all or nothing" responses related to S105 triggering at the single cell level are universal characteristics. Here, I monitored the integrity of the inner membrane of another class I holin, Y from phage P2, based on the primary structure differences but the same predicted topology as S105. Furthermore, I performed the same study on the class II holin $S^{21}68$ from phage 21, based on the different molecular function compared to canonical holin.

Chapter V: I identified and characterized the holin and antiholin in Mu, which is one of the few paradigm coliphages. Genetic studies coupled with physiological studies predict that the holin and antiholin in Mu have a new holin topology. I also surveyed phage genomes of Gram-negative hosts to identify the best holin candidate genes, based on criteria such as sequence similarity to known holins, predicted protein architecture,

and gene clustering. This survey revealed that there are several new topological classes of holin.

CHAPTER II

FUNCTIONAL ANALYSIS OF A CLASS I HOLIN, P2Y*

Introduction

Holins are small, phage-encoded membrane proteins that control the length of the bacteriophage infection cycle by determining the time of host lysis (5, 12). Almost all the experimental work done on holin function at the physiological or molecular level has been focused on a few holin genes: lambda *S*, phage 21 S^{21}, and T4 *t*. These genes encode the holins S105, $S^{21}68$ and T, which are respectively the prototypes of three distinct holin classes, distinguished by experimentally-determined membrane topology (Fig. 13A). S105 has class I topology, with three TMDs arranged N-out, C-in, whereas $S^{21}68$ has class II topology, with two TMDs arranged N-in, C-in. As of the last major review, ~ 90% of the more than 100 genes from phages and prophages of eubacteria that have been assigned as putative holins fit into either class I or class II, based on predicted topology using widely accepted algorithms (5). These predicted holins are very diverse, with 12 and 14 unrelated sequence families grouped into Class I and II, respectively. Class III topology, with a single TMD and a large periplasmic domain, is restricted to the sequence homologs of protein T, the holin of T4. These proteins are found only in phages of the large myophage class of T4-like phages and in the large siphophage class of T5-like phages.

---

*Reprinted with permission from "Functional analysis of a class I holin, P2 Y" by To KH, Dewey J, Weaver J, Park T, Young R, 2013. *Journal of Bacteriology*, 195 (6), 1346-1355, Copyright [2013] by American Society for Microbiology.

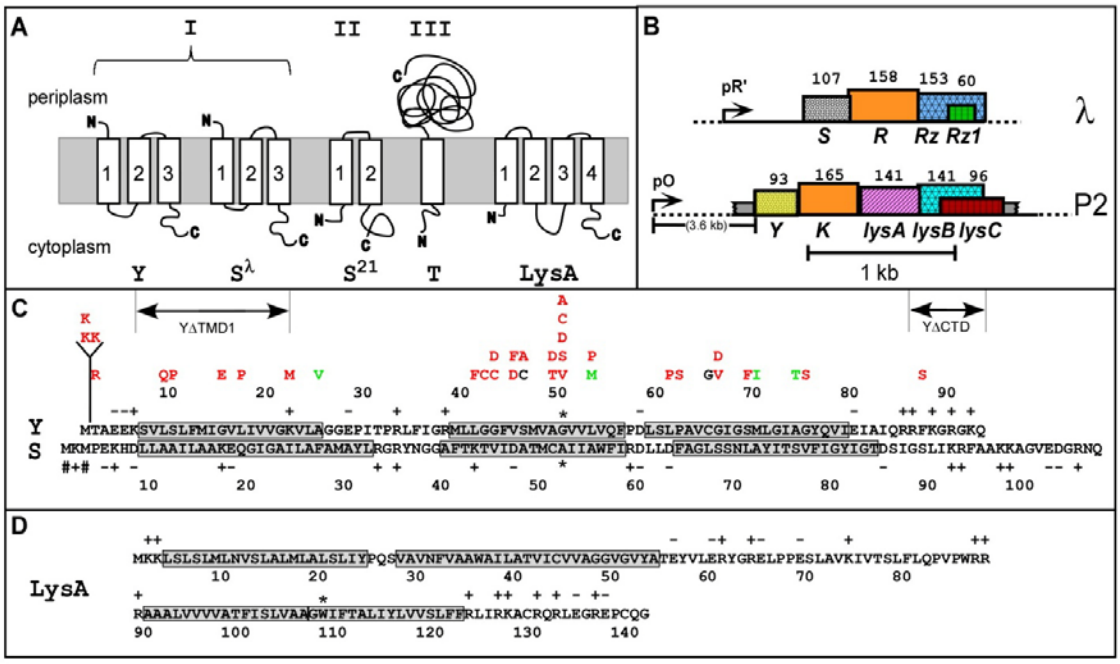

Fig. 13. Genes, sequences and topologies. (A) Membrane topologies of P2 Y, □ S105, phage 21 S^{21}, T4 T, and P2 LysA (see panel D). Three topology classes of the holins are indicated in roman numerals. (B) The lysis cassettes of P2 and λ. P2 genes *Y, K, lysA, lysB* and *lysC,* are shown, along with late promoter pO. λ genes *S, R, Rz,* and *Rz1* are shown, along with late promoter pR'. Genes are color-coded according to primary structure homology. Hatching pattern corresponds to homologous function. (C) Primary structures of Y (top) and S (bottom). Insertion, deletion and missense changes relevant to the text are shown. Mutants are indicated above the *Y* sequence, color-coded to phenotype: red = non-lytic; green = early lysis; black = normal lysis timing. Dual start residues for S107 and S105 are indicated by # signs, and the extent of the ΔTMD1 and ΔCTD deletions. Asterisks indicate Ala52 residue in S105 and Gly52 in Y, chosen as analogous small residues in central region of TMD2. (D) Primary structure and predicted topology of LysA. TMDs 1 and 2 were predicted by algorithms TMHMM (61) and TMPRED (62). Both programs predict a single TMD between 91 and 121. However, the homologous Y protein from a P2-like prophage in *Enterobacter sp. 638* has a Lys residue at the equivalent position to Trp108 in Y, indicated by asterisk, strongly suggesting that there are two TMDs in this hydrophobic stretch. This allows the highly basic N and C termini of LysA to be located in the cytoplasmic phase.

Despite this diversity, holins share some universal functional features, based mainly on extensive work done with these three holin paradigms. Holin genes are turned on at the beginning of viral morphogenesis, so that the holins accumulate continuously and harmlessly in the cytoplasmic membrane while the viral structural proteins are synthesized and virus particles are assembled. For the prototype class I and II prototypes, S105 and $S^{21}68$, the holin accumulates as harmless, mobile dimers in the membrane (37, 63-65). There is no detectable effect on membrane integrity, respiration, pmf or energy metabolism until suddenly there is an irreversible, holin-mediated halt of respiration, collapse of the pmf, release of small ions and small molecules, and loss of viability. This event is designated as holin "triggering" (5, 12). For the class I and class II prototypes S105 and $S^{21}68$, these effects have been shown to be caused by the sudden formation of lethal membrane lesions, or holes. Moreover, triggering of these paradigm holins can be provoked prematurely by treatment with energy poisons, indicating that the timing of hole formation depends on the energization of the membrane. In the absence of holin function, lysis does not occur, so that virions and endolysin continue to accumulate indefinitely in the cytoplasm; the lysis defect is instantly relieved by the addition of chloroform, in effect constituting chemical complementation of the hole-forming function.

Despite these phenotypic commonalities, the prototype holins also exhibit unique features. The holes formed by S105 were known to be very large, capable of releasing molecules of the size of the homotetrameric, 0.5 MDa β-galactosidase enzyme in vivo (66). Purified S105 forms highly multimeric rings in detergent, estimated at >70 copies

per ring (67). In contrast, the $S^{21}68$ holin forms very small heptameric pores, estimated to be ~ <2 nm in diameter, both in vivo and, when purified, in detergent (48). The S105 protein is extremely sensitive to missense mutations within the TMDs, with single amino acid changes causing dramatic acceleration, retardation, or abolishment of the triggering time (68, 69). This has been interpreted as indicating that the timing of holin triggering involves packing of the TMD helices within the bilayer.

No other class I holin protein has been studied in detail, so it has been unclear how many of the striking features of S105 would be characteristic of other holins of this class, unrelated in terms of sequence. Among the paradigm phages, coliphage P2 has a lysis cassette that includes a putative holin gene, Y, encoding a protein with predicted class I topology, adjacent to the gene for the endolysin, K (Fig. 13B). Ziermann et al. (70) constructed a P2 Yam mutant and showed that it exhibited a lysis defect suppressible by the addition of chloroform, indicating that Y functioned as the P2 holin. Here I report results of experiments designed to test this notion rigorously and to determine whether Y shares the extensively described physiological and biochemical features of the lambda S105 holin. The results are discussed in terms of a general model for holin topology and function.

Materials and methods

Materials, strains, bacteriophages, plasmids, and growth media

Strains, phages, plasmids and primers used in this work are listed and described in Table 1. The plasmids pYRRzRz1 (renamed pKT1-Y) and pYRamRzamRz1am (renamed pKT2-Y) are derivatives of the plasmid pS105, a medium copy plasmid which

has the entire λ lysis cassette under its native late promoter (31). The media, growth conditions, and thermal induction of the λ lysis genes from a prophage and/or plasmid, have been described previously (20, 31, 71). All associated mutants were made by commercial available error-prone PCR kit (Stratagene), and standard site-directed mutagenesis (71). Bacterial cultures were grown in standard LB medium supplemented with ampicillin (Amp; 100 μg/mL) and chloramphenicol (Cam; 10 μg/mL) for the maintenance of plasmids and prophage, respectively.

$Y_{\Delta TMD1}$ or $Y_{\Delta C\text{-}term}$ was expressed from pKT1-Y as mentioned above. The *Y* gene was replaced by truncated $Y_{\Delta TMD1}$ gene (renamed pKT1-$Y_{\Delta TMD1}$) or truncated $Y_{\Delta C\text{-}term}$ gene (renamed pKT1-$Y_{\Delta C\text{-}term}$). Either pKT1-$Y_{\Delta TMD1}$ or pKT1-$Y_{\Delta C\text{-}term}$ was expressed *in trans* to full-length Y. Two lysogenic strains were used for expression of *Y;* MC4100*ΔfhuA*(λ*Cam*Δ*SR*) was used for expression of *Y* and the lambda lysis genes *R*, *Rz*, and *Rz1* from pKT1, and MC4100 (λ-*Y*) was used for expression of *Y* and the lambda lysis genes *R*, *Rz*, and *Rz1* from the integrated prophage. *lysA* was expressed from pS105 without the entire lysis cassette and under lambda native late promoter (renamed plysA). plysA was expressed *in trans* to MC4100 (λ-*Y*) and MC4100 (λ-*S105*). MC4100 (λ-*S105*) was used for expression of S105 and the lambda lysis genes *R*, *Rz*, and *Rz1* from the integrated prophage.

Table 1. Strains, phages, plasmids and primers used in this work.

Phage	Genotypes and relevant features	Sources, references
λ Cm^r *Δ*(*SR*)	*stf*::*cat*::*tfa cI857 Δ(SR)*; replacement of *stf* and *tfa* genes (λ nt 19996–22220) with *cat* gene; *Δ(SR)*(72); deletion of λ nt 45136–45815 (68)	Laboratory stock
λ-*S105*	*cI857* S_{M1L}; encodes *S105* only	Laboratory stock
λ-*Y*	*cI857*; *S105* deleted, replaced by *Y (nt 6691-6997 of phage P2).*	This study
Strains	**Genotypes and relevant features**	**Sources, references**
MC4100	*E. coli* K-12 *F⁻ araD139* Δ(*argF-lac*)*U169 ΔfhuA rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR*	(73)
XL1-Blue	*E. coli* K-12 *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB* $lacI^qZ_{\Delta M15}$::Tn*10* (Tet^r)]	Stratagene
Plasmids	**Genotypes and relevant features**	**Sources, references**
pS105	lambda lysis cassette under pR' on pBR322 backbone; S_{M1L} allele, encodes *S105* only	(74)
pKT1-Y	Isogenic to pS105; *S105* replaced by *Y*; *(nt 6691-6997 of phage P2)*	This study
pKT2-Y	Isogenic to pKT1-Y; $R_{T54am,K60am}Rz_{S100am}Rzl_{W38am}$	This study
pKT2-Y_{am}	Isogenic to pKT2-Y; Y_{S46am}	This study
pKT2-S105	Isogenic to pS105; $R_{T54am,K60am}Rz_{S100am}Rzl_{W38am}$	This study
pKT1-$Y_{\Delta C\text{-}term}$	Isogenic to pKT1-Y; Arg86 to Gln93 deleted	This study
pKT1-$Y_{\Delta TMD1}$	Isogenic to pKT1-Y; Ser7 to Lys22 deleted	This study
pKT1-Y_{+K}	Isogenic to pKT1-Y; one Lys inserted after first Met	This study
pKT1-Y_{+2K}	Isogenic to pKT1-Y; two Lys inserted after first Met	This study
pLysA	Isogenic to pS105; S105*RRzRz1 deleted and replaced by lysA: nt 7487 to 7934 of phage P2*	This study
pc-mycR::lacZ	*mycR***::***lacZ* replacement of *luc* in pZA32-luc	(66)

Table 1. Continued

Primers	Sequence	Sources, references
pKT1-Yfor	5’GCGGTATTTCACACCGCACCTGGTGCACTCTCAG3’	This study
pKT1-YRev	5’CTGAGAGTGCACCAGGTGCGGTGTGAAATACCGC3’	This study

Mutagenesis of P2 Y

Error-prone PCR with Mutazyme II DNA polymerase was performed using the GeneMorph II Random Mutagenesis Kit (Stratagene) as recommended by the supplier. Each amplification reaction (50 µl) contained 0.1 µM each of the primers pKT1-Yfor and pKT1-Yrev (Table 1), 40 µM each of the dNTPs, 5µl of 10X Mutazyme II buffer, 4.5 µg of pKT1-Y template, and 1µl of Mutazyme II DNA polymerase. PCR was done with the following program: initial 5 min at 95°C, 20 cycles of amplification (30 sec at 95°C, 30 sec at 44°C, and 1 min at 72°C), and final 10 min hold at 72°C. PCR products were treated with DpnI, digested with NdeI and HindIII, which are the restriction sites flanking the *Y* gene, then re-ligated back into NdeI and HindIII-digested pKT1-Y as above. The ligated DNA was transformed into Xl1-Blue and the transformant colonies (~ 500 colonies) were slurried and used to prepare a plasmid library preparation. This DNA was transformed into the lysogenic host MC4100(λCm^r*Δ*(*SR*)). The transformants were again slurried in 25 ml of LBAmpCam and aerated at 30°C. After standard thermal induction at A_{550} ~ 0.2 and lysis, the lysate was filtered through a 0.45 µm filter to collect the cells that did not undergo lysis. The cells collected on the filter were resuspended in 5 ml LBAmpCam and the resuspended cells were used to prepare plasmid DNA in a final volume of 50 µl. For each iteration of the screening process, 1 µl of this lysis-defective-enriched library was transformed into MC4100(λCm^r*Δ*(*SR*)). Forty individual transformants were inoculated into 5 ml of LBAmpCam and grown at 30°C in a rollerdrum until the cultures were in mid-logarithmic phase. Fifty µl of each culture was reserved and then the tube cultures were transferred to a rack, shaken in a

42°C water bath for 15 min, and shaken in a 37°C water bath for 65 min. Typically, ~ 60% of the induced tube cultures underwent visible lysis. Non-lysed cultures were treated with 50 μl $CHCl_3$, to test for endolysin accumulation. Typically, 90% of the non-lysing cultures lysed upon addition of $CHCl_3$, consistent with an absolute defect in the function of the Y holin. These presumptive *Y* mutants were grown to saturation from the reserved culture and used for plasmid preparation and DNA sequencing. Of a total of 120 transformants, 72 exhibited an absolute lysis defect that was complemented by $CHCl_3$. Of these, 64 had at least one mutation in the *Y* gene. Of these, 43 had a single mutation, 15 had two mutations, and 6 had more than two. Of the 72 mutations, two were -1 frameshifts, and the rest were single base changes, approximately evenly distributed between transversions and transitions. All of the mutants which had a single, non-silent mutation are listed on page 55-56. The mutant library is not saturated because only two of the 18 nonsense mutations accessible by a single base change were isolated.

Growth conditions, inductions and TCA precipitation

For lysis curves and protein expression experiments, overnight cultures were used to inoculate 25 ml cultures of LB supplemented with the appropriate antibiotic. The cells were grown at 30°C to A_{550} ~ 0.2 (for lysis curves) and A_{550} ~ 0.4 (for protein expression experiments), thermally induced by aeration at 42°C for 15 min, and then returned to 37°C. Plaque-forming titer of progeny phage was determined as described previously (68). For protein expression experiments, 1 ml samples were removed from the growing cultures and added to 100% TCA on ice. The samples were incubated for 30 min, then centrifuged at maximum speed in a tabletop microcentrifuge at 4°C. The

supernatant was removed and the pellet was resuspended in 1 ml cold acetone, vortexed, and spun at maximum speed in the tabletop microcentrifuge at 4°C. The acetone wash was repeated and the supernatant was removed carefully after the final spin. The pellet was air-dried overnight and resuspended in 2X reducing sample loading buffer for SDS-PAGE. Western blotting was done with an antibody directed against the C-terminus of Y (Genscript).

Antibodies

Antibodies against the peptide CEIAIQRRFKGRGKQ corresponding to the C-terminal sequence of Y (Fig. 13C) were prepared and affinity purified by Genscript. For purification, the antibody used was commercial α-His_6 from Amersham.

Site-specific photocrosslinking

A 25 ml culture of MC4100Δ*fhuA*(λ*Cam*Δ*SR*) carrying the desired plasmid was thermally induced and harvested at $A_{550} \sim 0.5$ or 45 min after induction. One ml samples of different *Y* alleles was collected by centrifugation, and the cell pellet was resuspended and washed twice in 1 ml PBS, pH 7.2, 2 mM EDTA, pH 8.0 and 1 mM PMSF. The cell pellet was then resuspended in 200µl PBS and pipetted into a sterile Petri dish (15 mm diameter per well, 24 wells per dish) and 2 µl of 100 mM 4-maleimidobenzophenone, or MBP (Sigma), prepared in dimethylformamide, or 1 µl of dimethylformamide, for the mock treatment, was added. The mixture was incubated at room temperature for 30 minutes in the dark, after which 2 µl 1 M DTT was added to quench unreacted MBP. The samples were placed on ice and irradiated with 366 nm light using a hand held lamp

for 30 minutes. Samples were mixed with 2X sample loading buffer and analyzed by SDS-PAGE, and Western blotting using Y-specific antibody.

Over-expression and purification of P2 Y

All alleles of *Y* used in this work are isogenic to Y_{Q93}, which encodes Y with the sequence LEHHHHHH added after the C-terminal Gln93 residue. Overexpression of *Y* alleles was carried out as previously described (75) except that the vector pET30b was used instead of pET11a and the host used for induction was Escherichia coli C43(DE3) (76). The construct pET30b Y was made and transformed into competent cells for over-expression. Once induced for 45-50 min, the cells were harvested and lysed via passage through a French pressure cell. The membranes were pelleted, and the membrane proteins were extracted with detergent, and then purified over a Talon® resin column.

P2 Y was purified according to the protocol used for over-expression and purification of S105 (42, 67, 75) with the following exceptions. Tris was used for all buffers, and n-dodecyl β-D-maltoside, or DDM, detergent was used in the membrane extraction and subsequent purification buffers at final concentrations of 1% (w/v) and 0.1% (w/v), respectively. Membrane extraction was carried out at 37 °C for 2 hours with gentle shaking. Protein detection was carried out with SDS-PAGE followed by Western blotting using the commercially available His6 antibody (Amersham). Protein samples for cysteine modification experiments were purified as previously described but with a saturating concentration of 2,2′-dithiodipyridine, DTDP, included in the extraction buffer (42).

Gel permeation chromatography and negative stain electron microscopy

Gel permeation chromatography was carried out using Superose-6 on an AKTA FPLC (Pharmacia). Calibration standards were purchased from Bio-Rad and used according to the manufacturer's directions. Negatively stained specimens were prepared as described (77), by floating carbon-coated mica onto sample droplets followed by staining with aqueous uranyl acetate (2% w/v), then carefully blotted dry. Electron micrographs were recorded in a JEOL 1200 EX TEM (MIC, Texas A&M University) at a calibrated magnification of 24,380×. Micrographs were digitized using the Leafscan 45 microdensitometer to a final pixel size of 4.1 Å pixel^{-1}.

Results

Y exhibits holin function

As a first step in characterizing the lytic function of P2 Y, I cloned the *Y* gene into the plasmid pS105 (31), replacing the S105 gene in the lambda lysis cassette (Fig. 13B). Inserts in this plasmid are expressed in the proper temporal and transcriptional context by inducing a lysis-defective prophage, (λ*CamΔSR*), resulting in trans-activation of the plasmid-borne lambda late promoter, pR', by the phage-encoded late gene activator, Q. The resulting hybrid lysis cassette, *YRRzRz1*, in plasmid pKT1-Y, was found to cause abrupt lysis at ~ 30 min after induction (Fig. 14A). Since *R* and *RzRz1* are known to encode a soluble endolysin and a two-component spanin, respectively, this result strongly suggests Y is a functional holin. However, inappropriate expression of a membrane protein or membrane-active toxins can suppress a holin lysis defect, presumably by non-specific damage to the membrane allowing escape of sufficient

endolysin to destroy the cell wall (70). To eliminate this possibility, cells induced for the *Y-RRzRz1* chimera were tested for sensitivity to addition of energy poisons; premature triggering was unambiguously demonstrated (Fig. 14A), as previously shown for the prototype holins of classes I, II and III (1). Moreover, even in the absence of endolysin and spanin function, *Y* exhibits inducible lethality, and, like S105 and $S^{21}68$, addition of energy poisons triggers lethal function (Fig. 14B). Thus Y exhibits all the physiological characteristics unique to the well-characterized holins.

Single-amino-acid changes in Y have dramatic and unpredictable effects on lysis timing

For the prototype holins of class I and II, a striking common feature was that relatively conservative missense changes, especially in the TMD sequences, led to dramatically altered lysis timing compared to the wt, either advancing or retarding the triggering time (64, 68, 69, 78, 79). To see if the same were true for *Y*, initially site-directed mutagenesis was used to probe key positions in the TMDs for timing phenotypes. In S105, alleles with missense changes at a central residue in TMD2, Ala52, block plaque formation, for opposite reasons: $S105_{A52V}$ is completely non-lytic, whereas $S105_{A52G}$ causes lysis before the first virion is assembled (68, 69). An alignment of Y and S105 suggested that a residue analogous to Ala52 in S105 would

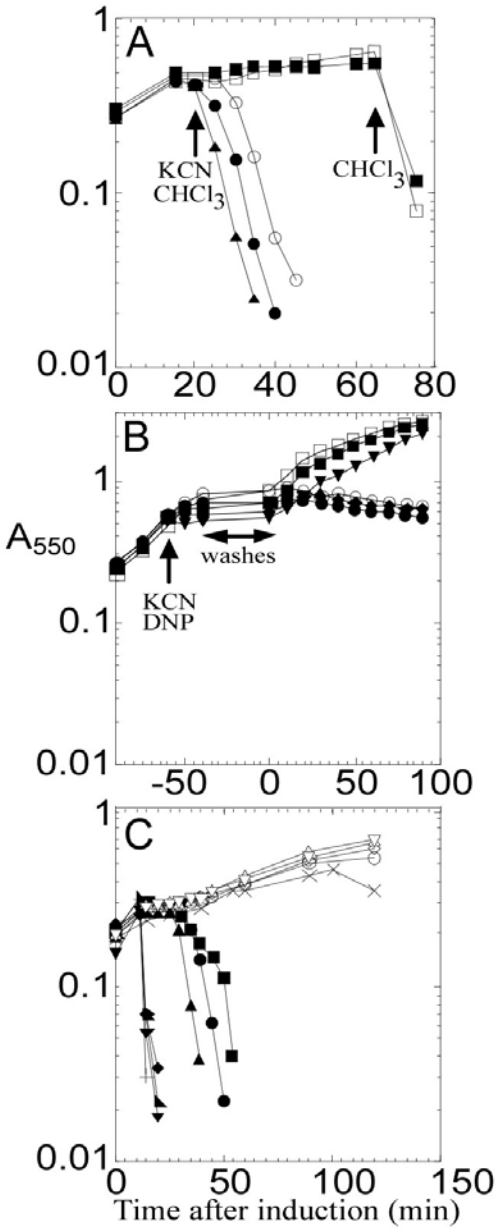

Fig. 14. P2 Y holin properties. (A) P2 Y has lytic properties of a holin. Inductions: pKT1-Y_{am}, with $CHCl_3$ added at 65 min (□); pKT1-Y_{am}, with cyanide added at 20 min and $CHCl_3$ at 65 min (■); pKT1-Y, with $CHCl_3$ added at 20 min (▲); pKT1-Y, with cyanide added at 20 min (●); pY (○). (B) Inducible lethality of P2 Y. Inductions: pKT2-Y (○); pKT2-Y, with DNP added at -60 min (●); pKT2-Y, with KCN added at -60 min (♦); pKT2-Y_{am} (□); pKT2-Y_{am}, with DNP added at -60 min (■); pKT2-Y_{am}, with KCN added at -60 min (▼). (C) Lysis phenotypes of P2 *Y* mutants. Inductions: pKT1-Y (▲); pKT1-Y_{L9Q} (×); pKT1-Y_{L17P} (○); pKT1-Y_{K22M} (Δ); pKT1-Y_{A25V} (+); pKT1-Y_{S46C} (●); pKT1-Y_{A49T} (◊); pKT1-Y_{L53M} (♦); pKT1-Y_{C65G} (■); pKT1-Y_{M70I} (▼); pKT1-Y_{A74T} (◣); pKT1-Y_{G75S} (▽).

be Gly50 in P2 Y (Fig. 13C; asterisks). Indeed, the G50V mutation in Y eliminates lytic function without affecting accumulation, similar to A52V in S105, but, unlike position 52 in S105, substituting small residues (Ala, Ser) in this position causes loss of lytic function (Fig. 14C; Table 2). To further address the mutational sensitivity of the P2 holin timing function, a library of point mutations was generated by error-prone PCR and selected for either defective or early lysis. A summary of the timing phenotypes of all the site-directed mutants and mutants obtained from the PCR library is shown in Table 2 and Fig. 13C, and representative lysis phenotypes are illustrated in Fig. 14C. As in lambda S105, missense changes in all three TMDs were found to abolish lysis, including some relatively conservative mutations, like L41F, V45F, S46A, and G75S. Moreover, a number of missense changes led to early lysis phenotypes, including A25V, L53M, M70I, and A74T (Fig. 14C; Table 2). These timing phenotypes could not be explained by profound changes in hydrophobic character of the substituted residues; both increases and decreases in hydrophobicity are associated with both early lysis and lysis-defective mutations (Table 2). Nor can the phenotypes be ascribed to differential accumulation of Y protein; all the alleles supported normal accumulation of Y, as determined by Western blot analysis (Fig. 15). Thus, although the mutant hunt was not saturated, it was

Table 2. Mutant alleles of *Y*. *Y* mutations isolated in this work are listed.

Position[1]	Residue change[2]	Codon change[3]	Lysis[4] (min)	ΔHydropathicity[5]	#[6]	Origin[7]
Wild-type			30			
insertion at 2	+Lys	+AAA	-			S
insertion at 2	+LysLys	+AAAAAG	-			S
2	Thr→Arg	A***C***A→A***G***A	-	-0.7→-4.5	1	
9	Leu→Gln	C***T***G→C***A***G	(90)	3.8→-3.5	1	
10	Ser→Pro	***T***CG→***C***CG	-	-0.8→-1.6	1	
15	Gly→Glu	G***G***G→G***A***G	-	-0.4→-3.5	1	
17	Leu→Pro	C***T***G→C***C***G	-	3.8→-1.6	1	
22	Lys→Met	A***A***G→A***T***G	-	**-3.9→1.9**	1	
25	Ala→Val	G***C***C→G***T***C	12	1.8→4.2	1	
41	Leu→Phe	***C***TC→***T***TC	-	3.8→2.8	3	
42	Gly→Cys	***G***GT→***T***GT	-	-0.4→2.5	1	
43	Gly→Cys	***G***GT→***T***GT	-	-0.4→2.5	3	
43	Gly→Asp	G***G***T→G***A***T	-	-0.4→-3.5	2	
45	Val→Asp	G***T***C→G***A***C	-	4.2→-3.5	2	
45	Val→Phe	***G***TC→***T***TC	-	4.2→2.8	2	
46	Ser→Ala	***T***CG→***G***CG	-	-0.8→1.8	1	
46	Ser→Cys	T***CG***→T***GC***	30	-0.8→2.5	1	S
49	Ala→Asp	G***C***C→G***A***C	-	1.8→-3.5	1	
49	Ala→Thr	***G***CC→***A***CC	-	1.8→-0.7	1	S
50	Gly→Ala	G***G***T→G***C***T	-	**-0.4→1.8**		S
50	Gly→Cys	***G***GT→***T***GT	-	**-0.4→2.5**		S
50	Gly→Asp	G***G***T→G***A***T	-	-0.4→-3.5	1	
50	Gly→Ser	***G***GT→***A***GT	-	-0.4→-0.8		S
50	Gly→Val	G***G***T→G***T***T	-	**-0.4→4.2**		S
53	Leu→Pro	C***T***G→C***C***G	-	3.8→-1.6	1	
53	Leu→Met	***C***TG→***A***TG	12	3.8→1.9	1	
54*	Val→fs	***G***TG→-TG	-		1	
(55)	Gln→stop	***C***AG→***T***AG	-	-3.5→NA	1	
61	Leu→Pro	C***T***G→C***C***G	-	3.8→-1.6	1	
62	Pro→Ser	***C***CA→***T***CA	-	-1.6→-0.8	1	
65	Cys→Gly	***T***GC→***G***GC	15	2.5→-0.4		S
(65)	Cys→stop	TG***C***→TG***A***	-	2.5→NA	1	
66	Gly→Asp	G***G***C→G***A***C	-	-0.4→-3.5	1	
66	Gly→Val	G***G***C→G***T***C	-	**-0.4→4.2**	1	
69	Ser→Phe	T***C***C→T***T***C	-	**-0.8→2.8**	2	
70	Met→lle	AT***G***→AT***T***	12	1.9→4.5	1	
74	Ala→Thr	***G***CC→***A***CC	12	1.8→-0.7	1	
75	Gly→Ser	***G***GT→***A***GT	-	-0.4→-0.8	2	
76*	Tyr→fs	***T***AT→-AT	-		2	

Table 2. Continued

Position[1]	Residue change[2]	Codon change[3]	Lysis[4] (min)	ΔHydropathicity[5]	#[6]	Origin[7]
87	Phe→Ser	T***T***T→T***C***T	-	2.8→-0.8	1	

[1]Aminoacid position (see Fig. 13C). All changes are missense changes except the two (Lys and LysLys) insertions at position 2, two frameshift mutants (asterisk), two nonsense mutants (parentheses). Heavy outlines enclose residues in TMDs 1, 2 and 3, from the top.
[2]Change in amino acid. fs=frameshift.
[3]Change in nucleotide sequence for mutation. Altered bases are in bold and italics.
[4]Triggering time in min after induction. - = non-lytic.
[5]Change in hydropathic character (80). Bold, underlined entries indicate increased hydrophobicity for lysis-defective alleles or decreased hyrophobicity for early-lysis alleles.
[6]Number of times isolated for error-prone PCR mutants.
[7]Origin of mutation. S = site-directed mutagenesis. All others were obtained by random error-prone PCR.

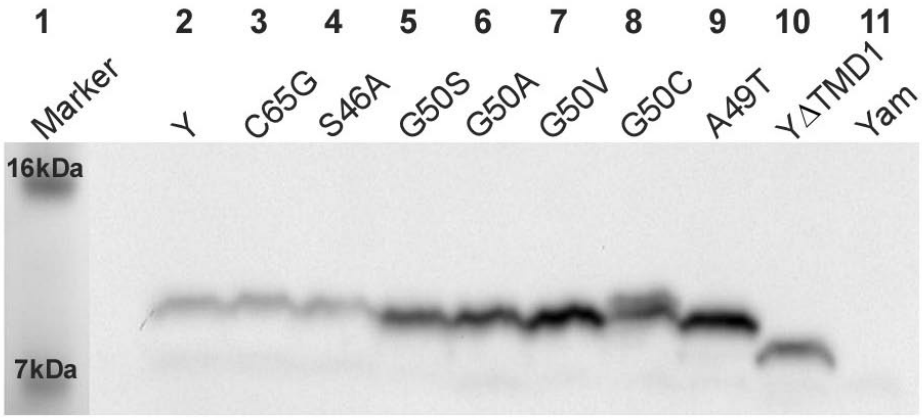

Fig. 15. Y mutants accumulate normally. Lane 1: molecular weight marker; lane 2: pKT1-Y; lane 3: pKT1-Y_{C65G}; lane 4: pKT1-Y_{S46A}; lane 5: pKT1-Y_{G50S}; lane 6: pKT1-Y_{G50A}; lane 7: pKT1-Y_{G50V}; lane 8: pKT1-Y_{G50C}; lane 9: pKT1-Y_{A49T}; lane 10: pKT1-$Y_{\Delta TMD1}$; lane 11: pKT1-Yam.

sufficient to show that, like S105, P2 Y appears to have the same extreme and unpredictable sensitivity to missense changes in the TMDs. Taken with the detailed studies on lambda S105 and $S^{21}68$, I conclude that the ability of conservative changes in TMD sequences to cause profound alteration of the triggering time is a common feature of holin proteins.

Requirement for C-terminal cytoplasmic domain

The most conserved feature of class I and II holins is that the most distal TMD is followed by a short, C-terminal cytoplasmic tail enriched in basic residues (5). It has been reported that the cytoplasmic C-terminal domain (CTD) of lambda S105 is not necessary for host cell lysis, but instead plays a regulatory role in lysis timing (22, 69, 81). Site-directed mutagenesis experiments with S105 indicated that lytic function was retained as long as one positive charge residue remained at the C-terminus, and that the timing of lysis was sensitive to the number of basic residues encoded in the C-terminal tail, with more positive charge associated with delayed triggering times (22). I tested whether the highly charged C-terminal tail of Y served a similar function by constructing $Y_{\Delta CTD}$, in which the last 8 codons, Arg86 to Gln93, were deleted (Fig. 13C), so that the $Y_{\Delta CTD}$ product has only a single basic residue at the C-terminus, instead of five in the parental protein. Unlike the $S105_{\Delta CTD}$, which terminates with the single positively-charged residue Arg93 and has an extreme early lysis phenotype (22), the $Y_{\Delta CTD}$ is an absolute lysis-defective (Fig. 16A). Since the antibody against Y recognizes

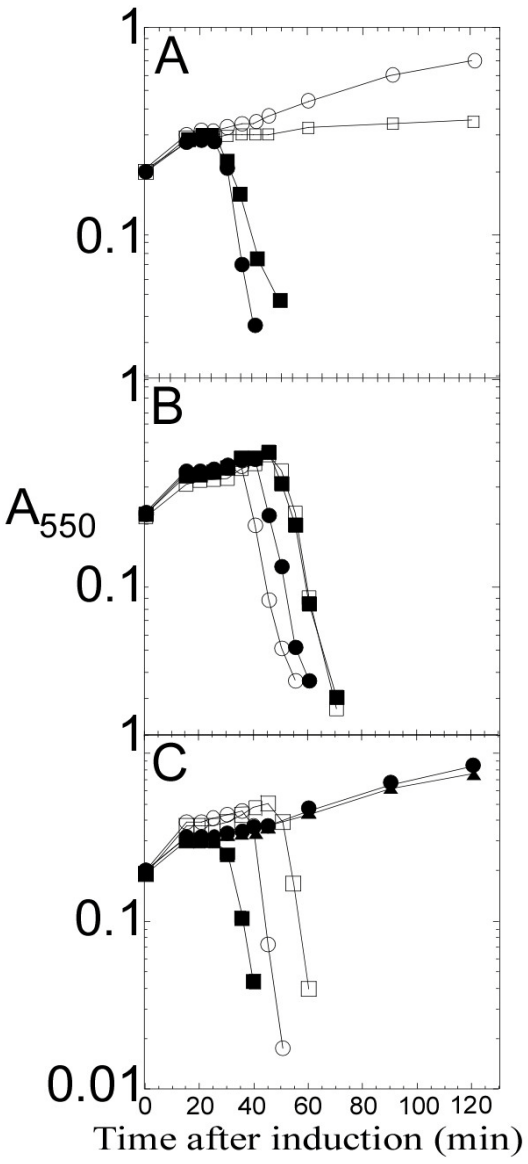

Fig. 16. Functions of Y CTD, LysA, and Y TMD1. (A) Deletion of the Y CTD inactivates holin function. Inductions *in trans* to λ*Cam*Δ*SR*: no plasmid, (○); pKT1-Y (●); pKT1-$Y_{\Delta C\text{-term}}$ (□); λ-*Y* lysogen harboring pKT1-$Y_{\Delta C\text{-term}}$ (■).(B) LysA delays Y-mediated lysis. Inductions *in trans* to λ-*Y*: no plasmid (○); plysA (●). Inductions *in trans* to λ-S105: no plasmid (□); plysA (■). (C) $Y_{\Delta TMD1}$ delays Y-mediated lysis. Inductions *in trans* to λ*Cam*Δ*SR*: no plasmid (●); pKT1-$Y_{\Delta TMD1}$ (▲); pKT1-Y (■). Inductions *in trans* to λ-*Y*: no plasmid (○); pKT1-$Y_{\Delta TMD1,}$ (□).

a C-terminal epitope, it was not possible to rule out whether $Y_{\Delta CTD}$ does not accumulate. However, *in trans* to the parental *Y* allele, $Y_{\Delta CTD}$ is as functional as the wt gene (Fig. 16A). The simplest interpretation is that the C-terminal domain must be present in some n-mer state for Y to function, but that not all protomers of the n-mer have to have the domain. In any case, I conclude that the C-terminal cytoplasmic domain of Y is essential for Y function in vivo, marking a distinct difference from the canonical class I holin, S105.

Sizing the Y hole

Early in the molecular characterization of the canonical holin class I holin S105, it was shown that the holes were lethal and non-specific for endolysins. An indication of the size of these holes was obtained using fusion constructs in which the lambda R endolysin gene was fused to the full-length *lacZ* gene, thus creating a chimera of more than 1100 codons encoding a hybrid R-βgal enzyme (66). In addition, the chimeric gene was modified to encode an N-terminal cmyc-tag, to allow monitoring of the production and stability of the hybrid protein. When this fusion gene was substituted for the parental *R* gene, the lytic and plating capacity of the phage were unaffected. Moreover, it was demonstrated that R-βgal was not only produced as an uncleaved, stable 120 kDa polypeptide but that the hybrid protein was fully active as an endolysin and also as a homotetrameric β-galactosidase of ~ 0.5MDa in mass. To ask whether Y also formed holes of such capacity, plasmids encoding the *cmycR-lacZ* fusion were induced in both *S+R-* and *Y+R-* contexts (Fig. 17A). The results show unambiguously that the Y holin

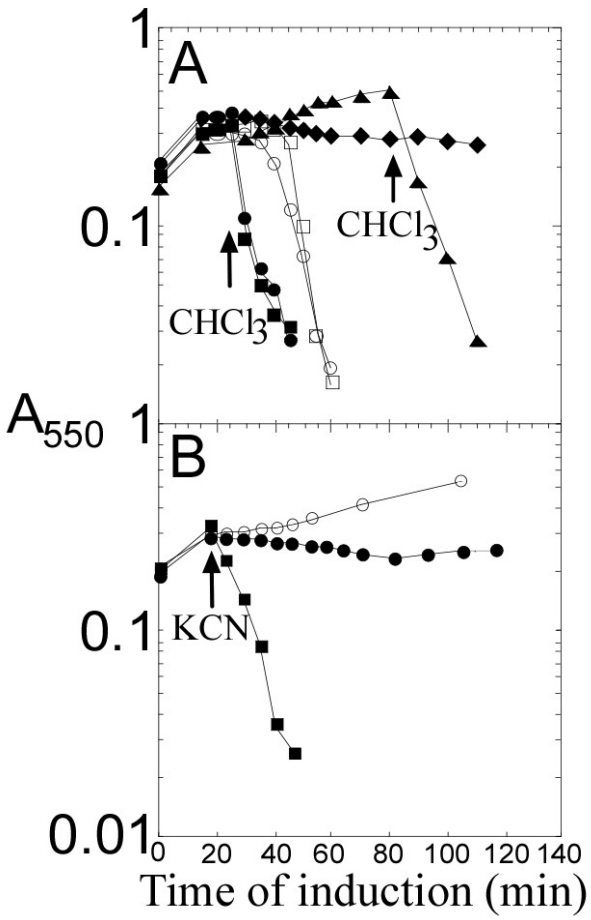

Fig. 17. Size of the Y hole. (A) Sizing the Y hole. Inductions *in trans* to λ*KanΔ(SR)*; pKT2-Y and pc-mycR::lacZ (○); pKT2-Y and pc-mycR::lacZ, with $CHCl_3$ added at 25 min (●); pKT2-S105 and pc-mycR::lacZ (□); pKT2-S105 and pc-mycR::lacZ, with $CHCl_3$ added at 25 min (■); pKT2-Y, with $CHCl_3$ added at 80 min (♦); pc-mycR::lacZ (▲). (B) Premature Y holes are permissive for R endolysin but not the endolysin-β-Gal hybrid. pKT2-Y and pc-mycR, with KCN added at 15 min(■); pKT2-Y and pc-mycR::lacZ, (●); pKT2-Yam(○).

is permissive for the R-βgal hybrid protein, suggesting that the non-specificity of the Y hole is, like S105, due to a very large pore size in the membrane. However, when the Y holin was prematurely triggered by addition of cyanide, lysis was observed with the wt R endolysin but not the R-βgal fusion (Fig. 17B). This indicates the holes formed by premature triggering are much smaller than those formed at the spontaneous triggering time, consistent with similar observations with the lambda holin (66). Furthermore, cryo-EM was used by Dr. Jill Dewey to visualize cells expressing P2 Y, in an effort to determine whether large lesions are formed. MC4100 Δ*tonA* (*Cam*Δ*SR*) competent cells transformed with the plasmid pYRamRzamRz1am were induced for expression of Y. The cells were collected, concentrated, and applied to prepared EM grids, then plunge-frozen into liquid ethane to fix the cells in a thin layer of vitreous ice. The cells were visualized by cryo-EM, and 65 cells were imaged. Fig. 18 shows a cell induced for expression of Y. Of 65 cells that were imaged, 30, (46%), had visible inner membrane lesions. Hole size ranged from ~ 114 nm to ~ 1.3 μm, and the major size class was 200 – 400 nm (Fig. 19A). In addition, the majority of cells with holes had a single visible hole, but there were instances where cells had multiple visible holes (Fig. 19B), and the geometry can be used to calculate the predicted actual number of holes per cell. Y holes were not localized to a specific region of the membrane, nor was there a correlation between hole size and localization within the membrane (Fig. 19C).

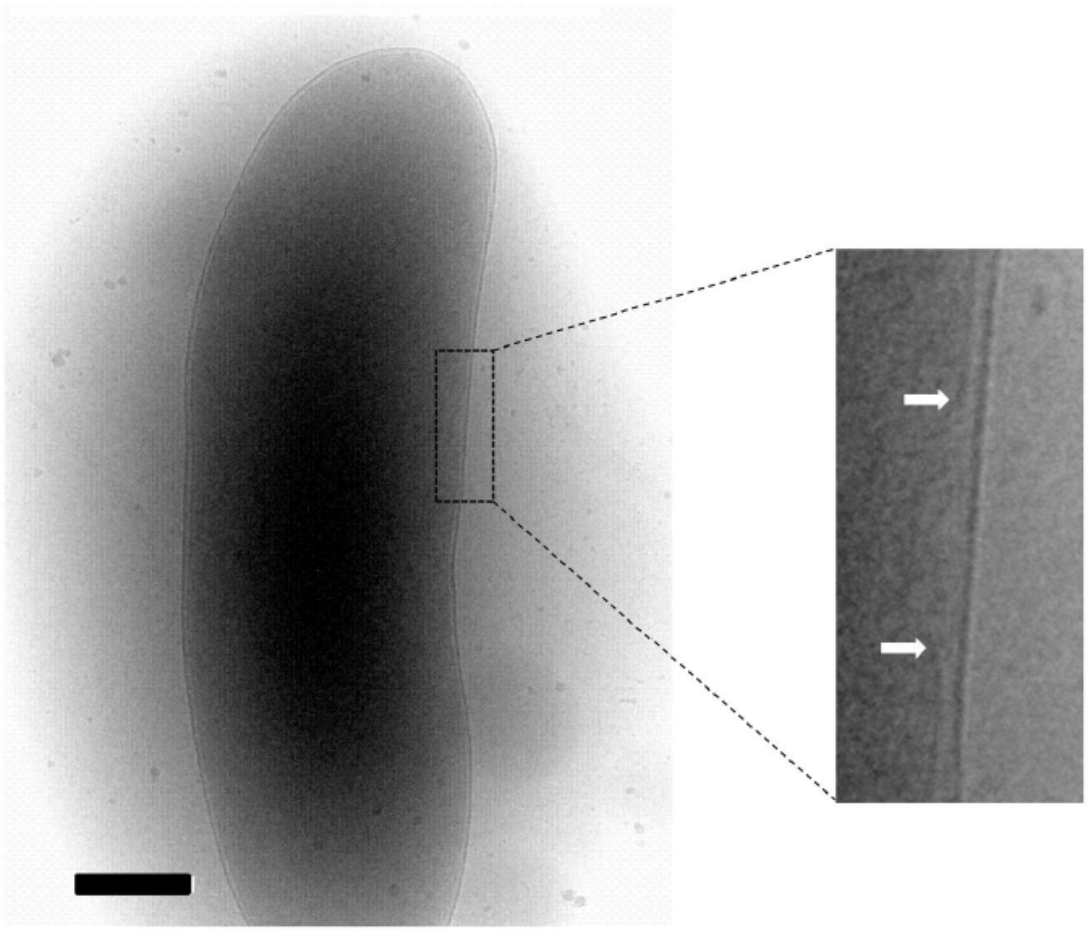

Fig. 18. Cryo-EM of *E coli* cell expressing P2 *Y*. The enlarged region shows a disruption in the inner membrane which is the holin-mediated lesion. White arrows indicate the discontinuity in the cytoplasmic membrane. The scale bar corresponds to 0.5 μm.

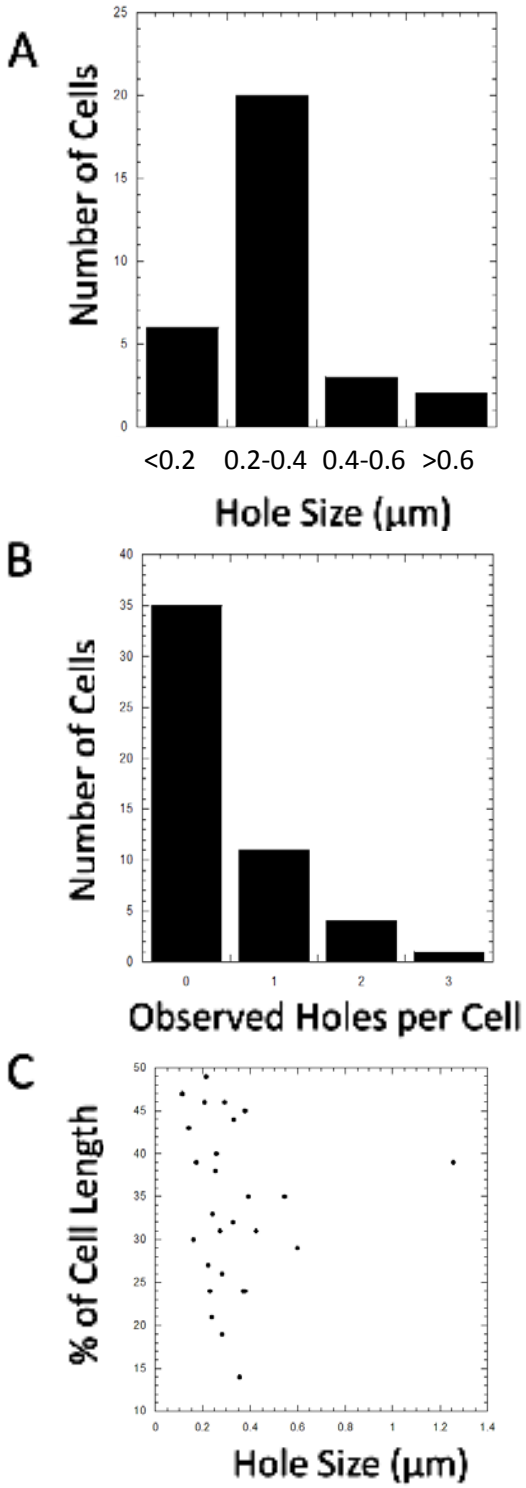

Fig. 19. Statistics for cells expressing *Y*. Including hole size (A), number of holes per cell (B), and hole localization with respect to hole size (C).

Lysis deficiency of Y mutants is associated with an oligomerization defect

If membranes are treated with the membrane-permeable, amine-specific cross-linker DSP, S105 can be trapped in covalent oligomeric states, up to at least n=5; by this criterion, almost all lysis-defective *S* missense alleles are defective in oligomerization (71). To distinguish the molecular differences between functional and nonfunctional alleles of *Y*, I attempted the same analysis, but treatment with DSP did not generate oligomeric species detectable by immunoblot (not shown). However, the presence of the single Cys residue at position 65 (Fig. 13C) made it possible to use the membrane-permeable, photoactive heterobifunctional cross-linker MBP (82). When whole cell samples that were treated with MBP and irradiated for increasing time were analyzed by immunoblot, I observed that most of the Y protein becomes cross-linked in species that correspond to a dimer and also to high order aggregates that are unable to enter the resolving gel, as is shown for representative alleles in Fig. 20A. In contrast, when the same treatment was performed with cells expressing lysis-defective missense *Y* mutants, irradiation had reduced or no effect on the amount of the monomeric species, the dimer species was more prevalent and the high-order aggregate was severely reduced or absent (Fig. 20B). The mutant Leu9Gln exhibits the weakest molecular phenotype, showing only a reduced amount of the higher order cross-linked species, which is consistent with the leaky lysis defect, with lysis beginning at about 90 min (Fig. 14C). The simplest interpretation is that Y undergoes dimerization and then higher order oligomerization, and that the mutants are blocked at different stages of oligomerization. Thus, for lethal function Y shares with S105 the requirement for high-order oligomerization detectable

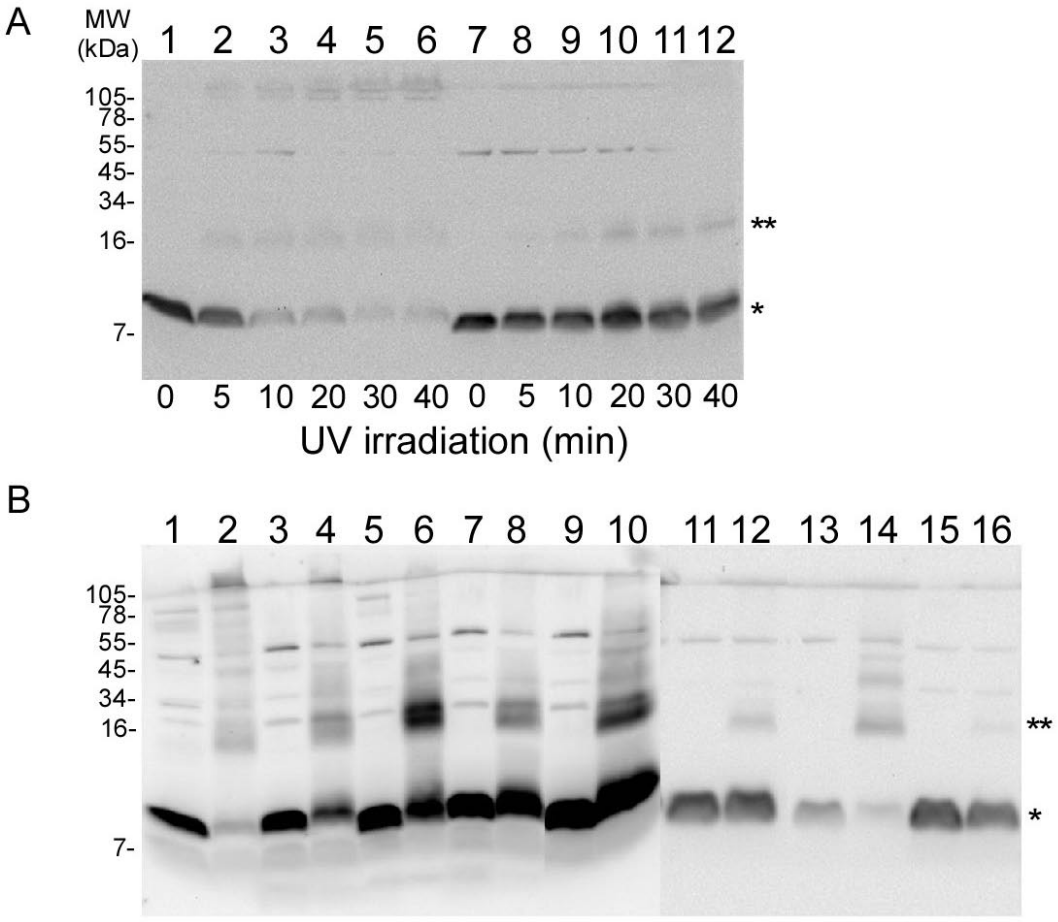

Fig. 20. Oligomeric state of different Y alleles in Western blot. (A) Y, but not defective Y proteins, can be cross-linked into high order oligomers. Photoactivated MBP crosslinking of whole-cell samples. Cells carrying either pKT2-Y (lanes 1-6) or pKT1-Y_{L17P} (lanes 7-12) were induced, labeled with MBP and UV activated for the indicated lengths of time, in minutes, as described in Materials and Methods. (B**)** Lysis-defective Y proteins lose the ability to be cross-linked into high MW species. Cultures carrying the indicated plasmids were induced, labeled with MBP, and UV-irradiated for 30 min (even-numbered lanes) or not (odd-numbered lanes). Lanes: 1, 2: Y; lane 3, 4: L9Q; lane 5, 6: L17P; lane 7, 8: A49T; lane 9, 10: G75S; lanes: 11, 12: K22M; lanes: 13, 14: L53P; lanes: 15, 16: G66D; monomer (asterisk), dimer (2 asterisk).

by cross-linking (71).

LysA exhibits Y-specific antiholin character

For the prototype holins of classes I, II and III, there is a specific antiholin. In the first two cases, lambda *S* gene and the S^{21} gene of lambdoid phage 21 encode both the holin and antiholin, by virtue of alternate translational start codons (e.g., holin S105 and antiholin S107 from translational initiations at codons 3 and 1, respectively (Fig. 13C). Such "dual start" motifs are a very common feature in genes identified as encoding class I and class II holins (5, 83). In contrast, in phage T4, the holin T is inhibited by a specific antiholin encoded by gene *rI*. Moreover, genetic evidence also suggests that in phage P1, the holin LydA is inhibited by LydB, encoded by the adjacent gene (84-86). Since *Y* has only a single start codon, I wondered whether one of the genes in the P2 lysis cassette encoded an antiholin. The P2 lysis cassette is syntenic with the lambdoid lysis cassette, with holin, endolysin and spanin genes arrayed in a cluster, except for a single gene, *lysA*, interposed between the endolysin gene *K* and the spanin genes *lysBC* (Fig. 13B). LysA is predicted to be an integral membrane protein with four TMDs (Fig. 13A, D). Indeed, Ziermann *et al* (70) constructed a P2*lysA*am and reported that cells infected with this construct underwent lysis ~ 10 min earlier than cells infected with the isogenic wt phage. To test the hypothesis that LysA could serve as a specific antiholin for P2 Y, I constructed a plasmid that would express *lysA* from the λ late promoter *in trans* either to induced λ-*Y* or the isogenic parental λ. Induction of λ-*Y* lysogens transformed with the *lysA* construct resulted in lysis ~ 10 min later than the same lysogen transformed with the isogenic vector (Fig. 16B), with a ~ 50% increase in

the yield of plaque-forming units (1.4×10^9 pfu/ml versus 9.3×10^8 pfu/ml) in the final lysate, without affecting accumulation of Y (not shown). These results indicate that LysA causes a delay in lysis by retarding the lethal action of Y, thus allowing continued virion morphogenesis. In contrast, the *lysA* plasmid had no effect on lysis of the induced parental λ lysogen, demonstrating that LysA-mediated inhibition of lysis is Y-specific and supporting the hypothesis that LysA is the P2 antiholin.

TMD1 is essential for holin activity and can be converted into an inhibitory domain

The S107 antiholin exerts its inhibitory function as a consequence of the ectopic localization of TMD1, which is prevented from entering the bilayer by virtue of the positively charged Lys residue at position two (Fig. 13C), which is missing in S107. Since *Y* does not have a dual start motif and appears to have a cognate antiholin encoded by a distinct gene, *lysA*, I wondered if there was something intrinsic about TMD1 of Y that prevents P2 from using the more genomically parsimonious dual start motif for encoding an antiholin. First I tested whether TMD1 was essential for Y function by constructing an in-frame deletion of codons Ser7 to Lys22 (Fig. 13A, C). The $Y_{\Delta TMD1}$ allele proved to be non-lytic in the context of the phage induction (Fig. 16C), although the deletion product accumulated normally (Fig. 15). Thus, as in S105, TMD1 is essential for lethal function of Y. Next I tested whether *Y* could be converted to dual start regulation, by constructing *Y* alleles with one (Y_{+K}) or two (Y_{+2K}) extra Lys residues at the N-terminus, making these the equivalent of the antiholin products of lambda *S* (S107, with one Lys residue at the N-terminus) and the homologous P22 gene *13* (gp13_{108}), with two Lys residues at the N-terminus (87). Both modified alleles exhibited

an absolute lysis defect (Fig. 21A) although the gene products accumulated normally in the membrane (not shown). Importantly, both modified *Y* alleles exhibited a specific antiholin phenotype, blocking triggering of Y but not S105, when expressed *in trans* (Fig. 21C). The simplest interpretation of these results is that, as for S105, an extra positively-charged residue at the N-terminus of Y prevents entry of TMD1 into the membrane, creating a negative-dominant mutant protein. Importantly, however, neither modified *Y* allele could be triggered by depolarization of the membrane with cyanide, either to release endolysin or to form lethal lesions in the membrane (Fig. 21B), indicating that a dual start motif may be necessary but not sufficient to generate a functional antiholin.

Purified Y holin oligomerizes as a helical structure in detergent

Cryo-EM studies showed that in the non-ionic detergent DDM, purified S105 forms ring structures, estimated to consist of 18-20 tetrameric protomers (67). These rings are not representative of S105 structures in vivo, in that their internal diameter of ~ 8 nm is too small to accommodate the proteins known to escape through S105 holes, and in detergent, they are found only in dimers and stacks of dimers. Nevertheless, the rings reflect the oligomerizing function of the holin in vivo, because mutant proteins that do not have lethal function in vivo do not form the rings in vitro. To facilitate biochemical analysis of Y, an allele encoding a C-terminal oligohistidine- tagged

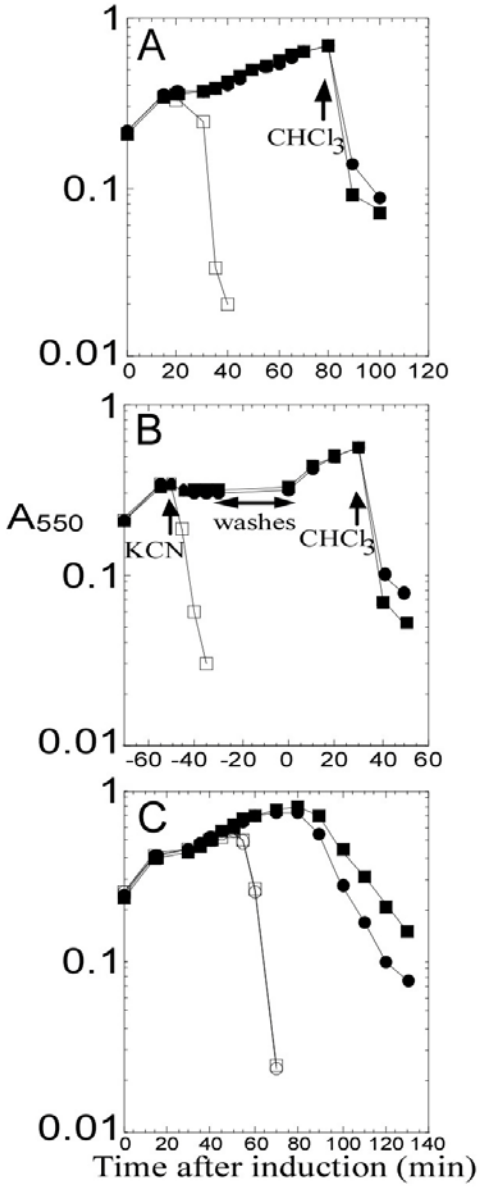

Fig. 21. The effect of Y_{+K} and Y_{+2K}. (A) Y_{+K} and Y_{+2K} are absolute defective holin. Inductions: pKT1-Y (□); pKT1-Y_{+K}, with $CHCl_3$ added at 80 min (●); pKT1-Y_{+2K}, with $CHCl_3$ added at 80 min (■). (B) Y_{+K} and Y_{+2K} have no lethality. Inductions; pKT1-Y (□); pKT1-Y_{+K}, with KCN added at -50 min and $CHCl_3$ at 30 min (●); pKT1-Y_{+2K}, with KCN added at -50 min and $CHCl_3$ at 30 min (■). (C) Y_{+K} and Y_{+2K} are negative dominant and Y-specific. Inductions: Lambda S105 lysogen harboring pKT1-Y_{+K} (○) or pKT1-Y_{+2K} (□). Lambda Y lysogen harboring pKT1-Y_{+K} (●) or pKT1-Y_{+2K} (■)

variant, Y_{H6}, was constructed (Fig. 13C), and shown to have normal lysis kinetics (Table 2). As a control, an A49T missense derivative of this allele was constructed and shown to faithfully reproduce the absolute lysis-defective phenotype observed in the parental allele. Both Y_{H6} and $Y_{H6,A49T}$ were then over-expressed in a pET vector system (See Materials and Methods), accumulating to 1300 molecules per cell at the time of lethal triggering (also consistent with Y molecules lining the Y holes) and ~ 0.25g/liter of induced culture. With the imprecision of quantitative Western blotting in mind, these calculations cannot be used to rule out the possibility of moderate "honey-combing" of the macroscopic holes. Despite attempts with multiple detergents, only a small fraction of either protein could be extracted from the membrane (Fig. 22A). After extraction in the non-ionic detergent DDM, both proteins were purified by IMAC and gel permeation chromatography, as previously described (42). Images of negatively stained samples taken from the peak fractions revealed that DDM-solubilized Y_{H6} formed helical rods of ~ 100nm average length and ~ 18nm width (Fig. 22B). Similar structures were observed for the non-functional variant, $Y_{H6,A49T}$, except that the rods were much longer (Fig. 22C).

Discussion

Prior to this work, only three holins had been studied in detail: S105, $S^{21}68$ and T, the prototypes of topology classes I, II and III. Of these, only the first two have been purified biochemically (31, 67), and only one, S105 is a canonical, large-hole forming holin that allows release of a prefolded soluble endolysin. Class I holins are the most

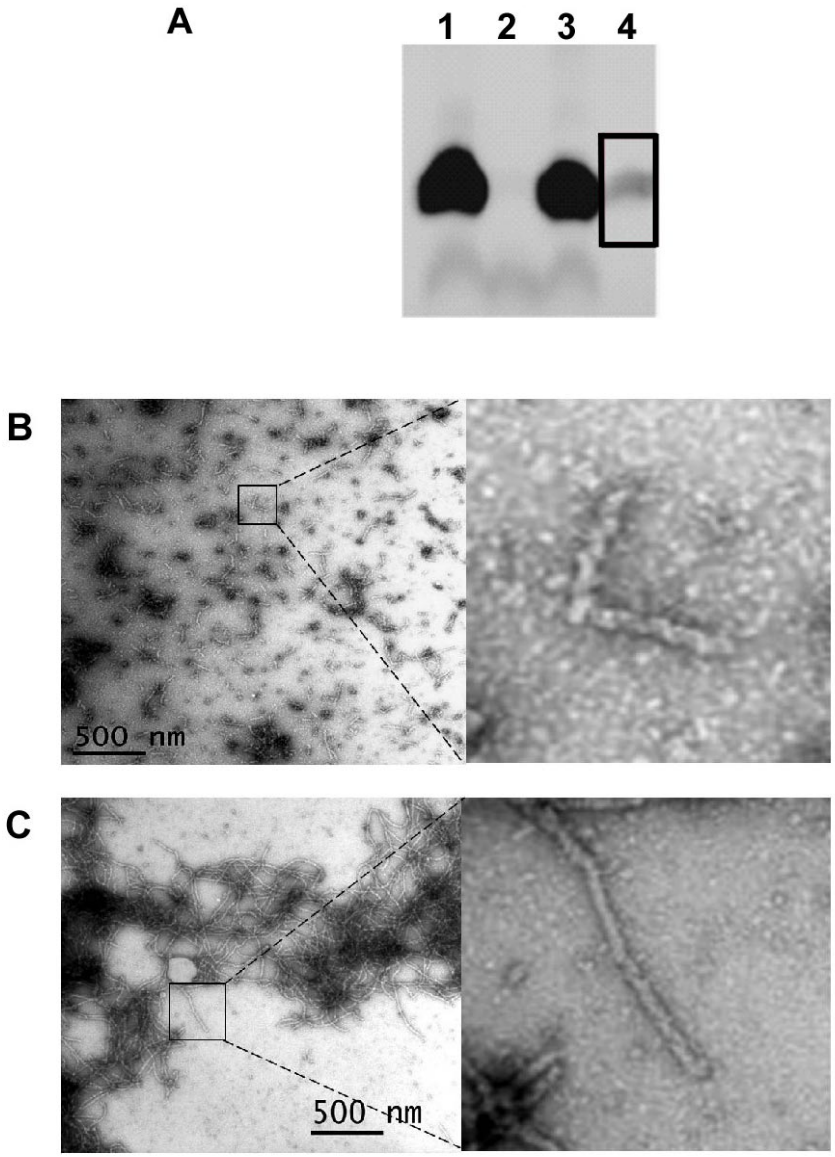

Fig. 22. Purification and negative-stain EM of Y. (A) Purification of P2 Y in DDM. Samples from successive steps in the purification of Y-His6 were analyzed by Western blot using anti-His6 antibody. Lanes: 1, total membrane; lane 2: detergent extract; lane 3, pellet from detergent extraction; lane 4, elution from Talon column.
(B, C) Negative-stain electron microscopy of DDM-solubilized Y (panel B) and the non-lytic mutant protein, Y_{A49T} (panel C). The helical structures average ~ 100nm in length and ~ 18 nm width for Y, but the mutant forms longer, branched structures. The scale bar corresponds to 500 nm.

numerous in terms of holins annotated in the database (5, 12), and thus, it was important to see if the many striking features of the well-studied lambda S105 was representative of this important topological class. The predicted membrane topology for Y is three TMDs with N-out, C-in, as demonstrated for S105. The results presented here show that some of these features are common to both S105 and Y and thus may be considered representative of type I holins:

Potential-sensitive lethality

It was previously shown that, the S105 holin exerted potential-sensitive lethality when induced independent of other proteins required for lysis, namely the endolysin and the spanin. The lethality is manifested as a sudden halt in culture growth after induction and associated with a loss of viability, i.e., the triggering phenomenon. Y fully shares this characteristic (Fig. 14B). The potential sensitivity is reflected in the fact that collapse of membrane energization with energy poisons causes premature triggering; Y also exhibits this feature (Fig. 14A). Since accumulation of any membrane protein can lead to lethality and membrane dysfunction, the susceptibility of S105 and Y to premature lethal triggering by the addition of energy poisons can now be regarded as the definitive feature of holins, and especially the common class I holins.

Sensitivity of the timing function to conservative amino acid changes in the TMDs

Perhaps the most striking aspect of S105 is the extreme sensitivity of the triggering time even to conservative missense substitutions, especially in the TMDs (68). For example, in S105, changes in a central residue in TMD2, Ala52, to Gly or Val both

result in the loss of plaque-forming ability, but for opposite reasons (69). The Ala52Val protein accumulates in the membrane but never causes triggering, and thus plaque-formation is abolished because lysis does not occur. By contrast, the Ala52Gly protein causes lysis at 20 min, at about the time that the first progeny virion appears in the cytoplasm; plaque-formation is thus squelched because lysis is too early. Our initial attempts to see if Y exhibited similar phenotypic traits led us to do site-directed mutagenesis on the TMD2 residue predicted to be analogous to Ala52 (Fig. 13C). Although I was correct in judging Gly50 to be an residue important for lysis function, it turned out that all changes there resulted in an absolute lysis-defective phenotype, including Gly50Ala and Gly50Val, without affecting accumulation of the holin (Fig. 13C; Table 2). However, when I conducted random mutagenesis of *Y*, I was able to isolate both lysis-defective and early lysis alleles with rather conservative missense changes. For example, Leu41Phe, Val45Phe, Gly66Val and Gly75Ser inactivate Y, whereas Ala25Val, Leu53Met, Cys65Gly, Ala75Thr and Met79Ile cause extremely early lysis. Importantly, these changes in lysis timing are not associated with changes in protein accumulation or membrane localization for either S105 or Y (Fig. 15). It should be noted that early lysis phenotypes were also found for some missense and suppressed amber mutations in the TMDs of another putative class I holin, gp*XXXV* of the tectivirus PRD1, although no protein accumulation studies were done (79). Based on the mutagenesis studies on the *S* gene, as well as studies with a GFP-tagged S105 holin, a model has been proposed in which the triggering time of S105 depends on the kinetics of formation of large, two-dimensional arrays, or "rafts" of holin proteins. These rafts are

proposed to exclude lipids as a result of intimate intermolecular helical packing interactions between the TMDs of the holin proteins. Conceptually, the exclusion of lipids from within the rafts would allow hole formation to occur by re-orientation of the holin molecules such that a relatively hydrophilic surface of one or more TMDs faces the lumen of the hole. In this scenario, the radical phenotypic variation observed for relatively conservative missense changes would derive from disruption or enhancement of helix-helix packing in the raft state. The results reported, that Y shares with S105 the same extreme sensitivity to missense changes within the TMDs provide additional support for this perspective and suggest that this model might be applicable to all class I holins.

Antiholin function and sensitivity of hole formation to the topology of TMD1

The data presented here show that P2 LysA expressed *in trans* can retard Y triggering but not S105 triggering (Fig. 16B), indicating that LysA is the P2 antiholin. In this light, the rather weak lysis timing phenotype associated with P2$lysA_{am}$, with only a 5-10 min early onset of lysis, is not unexpected. Similar small advances in triggering times were associated with the equivalent antiholin-defective mutants of *S* and S^{21}, and T4*rI*- lysis timing is unaltered compared to the wt. It should be noted that in all three cases, over-expression of the antiholin resulted in a complete blockage of holin triggering. Repeated efforts to achieve a greater lysis delay using over-expression constructs for *lysA in trans* to the P2 lysis cassette were unsuccessful (data not shown). It is possible that a physiological signal is needed to activate LysA, as in the case of T4

RI during T4 lysis inhibition; efforts are underway to look for *lysA* missense mutants that are non-permissive for lysis.

It is interesting that despite the identical topology, TMD1 is required for Y function and deletion of TMD1 creates a negative-dominant allele. It was therefore not possible to convert *Y* to a functional antiholin by adding one or two basic residues to the N-terminus, as is the case in the S107 and gp13_{108} antiholins of λ and P22. Although the altered Y proteins had Y-specific inhibitory character, collapsing the membrane potential did not rescue lytic function. Thus the addition of a positively charged residue at the N-terminus of TMD1 is not sufficient to generate a useful antiholin protein, which presumably must be both inhibitory during the latent period and activatable by membrane depolarization at triggering time. The simplest interpretations are that the addition of the basic residues do not prevent entry of TMD1 into the bilayer but abrogate Y function in another fashion, or that depolarization of the membrane does not restore proper TMD1 localization. Thus, it seems that *Y* would have to have more changes than just a dual start to encode a useful antiholin product, presumably accounting for the necessity of having a distinct gene to encode an antiholin. The most striking difference between the TMD1s encoded by *S* and *Y* is that the former has the sequence KEQGIG, which is both hydrophilic and high in turn-propensity, in the middle of the helix, where Y has the much more hydrophobic sequence GVLIVV. It will be of interest to select suppressors of the Y_{+K} lysis defect to see if they confer intrinsic antiholin properties.

Involvement of the cytoplasmic C-terminal domain

In S105, the cytoplasmic CTD is largely dispensable, requiring only that at least one positively charged residue, Lys92, be present, presumably to secure TMD3 in its proper topology. Indeed, shortening the C-terminal domain caused the lysis timing to become progressively earlier, suggesting that the highly positively-charged cytoplasmic tail played a regulatory role, possibly by interacting with the anionic surface of the cytoplasmic membrane. However, in Y, similar deletions inactivate the holin function, suggesting that the C-terminal cytoplasmic domain of Y plays a direct role in hole-formation, rather than a regulatory one. Oddly, the appending of an oligohistidine tag had no effect on lysis timing for Y, whereas the analogous tagging inactivated S105 (88). Taken together, this suggests that the dispensability and regulatory function of the cytoplasmic CTD cannot be taken as a general property of class I holins. In this light, it should be noted that the λ CTD (23 residues) is much longer than the Y CTD (14 residues); it will be interesting to see if the size difference is indicative of regulatory or hole-forming function.

Hole formation

A key functional difference between holins and other proteins that are involved in export of proteins across the bilayer is the size of the membrane channels involved. As noted above, in the case of S105, it was shown that a pre-folded R-βgal chimera of nearly 0.5 MDa mass was fully functional as an endolysin. Here I have shown that the same is true for P2 Y (Fig. 17A), suggesting that this is a feature common to class I holins. Recently, cryo-electron microscopy and tomography has revealed that the S105

lesions could be visualized as interruptions in the cytoplasmic membrane. These lesions were massive and heterogeneous, with an average diameter of >~ 300 nm and some extending to micron size. I presume that these interruptions constitute the large, non-specific holes required for escape of the R-βgal chimera to the periplasm. Although the fine structure within the large interruptions is still unclear, the sizes observed are consistent with the notion that all of the holin protein present at the time of triggering goes into lining the holes. It will be important to conduct similar cryo-EM studies on Y and other holins to determine whether these lesions are common feature for class I holins and, if so, how the holin molecules are arranged within them. In addition, the fact that both Y and S105 apparently form much smaller holes when triggered prematurely suggests that the surprising size of the holes, far bigger than needed to effect release of the typical phage endolysin, is a consequence of the timing mechanism rather than an evolutionarily selected feature.

Oligomerization

The results presented here show that the functional Y protein can be cross-linked into multimers in vivo, and that lysis-defective Y mutants show a defect in this property, both features shared by S105 and S105 lysis-defective mutants, respectively. In addition, in both cases, a dimeric species can be generated by cross-linking, and this species can be detected irrespective of the lytic function of a *Y* or *S* allele. This suggests that one of the fundamental properties of class I holins is the propensity for homo-oligomerization. In addition, multiple lines of evidence indicate that for S105, a dimeric species appears to be the mobile intermediate that exists before triggering in the lysis

pathway. The presence of the dimeric cross-linking product in both Y and Y mutants is at least consistent with the notion that this too is a common feature of class I holins.

There are some gene-specific differences, however. First, the bulk of the Y holin cross-links into very large multimeric species, unresolvable by SDS-PAGE, whereas cross-linking of S105 generates a regular oligomeric ladder up to at least 6-mers. Since the cross-linking agents were not the same, the results are not directly comparable. Nevertheless, it is striking that the Y protein was much less extractable by detergent in a form that could retain solubility under 100,000g centrifugation. S105 purified in the non-ionic detergent DDM forms multimeric rings of ~ 1 MDa mass, estimated to contain >70 S105 molecules but small enough to retain solubility. Despite the suggestive morphology and the fact that the purified S105 protein carrying the lysis-defective A52V substitution was restricted to dimers in the same detergent, these rings were judged not to reflect the structure of the S105 holes formed in vivo because the diameter of the lumen was far too small to accommodate the endolysin-βgalactosidase chimeras used to size the hole in vivo. Instead, it was concluded that the ring formation was a detergent artifact that nevertheless reflected the ability of the functional S105 holin, but not the defective $S105_{a52v}$ protein, to homo-oligomerize. I posit the same thing is true here for Y, which forms apparent helical structures in DDM, highly unlikely to be representative of an in vivo structure. Because these helices are not closed rings, much of the detergent- extracted material is likely to be too massive to persist in the soluble state after 100,000g centrifugation. Unlike lambda, however, the lysis-defective mutant I tested was not restricted to a monomer or dimer state; nevertheless, the distinct

difference between the oligomers formed by the wt and mutant proteins serves to underscore the point that holin quaternary structure in detergent is unlikely to retain features established in the context of the membrane.

CHAPTER III

THE HOLE STORY: DEFINING THE STRUCTURE OF THE S-HOLE

Introduction

Phage lysis begins when the phage-encoded holins suddenly form lethal lesions in the cytoplasmic membrane. Early in the study of phage lysis, these lesions were designated "holes" to distinguish them from channels and pores, the major structures that had already been characterized in bacterial membranes (16). In addition, the term "hole" implies that these lethal membrane protein complexes allowed non-specific escape of muralytic proteins from the cytoplasm of the infected cell. The two most important questions pertaining to holins and their role in phage lysis are: (a) what is the mechanism of hole-formation and how is its temporal schedule established; (b) what is the structure of the hole and how does this allow non-specific escape of large folded proteins. As is common for membrane complexes, little or no high-resolution structural information is available for the purified proteins. Instead, mechanistic understanding of such complexes usually requires a combination of genetics, ultrastructural studies, and topological investigations. Two key physiological characteristics have been established that must be explained in any model for holin function. First, holin triggering, the event that corresponds to sudden hole formation, can be prematurely imposed by partial depolarization of the membrane (5), suggesting a role for the membrane potential in regulating holin function. Second, for all three canonical holins that have been studied in detail, i.e., lambda S105, P2 Y and T4 gp*t*, the holes that are formed are sufficiently

large to allow release of fully functional 0.5 MDa endolysin-βgalactosidase chimeras (12, 66). The ability to release such large proteins suggested that the holin complexes formed lesions of an unprecedented size. A physical basis for this was provided for S105 by results of cryo-EM studies of cells in which the *S105* gene had been induced to form its lethal holes (6). The holes were visualized as interruptions in cross-sectional views of the cytoplasmic membrane and also by tomography. The holes were of micron scale, with an average diameter of >340 nm and ranging in size up to >1 μm. Similar holes were identified for the unrelated class I holin Y of the phage P2 and also for gp*t*. A cytological basis for these cryo-EM results was reported by White et al., who used S105-GFP chimeras to follow the holin-mediated lysis pathway (37). These fusion proteins were seen to accumulate uniformly within the membrane until suddenly redistributing to large foci, or rafts, at the time of triggering. A mutant S105-GFP chimera carrying a missense change that blocks S105 oligomerization and hole formation failed to undergo either triggering or the sudden redistribution, and instead accumulated indefinitely in a uniformly dispersed manner. A model that integrates these physiological and ultrastructural results is that triggering involves the sudden nucleation of two-dimensional aggregates, the holin rafts, which grow rapidly and then lead to local depolarization, followed by a re-organization of the holin into the lethal hole structures.

Many aspects of this integrated model rely on inferences derived from the phenotypes of a number of *S* mutants. This chapter reports studies involving phenotypic analysis of a systematic library of clustered site-directed mutants of S105, in which every residue of a complete turn of the central region of TMD2 is changed to all possible

residues. To establish a context for such a systematic mutational analysis, it is first necessary to review the previous mutational analyses of *S*. For reference, a master table listing all the known missense mutants of the *S* gene is provided (Table 3), including parameters that describe the lysis phenotype (lytic or non-lytic) and, if lytic, the lysis time (triggering time). The chapter concludes with experiments designed to probe the structure of the mature "S-hole" in the membrane of the cell, using chemical probes.

Materials and methods

Materials, strains, bacteriophage, plasmids, and growth media

The *E. coli* strain XL1Blue, the lysis-defective thermo-inducible prophages *λΔSR,* and the lysis-proficient thermoinducible prophages λ*S105* (expressing *S105*) have been described previously (15, 31, 32, 63). For experiments requiring coexpression of the wt *S105* with certain derivative alleles, prophage *S105* was used to provide one S105 allele and Q, which transactivates plasmids pS105. In general, *S105* was expressed from plasmid pS105, which carries the phage lambda lysis gene cassette, genes (SRRzRz1) under its native promoter pR'. The other lysis genes were inactivated by nonsense mutations, so that only *S105* is expressed. Bacterial strains, phages, and plasmids used in this work are listed in Table **4**.

Media, growth conditions, IPTG induction and thermal induction of the λ lysis genes from a prophage and/or plasmid have been described previously (20, 24, 29, 32). Bacterial cultures were grown in standard LB medium supplemented with ampicillin (100 μg/ml), kanamycin (40 μg/ml) and chloramphenicol (10 μg/ml) for the maintenance of plasmids and prophage, respectively. Lysis curves were repeated a

Table 3. Lysis times and dominant/recessive phenotypes of *S* mutants.

Codon	Mutant	Lysis time *in trans* to (λ*ΔSR*)	Lysis time *in trans* to (λ*S105*)	Phenotypic class
	Sam7 (W55am)	non-lytic	50’	R
	Wt (S105)	30’	25’	
12	A12T	30’	35’	--
16	A16N	30’	30’	--
18	E18K	non-lytic	non-lytic	D
21	I21T	25’	25’	--
22	G22R	non-lytic	non-lytic	D
22	G22W	non-lytic	60’	d
22	G22E	20’	25’	--
23	A23T	80’	55’	d
23	A23V	non-lytic	non-lytic	D
25	L25G	non-lytic	70’	d
25	L25V	non-lytic	non-lytic	D
28	A28T	25’	25’	--
29	M29I	72’	45’	--
30	A30S	38’	55’	d%
30	A30V	non-lytic	non-lytic	D
31	Y31I	30’	30’	--
33	R33C	non-lytic	non-lytic	D
33	R33H	non-lytic	non-lytic	D
33	R33L	non-lytic	non-lytic	D
34	G34S	non-lytic	non-lytic	D
35	R35K	non-lytic	55’	R
38	G38S	non-lytic	non-lytic	D
39	G39D	20’	16’	--
39	G39V	23’	22’	--
47	D47Y	non-lytic	35’	s
48	A48T	non-lytic	25’	S
48	A48V	non-lytic	40’	R
49	T49A	75’	50’	--
49	T49C	70’	45’	--
49	T49D	60’	40’	--
49	T49E	40’	35’	--
49	T49F	90’	45’	R
49	T49G	85’	50’	R

Table 3. Continued

Codon	Mutant	Lysis time *in trans* to (λ*ΔSR*)	Lysis time *in trans* to (λ*S105*)	Phenotypic class
49	T49H	non-lytic	50’	d
49	T49I	50’	40’	--
49	T49K	60’	45’	--
49	T49L	35’	35’	--
49	T49M	50’	40’	--
49	T49N	60’	45’	--
49	T49P	non-lytic	40’	s
49	T49Q	70’	45’	--
49	T49R	50’	35’	--
49	T49S	40’	35’	--
49	T49V	50’	45’	--
49	T49W	90’	50’	R
49	T49Y	non-lytic	50’	R
50	M50A	90’	25’	S
50	M50C	80’	30’	S
50	M50D	90’	35’	s
50	M50E	50’	30’	--
50	M50F	non-lytic	45’	s
50	M50G	65’	25’	--
50	M50H	non-lytic	55’	R
50	M50I	85’	50’	R
50	M50K	90’	30’	S
50	M50L	non-lytic	45’	s
50	M50N	non-lytic	30’	S
50	M50P	non-lytic	25’	S
50	M50Q	65’	45’	--
50	M50R	non-lytic	35’	s
50	M50S	85’	25’	s
50	M50T	non-lytic	45’	R
50	M50V	non-lytic	30’	S
50	M50W	non-lytic	35’	s
50	M50Y	non-lytic	35’	s
51	C51A	36’	22’	--
51	C51D	77’ (slow)	35’	--
51	C51E	non-lytic	45’	R
51	C51F	non-lytic	45’	R
51	C51G	70’	47’	--
51	C51H	non-lytic	24’	S

Table 3. Continued

Codon	Mutant	Lysis time *in trans* to (λ*ΔSR*)	Lysis time *in trans* to (λ*S105*)	Phenotypic class
51	C51I	53’	27’	--
51	C51K	50’	30’	--
51	C51L	53’	25’	--
51	C51M	16’	16’	--
51	C51Q	non-lytic	25’	S
51	C51R	non-lytic	50’	R
51	C51S	25’	25’	--
51	C51T	non-lytic	45’	R
51	C51V	33’	24’	--
51	C51W	non-lytic	50’	R
51	C51Y	non-lytic	40’	R
52	A52C	80’ (slow)	55’ (slow)	d
52	A52D	non-lytic	non-lytic	D
52	A52E	non-lytic	50’	R
52	A52F	11’	11’	--
52	A52G	17’	17’	--
52	A52H	53’ (slow)	50’	--
52	A52I	non-lytic	55’	R
52	A52K	non-lytic	52’	R
52	A52L	15’	17’	--
52	A52M	23’	25’	--
52	A52N	non-lytic	non-lytic	D
52	A52P	non-lytic	45’	R
52	A52Q	75’ (slow)	38’	--
52	A52R	53’ (slow)	55’	--
52	A52S	non-lytic	75’ (slow)	d
52	A52T	non-lytic	non-lytic	D
52	A52V	non-lytic	non-lytic	D
52	A52W	60’	30’	--
52	A52Y	15’	15’	--
53	I53Y	non-lytic	non-lytic	D
55	A55T	70’	30’	--
56	W56S	non-lytic	non-lytic	D
56	W56Y	21’	19’	--
59	R59C	non-lytic	50’	R
59	R59H	75’	50’	--
59	R59L	60’	35’	--
60	D60N	non-lytic	57’	R

Table 3. Continued

Codon	Mutant	Lysis time (λ*ΔSR*)	Lysis time (λ*S105*)	Phenotypic class
62	L62F	75’	35’	--
66	G66E	25’	25’	--
71	L71F	55’	50’	--
72	A72C*	50’	15’	s
73	Y73C*	80’	15’	S
73	Y73F	non-lytic	non-lytic	D
73	Y73T	10’	15’	--
75	T75C*	24’	18’	--
76	S76C*	20’	20’	--
77	V77T	35’	35’	--
78	F78G	20’	20’	--
79	I79C*	90’	15’	S
80	G80C*	75’	35’	--
80	G80S	non-lytic	non-lytic	D
83	G83D	72’	50’	--
83	G83I	non-lytic	non-lytic	D
87	I87Y	non-lytic	65’	d
88	G88K	non-lytic	non-lytic	D
89	S89C*	20’	19’	--
99	A99C*	22’	15’	--
101	V101T	35’	35’	--
102	E102K	37’	38’	--
104	G104C*	35’	22’	--
108C	108C*	22’	18’	--

*These mutants also have the mutation *C51S* in cis.
% This lytic mutant exhibits interference with wt S105 timing.
Key:
D = negative dominant.
d = weakly dominant.
R = recessive.
S = synergistic (also shaded). This phenotypic class is for mutants that exhibit an absolute lysis defect or severely retarded lysis but, in the presence of wt, cause lysis as early or earlier than the case where wt S105 is *in trans* to the prophage S105.
S = weakly synergistic.

Table 4. Strains, phage and plasmids.

Strain, phage or plasmid	Genotype/features	Source or reference
Strains		
MC4100 *ΔtonA*	*E. coli* K-12 F$^-$ *araD139 Δ(argF-lac)U169 rpsL15relA1flbB3501 deo pstF25 rbsR ΔtonA*	Lab stock
MG1655 *ΔtonA lacI^q*	F$^-$ *ilvG rfb50 rph1 ΔtonA lacI^q*	Lab stock
MDS12 *ΔtonA*	MG1655 with 12 deletions, totaling 376,180 nt including cryptic prophages; *ΔtonA*	(89)
XL1Blue	*E. coli* K-12 *recA endA1 gyrA96 thi1 hsdR17 supE44 relA1 lac* [F' *proAB lacZΔM15::tn10*]	Stratagene
Phages		
λΔSR	*Δ(stf-tfa)::cat cI$_{857}$* Δ(*SR*)	(32)
λ*S105*	*Δ(stf-tfa)::cat cI$_{857}$ S$_{M1L}$*	(31)
Plasmids		
pQ	λQ cloned into pZS*24, *kan^R*	(29)
pS105	λ lysis gene region with *S$_{M1L}$* cloned into pBR322, *amp^R*	(25)
pS105$_{C51S}$	Same as pS105, but with *C51S* mutation	(24)
pKB1	Same as pS105, but with *Sam7* mutation	(25)

minimum of three times, and the lysis times listed in for each mutant in Table 3 are accurate within ± 2 minute.

Standard DNA manipulations, PCR, site-directed mutagenesis, and DNA sequencing

Isolation of plasmid DNA, DNA amplification by PCR, DNA transformation, and DNA sequencing were performed as previously described (24). Primers were obtained from Integrated DNA Technologies, Coralville, IA, and were used without further purification. Restriction and DNA-modifying enzymes were purchased from New England Biolabs; all reactions using these enzymes were performed according to the manufacturer's instructions. Single-cysteine substitutions were made in corresponding plasmids by site-directed mutagenesis using the QuikChange kit from Stratagene as described previously (24). pS105 was used as the template for all primers, except the cysteine mutation primers (excluding A52C For/Rev). $pS105_{C51S}$ was used as the template for all of the cysteine mutation primers. The DNA sequence of all constructs was verified by automated fluorescence sequencing performed at Eton Bioscience Inc.

Generation of recombinant bacteriophage

Recombinant phage was constructed as follows: Plasmids were transformed into MDS12 *ΔtonA* (λ*ΔSR*) and individual transformants were grown up and induced as previously described (24). Chloroform (final concentration, 1%) was added to the cultures 2 hours after induction, or when lysis was complete. Lysates were cleared by centrifugation (3,000 rpm in a clinical centrifuge) and used to reinfect early-log (A_{550} =

0.2) cultures of MDS12 *ΔtonA* at a multiplicity of infection (m.o.i.) of 5. Two hours after infection, chloroform was added and the lysate cleared. This enrichment step was repeated twice more, and the final lysate was used to lysogenize MC4100 *ΔtonA*. Lysogens were selected for by plating on LB plates containing chloramphenicol and were screened using Single Prophage PCR (SPP) (90) to determine which candidates contained a single prophage. These candidates were then cross-streaked against λ86 (λ^{imm434} c^- $R_{am54am60}$) to verify the presence of a functional endolysin; positive candidates were used for lysis curves and to create phage stocks.

TCA precipitation

1 ml or 5 ml culture aliquots were added to 111 µl or 555 µl, respectively, of cold, 6.1 N trichloroacetic acid, then placed on ice for 30 minutes. The precipitate was collected by centrifugation (15,000 rpm in a tabletop microcentrifuge or 3,000 rpm in a clinical centrifuge, respectively) and washed once with acetone, resuspending the pellet completely. Pellets were air-dried and resuspended in SDS-PAGE loading buffer. Proteins were separated on 16.5% SDS-PAGE with a 4% stacking gel. Western blotting and immunodetection with anti-S antibodies were performed as previously described (29).

Cysteine modification

Cultures were grown to A_{550}=0.4, induced by a thermal shift, as appropriate, and aerated past the time of holin triggering or, in the case of non-lethal S105 alleles, for 50 min. Cells corresponding to 0.25 A_{550} units were collected by centrifugation, washed

twice with 1 mL TBS buffer (25 mM Tris-HCl, 150 mM NaCl, pH 7.2), and then resuspended in 0.25 mL of TBS. Each sample was divided into two 125 μL aliquots. To one, 10 mM 4-acetamido-4'-((iodoacetyl) amino) stilbene-2,2'-disulfonic acid, disodium salt (IASD) from Invitrogen was added and to the other, an equivalent amount of water was added. After 30 min at room temperature, 50 mM L-cysteine was added to quench any unreacted IASD. After 10 min, the cells were diluted by the addition of 0.75 mL TBS, collected by centrifugation, washed twice with 1 mL TBS, resuspended in 100 μL PB, and extracted with of 750 μL chloroform:methanol:water (1:4:1). After incubation on ice for 30 min, the denatured and delipidated proteins were collected by centrifugation at 13,000g for 5min at 4°C. The protein pellets were washed once with 400 μL 95% methanol and resuspended in 100 μL PEGylation buffer [10 M urea/1% SDS/1 mM EDTA/0.6 M Tris, pH 7; adapted from Lu and Deutsch (91)]. 50 μL of each sample was transferred to a clean tube and treated with 0.2 mM PEG-maleimide (Creative Biochem) for 30 min at room temperature, and then precipitated with 1 mL cold ethanol. After overnight at 20 °C, proteins were collected by centrifugation at 13,000g for 15 min at 4 °C. The pellets were air-dried and resuspended in sample loading buffer for analysis by SDS/PAGE (92). For experiments involving depolarization of the membrane, 1 mM DNP was added into the culture at the time of harvesting and used to supplement TBS in each step until quenching with L-cysteine.

Western blotting

SDS/PAGE and Western blotting was performed as described (47), and can be found in SI Materials and Methods; 16.5% Tris-Tricine gels were used to separate

protein samples. An antibody raised in rabbit against the S105 C-terminal peptide was used as primary antibody to detect S105 protein variants (46). Horseradish peroxidase conjugated goat-anti-rabbit secondary antibody was from Pierce.

Results and discussion

Mutational analyses of the lambda holin gene

Over a period of more than 50 years, there have been several different mutational studies of the *S* gene, beginning with the first identification of *S* as a locus for mutations that blocked lysis and plaque-formation (11). One of the mutants isolated in this hunt was Ala55Thr, a relatively conservative change in the middle of TMD2 and still the only temperature-sensitive allele known for this gene. In the most extensive mutant hunt, Raab et al. used a system with a medium-copy plasmid clone of *S* expressed from the *lac* promoter to isolate survivors of induction (26). These mutations mapped to all three putative TMDs, plus a few in the linkers connecting the TMDs. They then recombined these mutations back into the lambda context and tested the lysis phenotype, as single lysogens and also in the presence of a second prophage with the parental *S* gene, thus getting an indication of dominance or recessiveness. The most unexpected mutations were dominant single base changes upstream of the start codon. The analysis of these mutations led to the discovery that *S* produced both the S105 holin and S107 antiholin, with the choice of start codons controlled by the secondary structure in which the original mutations were found. Most of missense mutations in the coding sequence had an absolute lysis defect, but some were restored to lytic function, albeit with a delayed triggering time, in the phage context, presumably because the expression level was

higher, and thus, compensated for the defect. This was an early indication that S holin function was tightly linked to quantitative parameters. It was striking that numerous mutations in the TMDs were conservative in terms of hydrophobicity; i.e., Leu25Val. However, the biggest surprise was obtained when this collection of lysis-defective and lysis-delayed alleles were tested for dominance/recessiveness. Both dominant mutants and recessive mutants were obtained in all three TMDs as well as the connector loops. This was expected since it was anticipated that, given its size, the *S* holin would have to be oligomeric to have lethal hole-forming function; some mutant proteins would poison the oligomerization process and be dominant and others might simply not participate and be recessive. Surprisingly, a third class was also found, called "early dominant", in which the defective allele, i.e., Ala48Val, when tested with equal expression of the wt *S* gene, caused lysis either at the same time or earlier than was observed for two wt alleles. (These are now called "S" for synergistic in Table 3). These observations led to the notion that intermolecular interactions between TMDs, presumably helical packing, played an important role in determining lysis timing. Helical packing interactions would be expected to be exquisitely side-chain specific, so mixed pools of different missense mutants and wt proteins could have unpredictable association/dissociation properties (93, 94).

A single important mutant was obtained by a suppressor selection, starting with the lysis-defective Ala52Val and selecting for a suppressor mutation that restored lysis (Fig. 23). Only a single mutant was obtained, which turned out to be a change in the same codon, converting the allele to Ala52Gly. This change restored lysis but not

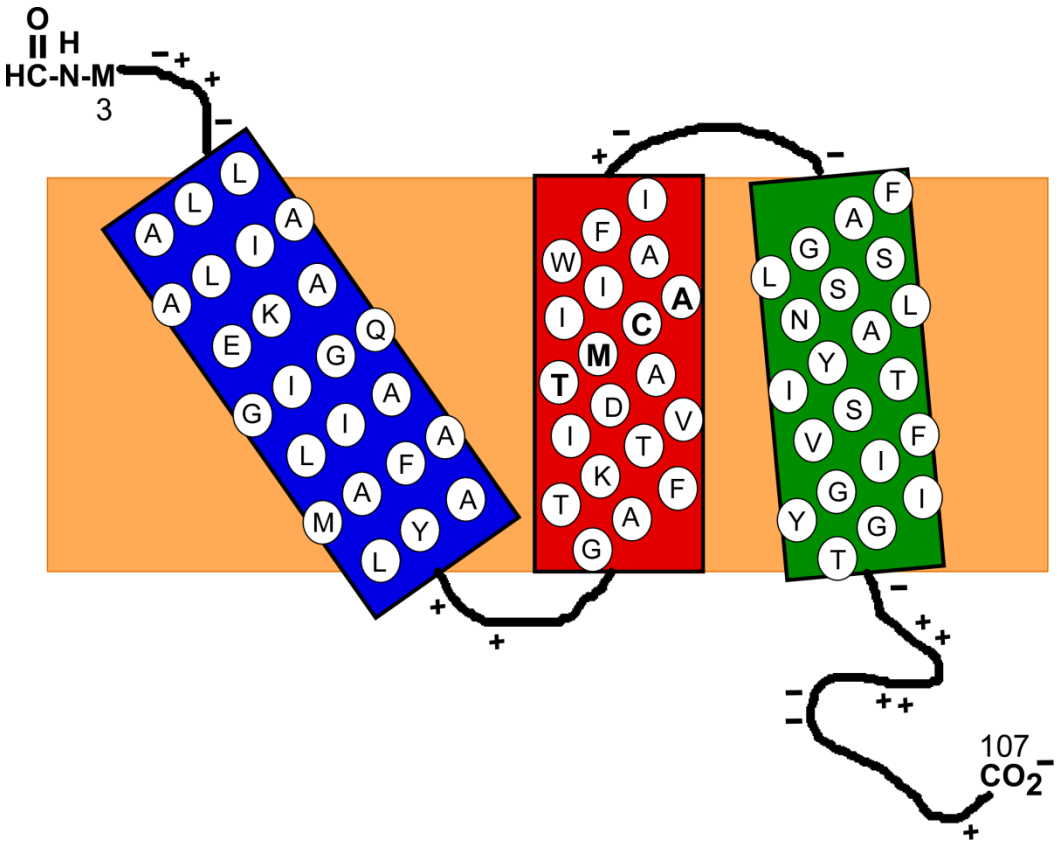

Fig. 23. Topology of S105. Amino acids are indicated in each TMD. TMD1 (blue); TMD2 (red); TMD3 (green). Bolded amino acids in TMD2 are Thr49, Met50, Cys51, and Ala52 that are the main focus of this genetic study.

plaque-formation, because the lysis time was, at 20 min, so early that no phage particles were assembled. The existence of the two non-plaque forming alleles Ala52Val and Ala52Gly, with opposite lytic phenotypes (i.e., non-lysis and early lysis) was formal proof that the function of the *S* gene was to (a) cause lysis and (b) provide a viable termination time to the infection cycle.

It should be pointed out that after the Raab study, which first identified the dual start motif, it was realized that the wt *S* gene was a poor subject for genetic studies because it produced two protein products with opposite functions, to effect and prevent lysis. Thus, in principle, every lysis or timing phenotype associated with a missense change had to be interpreted in terms of positive or negative effects on the holin and the antiholin. Thus, from then on, genetic and physiological studies used an altered parental gene, *S105*, in which the start codon for the S107 antiholin was inactivated and thus only one protein product, the S105 holin, was produced.

Three other mutational studies of *S105* involved site-directed mutagenesis for various reasons (24, 31, 95). As part of a process to establish a purification system for the holin, Smith conducted a "his-tag" mutagenesis of *S* by inserting an oligohistidine linker at 13 positions, including the N and C-termini (31). Almost all these insertions inactivated the function of the gene. However, two insertions, at positions in the C-terminal cytoplasmic domain, just after the end of TMD3, were found to retain function. One of these was used for large-scale overproduction and purification of the holin. Gruendling was studying the topology of the S105 protein and, to this end, wanted to have a library of *S105* mutants with a single Cys residue at different positions, for use in

cys-accessibility studies (24). The parental *S105* allele has a single Cys codon at position 51, which was changed to Ser as the parental allele for a series of single-Cys substitutions throughout the *S105* reading frame. The parental Cys51Ser retained plaque-forming ability although the result was that every Cys substitution, including the parental Cys51Ser, produced a significant change in lysis timing, although whether the change was to advance or delay lysis timing was unpredictable. This suggested that a fundamental character of the S105 gene, and of holin genes in general by extension, was that they can be easily mutated to radically different lysis timing by a single missense change. This conclusion was reinforced by the work of Zheng (2008). Zheng compared *S105* to the *E* lysis gene of the microvirus ϕX174, which causes lysis by inhibiting a step in cell wall synthesis (96). Single codon changes were chosen randomly in both genes by a computer algorithm and then tested for lysis timing. In *E*, most changes either had no effect or disrupted E function. For the 10 random codon changes in *S*, two had unchanged lysis time but 8 others, scattered throughout all three TMDs, exhibited large changes in lysis timing, either earlier or later than wt. This reinforced the idea that a fitness parameter for holins is their easy mutability to altered lysis timing.

Overview of the directed mutagenesis system

As noted above, some provocative mutant phenotypes had been found for mutations in TMD2, with alleles obtained either by selection or by site-directed mutagenesis; the best examples are the opposite phenotypes of Ala52Gly and Ala52Val, which are both absolute defectives in terms of plaque-formation, for opposite reasons, as noted above. We reasoned that more dramatic timing phenotypes might be obtained by

site-directed mutagenesis in this region and also that, potentially, one or more faces of TMD2 might be genetically insensitive to conservative (hydrophobic) substitutions if it faced the lipid, rather than forming a protein-protein interaction surface in the S105 lysis pathway. To test this idea, an entire alpha-helical turn of TMD2, spanning from residues 49 through 52, was subjected to complete site-directed mutagenesis, in that each of the four residues was changed to all possible missense codons. For simplicity, the approach was to conduct the mutagenesis on an *S105* gene mounted in a medium copy plasmid under its native transcriptional control, *in trans* to a prophage deleted for the *S* gene. Thermal induction of the prophage causes trans-activation of the cloned lysis cassette and gives reproducible lysis timing, slightly earlier than with an induced *S105* prophage (16). To serve as a basis for comparison, the lysis phenotypes of all mutant alleles of the *S* gene obtained from the studies summarized above are listed in Table 3. For each allele, the lysis timing obtained by inducing the allele in the trans-activation system is shown. Also listed is the lysis time for induction the same system *in trans* to the *S105* wt gene, to assess dominance/recessiveness. Finally, for comparison, the phenotype obtained for many of these alleles in the double prophage system used by Raab (1986) is listed (26). Some minor differences are observed, presumably due to small differences in the relative expression levels from the medium copy plasmid and the induced prophage, despite the identical transcriptional context.

Overview of the results

The results were incorporated into Table 3, and are depicted as lysis curves in Fig. 24. In each panel of the figure, the curves are color and symbol coded for all 20 AA, with the wt for each position indicated by an asterisk. Triggering times were estimated as the earliest point at which the A_{550} values began to decrease. Lysis curves were determined to be normal if the profile exhibited sharpness comparable to the wt allele, irrespective of triggering time up to ~ 90 min. Partial lysis observed beyond that time was judged to be lysis-defective and possibly due to non-specific membrane damage incurred as a result of the prophage induction. The lysis-defective mutants were screened for the accumulation of S105 protein by Western blot. In most cases of lysis-defective alleles, the accumulation was normal within the accuracy of the Western blot, with the exception of His (Fig. 25).

The results lead to several general conclusions. Overall, the hypothesis that any face of TMD2 was genetically insensitive to substitutions that preserved or increased hydrophobic character, as judged by the Wimley-White interface or octanol scales was clearly false (97). This is best seen by looking at the phenotypes of 7 substitutions (V, M, I, L, Y, F and W) that are judged to be of equal or higher hydrophobic character compared to the wt residue in each of these four positions. Overall, among these 28 substitutions, nearly half (13) abolished holin function entirely, and four more exhibited retarded lysis. Nonetheless, each position, representing different faces of TMD2, had different patterns in these substitutions. All seven of these changes inactivated the holin in position M50, whereas the other three positions had some retarded lysis and non-lytic

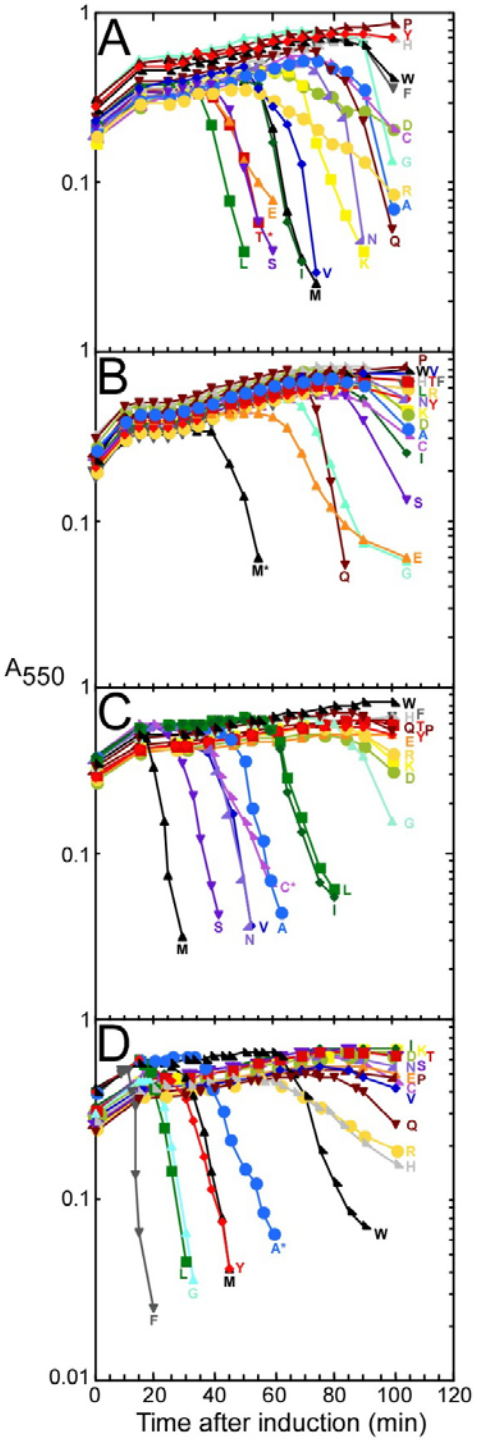

Fig. 24. Lysis curves of position 49, 50, 51 and 52 mutant series. Mutations are labeled on the graph. A) Position 49 series; B) Position 50 series; C) Position 51 series D) Position 52 series. Asterisk indicate wild-type.

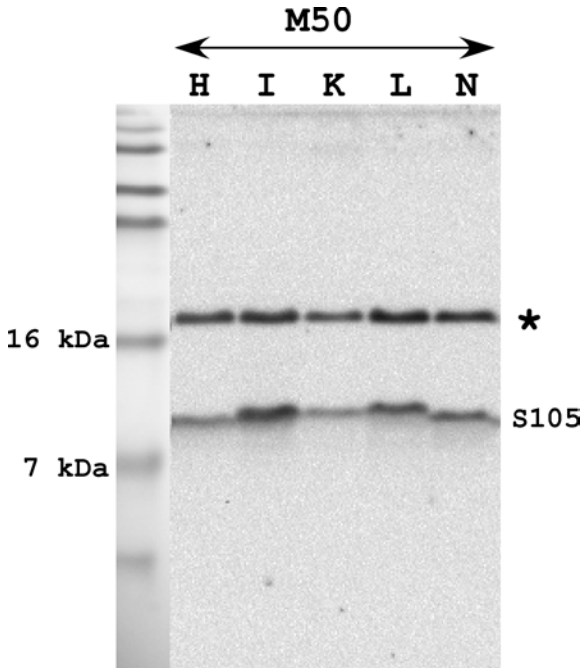

Fig. 25. Western blot of lysis-defective alleles. Five representative lysis-defective alleles are shown: M50H, M50I, M50K, M50L, and M50N. Asterisk indicates the loading control.

alleles. T49I, T49V, and T49M showed later lysis, whereas T49 to W, F and Y were non-lytic. C51I and C51L show later lysis times; C51Y, C51F and C51F are non-lytic. A52W has a later lysis time than wt; A52V and A52I are non-lytic. Taken together, this suggests that all faces of TMD2 participate in helix-helix interactions at some point in the hole-formation pathway, and, as would be expected from close helical packing, these interactions are sensitive to the bulk or shape of the side-chains at each position.

Hyper-sensitivity at position 52

The sensitivity to side-chain structure at position 52 was extraordinary. In view of the strikingly opposite phenotypes of Ala52Gly and Ala52Val, the simplest expectation was that large substitutions, like Ala to Val, here would impede lytic function. This hypothesis was wrong, as two of the largest possible substitutions, Ala52Phe and Ala52W had opposing changes, the former causing triggering at ~ 11 min, whereas the latter exhibited significantly delayed lysis. Even more remarkable in terms of extremely different phenotypes were Ala52Leu and Ala52Ile, with the former exhibiting very early lysis and the latter being completely non-lytic, with even a tighter defective phenotype than the Ala52Val. These two alleles, in the context of the phage, would certainly both be non-plaque formers, like Ala52Gly and Ala52Val, for reasons of too early and non-lysis, respectively. Altogether, five missense substitutions, all with non-polar residues, resulted in accelerated lysis (Phe, Leu, Gly, Met, Tyr) and one exhibited retarded lysis (Trp). All other substitutions either completely inactivated hole formation or resulted in gradual and partial lysis. As expected, all the charged and highly polar substitutions in this position fall into the inactive category.

Phenotypic spectrum at positions 49, 50 and 51

No other position exhibited comparably dramatic phenotypic variability. Only two other alleles out of the 57 tested in positions Thr49, Met50 and Cys51 had a significant early lysis phenotype and both of these (C51S and C51M) were in the Cys51 position. Cys51 did exhibit phenotypic eccentricity similar in character to that observed at position Ala52 but reduced in scale: C51I, C51L showed a ~ 20 min delay in triggering, whereas all three bulky aromatic substitutions (C51W, C51F and C51Y) were non-functional. In both positions Thr49 and Met50, which should be on the opposite side of the TMD2 helix as Cys51 and Ala52, almost all substitutions caused a significant lysis delay or inactivated the holin. Even here, there were marked differences, with Thr49 tolerating 13 of 19 different substitutions, albeit with retarded triggering times, whereas all but 4 of the 19 substitutions inactivated Met50.

Common features

One constant was that in each case, a Pro substitution inactivated the holin protein; the simplest interpretation is that a proline "kink" in the middle of TMD2 abolishes necessary helix-helix interactions required for the holin pathway. However, there were position-dependent differences in the dominance tests for the four defective proline substitutions. Pro substitutions at positions 49, 50 and 51 had early dominant or synergistic phenotypes, whereas the position 52 substitution had negative dominant character.

A similar disruptive effect to holin function was expected for polar and charged substitutions, considering this is in the middle of TMD2. Overall, this expectation was

met. Of the 24 cases where a charged (Glu, Asp, Arg, Lys) or highly polar (Asn, Gln) residue was substituted, all but six resulted in lysis-defective alleles, including, as noted above, all such substitutions at position Ala52. Nevertheless, the exceptions were striking. The substitutions T49E and C51N exhibited wt lysis timing, and T49K, T49Q, M50E, and M50Q showed normal but delayed lysis curves. TMD2 does have a predicted salt-bridged ion pair in Lys43/Asp47, so if the latter residue switched partners, the positive charge on Lys side chain in the T49K mutant could neutralized; the charged epsilon amino group on the embedded Lys43 residue could be accommodated by "snorkeling". However, no similar salt-bridging rationale can be invoked for T49E. It remains completely unclear how a product like $S105_{T49E}$, with an acidic residue in the middle of the bilayer, is effectively integrated in the bilayer, much less functions in the complex holin pathway. One possibility is that, during synthesis and localization of the $S105_{T49E}$ protein, TMD2 integration in the bilayer is delayed until the Glu side chain is protonated. Considering the small fraction of Glu residues that would be protonated at any given time at cytoplasmic pH, this might involve minutes of delay; however, the time scale of holin function is so long that this might not be readily apparent.

Interpretation of phenotypic variability

Overall, for these four positions 49, 50, 51 and 52 series, the comparative analysis of lysis times (listed in Table 3; Fig. 24) yielded no discernible patterns in terms of common effects of side chain bulk, shaper, polarity, hydrophobicity or charge. The parental side chains Thr, Met, Cys and Ala, are all rather neutral in terms of hydrophobicity, so the substitutions of the aliphatic, aromatic and polar residues have

dramatic effects on the predicted overall membrane propensity, either increased or decreased. The simplest interpretation is that all four surfaces are involved in intimate helix-helix interactions at some point in the holin pathway. In addition, the fact that the character of the mutant phenotypic spectrum is specific to the position suggests that the phenotypes are dominated by homotypic interactions. Position Met50 appears to be the least tolerant of substitutions, with no substitutions leading to early lysis and most causing inactivation of the holin. This suggests that a specific homotypic interaction on this face is required for forming the hole. Position Thr49 is somewhat more permissive for substitutions, maintaining some ability to trigger with nearly all changes except large bulky residues (Tyr, Phe, Trp, His) and Pro; however, only one substitution, T49L, has an early lysis phenotype, triggering ~ 5 min early. In contrast, both Cys51 and Ala52 are much more permissive. Leaving out the substitutions of charged residues or proline that would be problematical for maintaining the TMD helix, 8 of 14 substitutions at Cys51 and 6 of 14 substitutions at Ala52 retained normal sharply-triggered lysis phenotypes. Even more strikingly, of these 14 mutant alleles with normal lysis phenotypes, 7 were early-triggering mutants, 2 of 8 at Cys51 and 5 of 6 at Ala52. This suggests a face of TMD2 occupied by these residues is involved in homotypic interactions that must break or form during the pathway. It is reasonable to assume that most mutational changes to a homotypic interaction surface would be deleterious to the interaction, so that the preponderance of early lysis phenotypes, especially at Ala52, further suggests that weakening a homotypic interaction at this surface accelerates the triggering pathway. The exquisite sensitivity of triggering time to the shape of the side chain at position 52 is

worth noting again. It is difficult to imagine anything but intimate helical packing that could detect the differences between Leu, Ile and Val side chains, leading to catastrophically early lysis in the case of Leu but a completely non-lytic phenotype for Ile and Val. The same can be said for the dramatic differences between the Phe substitution, which triggers at 11 min after induction, compared to the Trp substitution, which triggers 25 min later than wt.

Predictions for and exploitation of the earliest lysis allele

It is important to know how many S105 molecules are present at the time of triggering of $S105_{A52F}$; however, S105 levels before about 10 min are undetectable using Western blots. Nevertheless, a rough estimate can be made from known parameters. Under the conditions used here, late protein synthesis starts at about 6-8 min after induction, so a triggering time of 11 min would mean only 3-5 min of expression of the holin gene, about one-tenth the normal time of expression. Because mRNA is unstable, the level of induced *S* mRNA would still be increasing at the time of triggering for the A52F allele, so the total amount of translation would be considerably less than ~ 10% of the wt, which triggers at ~ 40 min, about 32-34 min after induction and long after the steady state is reached for the *S105* mRNA. If we use 5% as a reasonable estimate, then it is likely that the total amount of $S105_{A52F}$ is on the order of ~ 200 molecules or less. Among other things, the total perimeter of the hole that could be occupied by the $S105_{A52F}$ protein would be <600 nm, so the S-holes formed by this allele should be significantly smaller or fewer in number than the wt.

Finally, the extreme early lysis phenotype of $S105_{A52F}$ may provide an approach for mapping interhelical contact points that are important in determining lysis timing. As noted above, in the phage context this would have a tight non-plaque forming phenotype, since the first virion would not be assembled until ~ 20 min after induction. It should be possible to use this tight defective phenotype to find intragenic suppressor mutations that restore the normal helix-helix interactions by simply inducing a lysogen carrying this allele and isolating rare plaque-formers. Only four residues provide plaque-forming ability at this position: the wt Ala, the early lysis residues Met and Tyr, Ala, and the late lysis residue Trp. Only the Tyr codons are accessible by a single base change from the Phe codon. Thus, there is a low background of same-codon revertants, which might make isolation of second-site suppressor mutations possible. If the hypothesis is right that the dramatic lysis timing phenotypes at Ala52 reflect disruptions of an interhelical interaction that retards lysis, the second site mutations would be expected to be in a position to alleviate that defect. Presumably this would have to be on the same face of the helix at about the same depth in the membrane, i.e., at Cys51.

Analyzing the S-hole with cysteine-scanning accessibility: differences between lytic and non-lytic alleles

The genetic approaches mentioned above only give a partial picture on how holin functions. In order to examine the mechanism of hole formation *in vivo*, we used biochemical methods to investigate the structure of the lesion. To probe the structure of the holin in the membrane, we used a variation of cysteine-scanning-accessibility. This approach involves treating whole cells expressing the desired S105 single-cysteine

mutant allele (Fig. 26A) with membrane-impermeant thiol reagent IASD (Fig. 26B). After quenching the reaction, cells are solubilized with SDS/urea, and the extent of IASD modification is assessed by derivatizing the remaining free cysteines with a high molecular mass (5 kDa) derivative of methoxy-PEG maleimide. For the hole-forming S105 alleles, residues in both the periplasm and cytoplasm should be modified by IASD, given a luminal diameter needed for passage of the reagent through the bilayer (98). We first examined the accessibility of cysteines located in either the periplasmic loop or the cytoplasmic domain of $S105_{C51S}$. As can be seen in Fig. 26B, positions in the periplasmic loop of $S105_{C51S}$ (His7) show partial protection, from comparing the unmodified species (asterisk) with or without IASD, or by comparing the PEGylated species, with or without IASD. (The intensities of the unmodified and PEGylated forms cannot be directly compared with each other because of inherent differences in the efficiency of blot transfer and immunostaining.) For both $S105_{C51S}$ (lytic) and $S105_{C51S\ A52V}$ (non-lytic), the protection is as efficient. Importantly, the protection of the cytoplasmic C-terminal positions (Ala94 and Cys108) is efficient for both the hole-forming alleles, $S105_{C51S}$ and $S105_{C51S\ A52V}$. Also, membrane depolarization by 2,4-dinitrophenol (DNP) has no effect on the $S105_{C51S}$ C-terminal accessibility to IASD (not shown). Together with previous findings, we conclude that the cytoplasmic membrane of cells expressing $S105_{C51S}$ and $S105_{C51S\ A52V}$ retains its permeability to IASD, and the lumen of holes formed by both $S105_{C51S}$ is large enough to permit the diffusion of IASD into the cytoplasm.

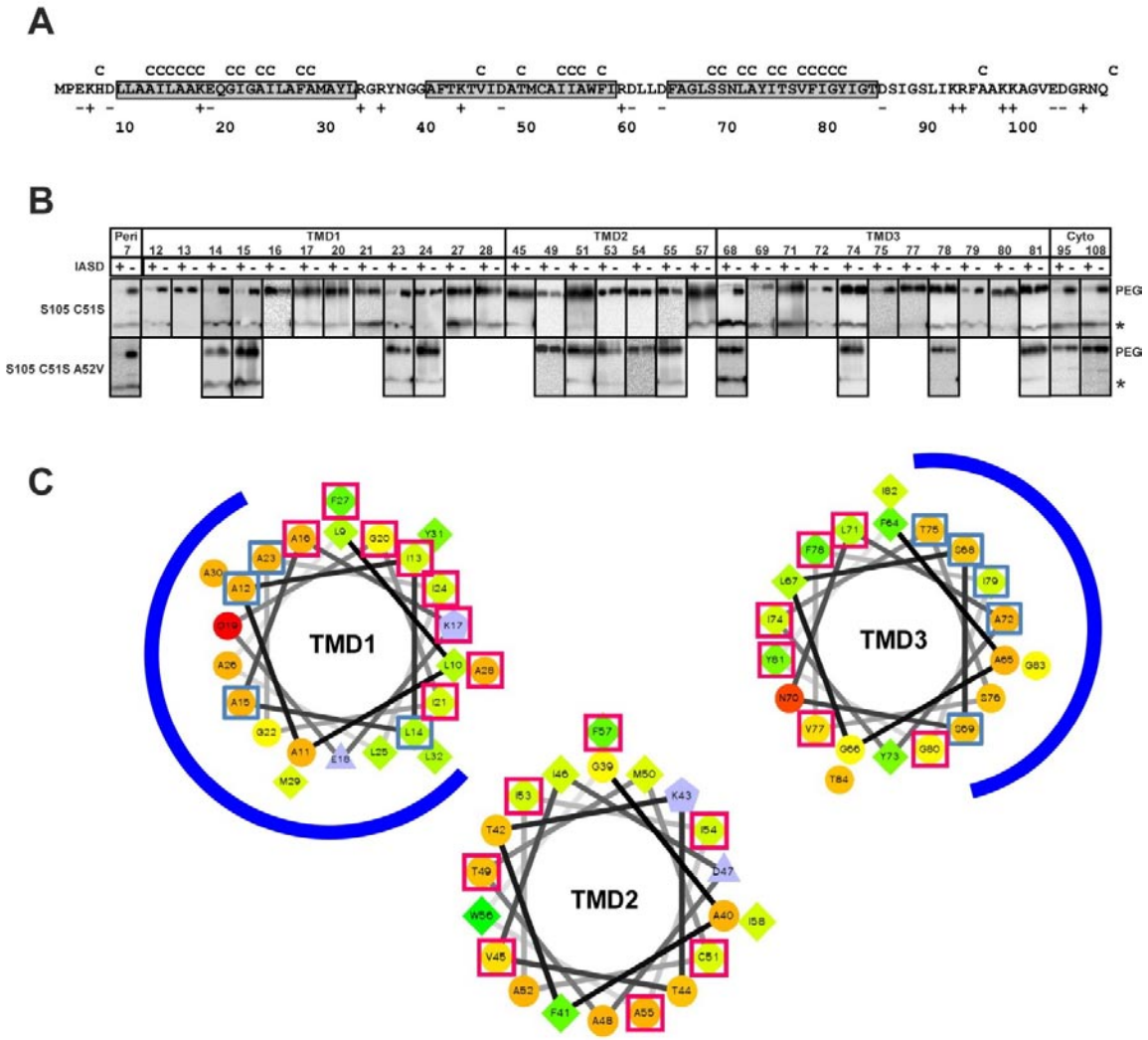

Fig. 26. IASD protection analysis of S105. (A) Whole cells carrying plasmids encoding single-Cys substitutions of S105 or its derivatives were induced, (B) Immunoblot analysis, as described in methods. The asterisk and ‘‘PEG’’ indicate the position of the unmodified monomer and the PEGylated species, respectively. IASD protection is indicated by an IASD-dependent decrease in the PEGylated species and increase in the unmodified species. NA, not available; Peri, periplasmic loop; Cyto, cytoplasmic tail. (C) IASD-protected positions in TMD 1, 2 and 3 map to its most hydrophilic face. Helical projections of TMD1, TMD2 and TMD3 are shown, with 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. The potentially charged residues are light blue. Blue square represent lumen facing side and pink square represent lipid facing side of the helix. The blue arches represent lumen facing side of TMD 1 and 3.

Data collected from the cysteine modification experiments involving holin mutant S105 $_{C51S\ A52V}$ strongly reinforces the assumption that all three TMDs of S105 $_{C51S\ A52V}$ are embedded in the membrane during the non-hole formation pathway. All residues around all three TMDs of S105 $_{C51S\ A52V}$, which does not trigger, are non-sensitive to IASD modification in cells (Fig. 26B). Also, it was shown that this non-hole forming topology completely block IASD from accessing the cytoplasm, where C-terminal cytoplasmic tail of S105 $_{C51S\ A52V}$ is not assessable to IASD (Fig. 26B). As expected, the N-terminal of S105 $_{C51S\ A52V}$ is still assessable to IASD modification due to its topology, where N-terminus is located in the periplasm (Fig. 23). Also the C-terminal cytoplasmic tail is assessable to IASD with hole forming allele, S105 $_{C51S}$, where ISAD is able to pass through the membrane lesions.

Identification of the TMDs residues lining the S105 channel

Since IASD penetrates the lesion formed by S105, it was reasonable to use the cysteine-scanning-accessibility method outlined above to determine which face of each TMD helix faces the lumen of the hole. Single cysteine alleles were engineered along the length of all three TMD helices encoded by S105$_{C51S}$, and S105$_{C51S\ A52V}$ (Fig. 26B). All alleles selected for single-cysteine-scanning-accessibility are as lytic as the wild-type; however, lysis timing varies for each of them. IASD did not react with any of the S105$_{C51S\ A52V}$ proteins carrying a single cysteine substitution in any of the three TMDs, which is consistent with its inability to form holes in the cytoplasmic membrane (Fig. 26B). By contrast, in S105$_{C51S}$ the same positions lining the hydrophilic face of TMD1 and TMD3 were found to be accessible to IASD (Fig. 26C), suggesting that these

residues face the lumen of the hole. The behavior of the lysis defective mutant $S105_{C51S\ A52V}$ confirms the connection between IASD sensitivity and lethal hole formation.

Conversely, for hole forming alleles of S105, positions along the surfaces of TMD1, TMD3, the N-terminus, and the C-terminus are assessable to IASD (Fig. 26B). As anticipated, the surfaces of TMD1 and TMD3 that become solvent exposed are their most hydrophilic surfaces (Fig. 26C). For TMD1, even though the arc of the accessibility begins with Leu (L14) and ends with Ala (A23), all but three (L25, M29, L32) residues on that face are either polar (Q19, E18) or of neutral hydrophobicity (A11, A12, A15, G22, A26, A30). For TMD3, the arc of the accessibility begins with Ser (S69) and ends with Thr (T75). All but one residue (I79) on that face are either polar (S68, S76) or of neutral hydrophobicity (A65, A72, G83). Nevertheless, cysteine was substituted for six residues that would occupy the different helical faces of TMD2. In the $S105_{C51S}$ context, all six positions around the helix were found to be inaccessible to IASD (Fig. 26C). These results suggest that during the pathway to hole formation, TMD2 is completely embedded in the lipid. Similarly, the same positions in the nonlethal $S105_{C51S\ A52V}$ derivative were completely insensitive to IASD, consistent with it being buried in the membrane.

Numerous lines of evidence suggest that the lesion generated by S105 helices must be adequately large enough to permit efficient IASD modification with luminal thiols (66). Its seems likely that several of the ~ 300 nm diameter holes is sufficient to depolarize the inner membrane within seconds and allow the endolysin to gain access to the cell wall. We suggest that this result reflects the biological mandate that at least two

TM helices are needed to line the large lesion in a straight line, assuming each TMD is ~ 1 nm diameter and half of the helix from both TMDs are facing the lumen (Fig. 27). Therefore, it allows instantaneous and complete depolarization of the membrane at an allele specific time. Importantly, protection for most of the lumen-facing residues approaches ~ 90% completion (Fig. 26B). This suggests that almost all the S105 protein, which is present at ~ 1000 – 3000 molecules per cell at the time of triggering is involved in the large holes based on quantitative immunoblotting (20).

General Model of the S105 hole formation pathway

Despite recent advances in the visualization of holin-triggered membrane lesions using cryo-EM (6), the arrangement of the TMDs that line the hole is still unknown. Here, the S hole formation pathway model in the biological circumstances comes from the thiol accessibility and mutagenesis data described above. At the molecular level, the lysis defect is associated with an oligomerization defect. In cross-linking experiments with the bifunctional agent DSP, Gründling et al. have showed that the parental S105 forms dimers and higher-order oligomers (25). And the authors concluded that the ability of S molecules to dimerize is not sufficient for the lytic step in holin function. All other defective *S* lysis-defective *S* mutants, which were tested by DSP cross-linking also efficient dimer formation, and except for S_{R59C}, all were deficient in oligomerization. Moreover, protein accumulation for S_{A48V} allele is much less than the wild-type S, and lower percentage of the total protein is in the disulfide dimer form. It is likely that S_{A48V} is proteolytically unstable, and suggesting that a reduced efficiency in dimer formation may lead to proteolysis (25). The model incorporates these findings

into a working model for hole formation. This model is depicted by a cartoon top-down view in Fig. 27. In this figure, TMD2 is a green circle with four surface quadrants numbered as 9 (Thr49), 0 (Met50), 1 (Cys51), and 2 (Ala52). TMD1 and TMD3 are shown as having blue (hydrophilic) and orange (hydrophobic) surfaces. Individual molecules are outlined in red or black, to distinguish intermolecular and intramolecular interfaces. The model depicts the fate of the holin, in this case lambda S105, during the entire morphogenesis period, i.e., from the outset of late gene expression to holin triggering and termination of the infection cycle.

In this model (Fig. 27), the holin accumulates as a stable, highly mobile dimer, with the dominant intermolecular interaction defined by helix-helix packing along face 1 of TMD2 and stabilized by the intramolecular sequestration of the hydrophilic faces of TMD1 and TMD3 against each other. (We posit that A48V is defective in formation of this dimer, although it is unclear whether the missense change directly affects TMD2-TMD2 interactions or imposes a steric defect on either TMD1 or TMD3, disfavoring the TMD1-TMD3 interaction.) The continued accumulation of holin molecules during the late gene expression phase gradually shifts more and more of the holin population into lower order oligomers, in dynamic equilibrium with the dimer and other low oligomer states. At a critical concentration, the dimers and low oligomers suddenly nucleate the formation of the two-dimensional aggregates (rafts), defined mainly by new homotypic intermolecular interactions on face 0 of TMD2. The rafts, which grow rapidly, are characterized by tight interhelical packing and are largely depleted of lipids. The loss of

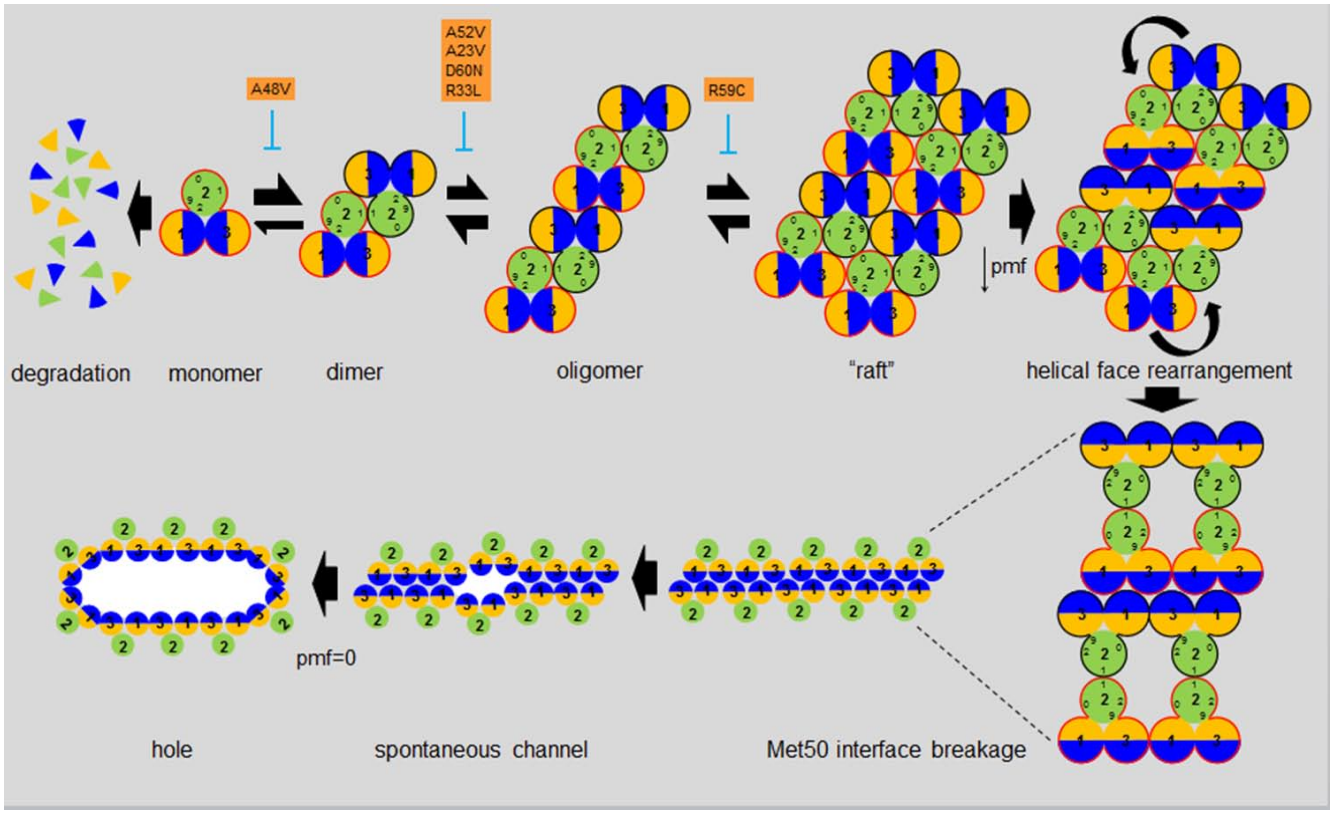

Fig. 27. General model of the S105 hole formation pathway. A top-down view of a region of the cytoplasmic membrane is shown; each circle represents a helical TMD. At least six steps, described in the text, are required for hole formation: first, dimerization of λ S; second, oligomerization; third, raft formation; fourth, helical-face rearrangements; fifth, Met50 interacting face breakage; sixth, spontaneous channel formations which lead to the hole. The *S* alleles which are blocked in different steps during the process of hole formation are indicated above the individual steps. The three TMDs (green, TMD2; sectored, TMD1 and TMD3) in a single S molecule are outlined with either red or black. TMD2 contains four interacting faces; 9, 0, 1, and 2 indicate the helical positions of T49, M50, C51, and A52, respectively. TMD 1 and 3 contain two interacting faces, hydrophilic (the face which also accessible to IASD) and hydrophobic, represented by blue and orange, respectively.

the lipid "insulation" leads to ion leakage and thus local deterioration of the pmf, providing the driving force for helical face rearrangements that result in the shift of the hydrophilic surfaces of TMD1 and TMD3 from intramolecular sequestration to intermolecular sequestration and breakage of the 0-face interactions, the weakest of the TMD2 homotypic interactions (last to form, first to break). This leads to large-scale formation of zipper-like structures, the immediate precursors to the holes, in which the local interfaces of the hydrophilic surfaces of TMD1 and TMD3 within the raft assume an energetically more favorable continuous interface. Conversion from the zipper arrangement to the actual hole would then be driven by progressive hydration from one or both surfaces of the bilayer, leading to the formation of spontaneous channels or micro-holes. Obviously, the formation of continuous aqueous channels would exacerbate the local collapse of the pmf, presumably leading immediately to depolarization throughout the cytoplasmic membrane. Note that ion leakage is proposed to favor changes that lead to more ion leakage and ultimately to total depolarization, presumably by favoring particular orientations of one or more TMDs of the holin molecule. Given the previous finding that holin triggering occurs in a saltatory fashion, in that the pmf is unaffected throughout the morphogenesis period while S105 accumulates, the model depends on having all of the events after raft formation be essentially concerted.

CHAPTER IV

HOLINS KILL WITHOUT WARNING – OR DO THEY

Introduction

Despite the exceptional diversity, conserved function, and genetic tractability of holin genes, the physical nature of the membrane lesion or the method in which this biological timer mechanism operates is unknown. During the end of the lytic cycle, at an allele-specific time, the holin triggers and disrupts the cytoplasmic membrane by the formation of non-specific lesions. This lytic process involves three essential proteins: holin, endolysin and the spanin complex for phages of a Gram-negative host (1, 99). Each of these proteins has a specific function to counteract the three barriers of the Gram-negative host in order to release phage progeny into the environment. The holin depolarizes the cell membrane potential by forming holes in the lipid bilayer. This halts macromolecular synthesis and also allows the endolysin to enter the periplasm. Endolysin has muralytic activity and is responsible for the degradation of peptidoglycan in the cell wall. Finally, the spanin complex completes lysis by disrupting the outer membrane. The entire process is finished within seconds after holin triggering (12).

The effects of the canonical S holin from phage lambda on the integrity of the host membrane has been studied using a tethered-cell flagellar rotation system (29). The results indicate that the integrity of the inner membrane remains fully intact until an allele-specific time, whereupon the holin triggers and induces hole formation; i.e., the S105 holin "kills without warning"(29). However, these results are only based on the

study of one specific holin, lambda S105. S105 is a class I holin composed of three TMDs arranged with N-out and C-in topology. Despite sequence diversity of holins, they share some universal functional features, which are currently categorized into three known classes of holin. However, no other classes of holin protein have been studied in detail, so it is unclear if allele-specific triggering time is a universal characteristic seen in all holin proteins. Among the paradigm phages, coliphage P2 has a lysis cassette that includes a putative holin gene, *Y*. This gene encodes a protein with predicted class I topology, adjacent to the gene for the endolysin, K. Ziermann et al. constructed a P2 Y_{am} mutant and showed that it exhibited a lysis defect suppressible by the addition of chloroform, indicating that Y functioned as the P2 holin (70). Lately, the Y holin functionality has been investigated in great detail (100).

Recently, an entire different class of holin-endolysin system has emerged (45, 47, 54). This new class, which is prototyped by phage 21, has a lysis cassette structure similar to that of lambda, containing S^{21}, R^{21}, Rz^{21}, and $Rz1^{21}$ genes downstream of the phage 21 late gene promoter. This alternative lysis system utilizes an endolysin that contains a novel secretory signal called a signal-anchor release (SAR) domain. SAR sequences initially act as signal-anchor domains and promote the integration of proteins into the cytoplasmic membrane (45, 54). The S^{21} gene encodes $S^{21}68$, the phage 21 holin that is referred to as pinholin as it forms small, pin-like holes rather than large nonspecific holes like the ones created by lambda holin. The pinholin protein is composed of 68 residues, and it accumulates in the cytoplasmic membrane during the beginning of the morphogenesis period, which begins at 10 minutes, until the cell lyses

at about 45 minutes. $S^{21}68$ pinholin has a class II holin topology with two TMDs arranged an N-in, C-in orientation. It has been studied extensively, with recent evidence indicating that topological dynamics of the protein affect its lysis behavior and that only the second TMD is necessary for pinholin functionality (48). The data suggest that $S^{21}68$ ΔTMD1 with only 44 remaining residues were sufficient for hole formation in the cytoplasmic membrane. In addition, biochemical studies have suggested that the $S^{21}68$ pinholin forms lesions only large enough for the passage of ions, and it does not allow the passage of cytoplasmic endolysin (45).

In this study, I address whether the Y holin and the $S^{21}68$ pinholin cause membrane depolarization with the same all-or-nothing fashion as S105, using the same tethered-cell assay previously employed for studying S105. The results are discussed in terms of a general model for lysis timing in which holin accumulates to a critical concentration that begins the formation of an oligomeric membrane lesion structure. Furthermore, phage shock protein (Psp) is the only known system for repairing or preventing damage to the cytoplasmic membrane (101). Based on extensive studies on phage-shock-protein or Psp in recent years, I hypothesized that functions in extra cytoplasmic stress response (101), may inhibit proton leakage in the cytoplasmic membrane by forming complexes to plug the holes. Psp proteins are conserved in many bacteria, including *E. coli*. In other bacterial genomes, there are two or more adjacent *psp* genes that together comprise a *psp* system (101). Dissipation of the membrane pmf is an inducing signal for the *psp* system. Therefore, it's possible that these Psp protein complexes are induced immediately after the cytoplasmic membrane is depolarized due

to the first pinhole formation. The Psp complex would serve to neutralize the pmf dissipation caused by the tiny hole formations by plugging the pinholes, thus prolonging the time of lysis. This would explain the phenomenon of "on and off" rotation of the cells mentioned before. I hypothesized that this phenomenon was not present in cells expressing the S and Y canonical holin, in which lysis occurred sharply because there was simply not enough Psp to block the large lesions formation by both functional holins.

Materials and methods

Strains and plasmids

The *E. coli* strain RY8797 is an *ara*$^{+}$ *lac*$^{+}$ *lacI*^{q} *eda*$^{+}$*ΔcheA* derivative of the strain RP437 (102), which is wild-type for motility. The plasmids used in this study were derivatives of pS105, which originally contained the lambda lysis gene cassette, *S105RRzRz1,* downstream of the pR', its native late promoter (32). Here, I replaced the lambda holin gene, *S105*, with *Y* gene from phage P2 or *S*^{21}*68* gene from phage 21. Several plasmids carrying different *Y* and *S*^{21}*68* alleles were used in this study (Table 5). The plasmids pY$_{M70I}$ and pY$_{L53M}$, containing the early lysis alleles Y_{M70I} and Y_{L53M}, were used for most tethering assays and lysis inductions. To construct plasmid pQ, the *Q* gene was PCR-amplified from λ DNA. The PCR fragment was digested with *Kpn*I and *Cla*I and cloned under control of the P$_{lac/ara-1}$ promoter in the *Kan*^{R} plasmid pZS-24 (103). Transcription of the *Q* gene from the *lac*/*ara* hybrid promoter is tightly repressed and can be induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) and arabinose.

Table 5. Strains, phages, plasmids and primers used in this study.

Strains	Genotypes and relevant features	Sources, references
MC4100	*E. coli* K-12 *F⁻ araD139* Δ(*argF-lac*)*U169 ΔfhuA rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR*	(73)
XL1-Blue	*E. coli* K-12 *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*^{q}*Z*$_{\Delta M15}$::Tn*10* (Tet^{r})]	Stratagene
Plasmids	**Genotypes and relevant features**	**Sources, references**
pS105	lambda lysis cassette under pR' on pBR322 backbone; S_{M1L} allele, encodes *S105* only	(74)
pSam7	Isogenic to pS105; Sam	This study
pY	Isogenic to pS105; *S105* replaced by *Y*; *(nt 6691-6997 of phage P2)*	This study
pY$_{M70I}$	Isogenic to pY; Y_{M70I}	This study
pY$_{L53M}$	Isogenic to pY; Y_{L53M}	This study
pY$_{\Delta TMD1}$	Isogenic to pY; Ser7 to Lys22 deleted	This study
pR'rbs S^{21}68	Isogenic to pS105; *S105* replaced by *S^{21}68*; and rbs is enhanced	This study
pR'rbs S^{21}68V43am	Isogenic to pR'rbs S^{21}68; S^{21}68V43am	This study
pR'rbs S^{21}68ΔTMD1	Isogenic to pR'rbs S^{21}68; TMD1 of S^{21}68 deleted	This study
pR'rbs S^{21}68 (R^{21}E/Q)	Isogenic to pR'rbs S^{21}68; (R^{21}E/Q)	This study
Primers	**Sequence**	**Sources, references**
pspA Cam For	5′CCTGAACGCAGAAATCAAGAGGACAACATTATGGGTATTTTTTCTCGCTTTGTGTAGGCTGGAGCTGCTTC3′	This study
pspA Cam Rev	5′CGGCGGCGCGTTCGGACGCCGCCTGGATGTTATTATTGATTGTCTTGCTTCATGGGAATTAGCCATGGTCC3′	This study
pspC Cam Rev	5′GCCTGTTGCCAGCGAGTATTCATAACTTTCCTCACAGTTGACGGAAACGGCATGGGAATTAGCCATGGTCC3′	This study
pspA For	5′CCTGAACGCAGAAATCAAGAGGAC3′	This study
pspC Rev	5′CCGGCCTGTTGCCAGCGAGTATTC3′	This study

Tethering Assay, microscopic analysis, and video recordings

The Δ*cheA* strain RY8797, in which flagella only rotate counterclockwise, was used in these studies. Overnight cultures of strain RY8797 pQ carrying pY or $pS^{21}68$ or an isogenic derivatives were grown in terrific broth (TB) medium at 30°C. For most of the tethering assays, early lysis constructs: Y_{L53M} or Y_{M70I} were used to shorten the induction time. Freshly grown overnight cultures were diluted 33-fold into fresh TB medium and grown at room temperature until the culture reached on A_{550} of 0.2. Motile cells were sheared and tethered to a glass coverslip through a flagellar filament-specific antibody, as described previously (104). The behavior of the tethered cells was observed in a flow chamber (105) by reverse phase-contrast microscopy. Before recording the baseline behavior of tethered cells digitally on the computer, cells were equilibrated for several minutes by the flow of pre-warmed TB medium at a flow rate of ~ 0.2 ml/min at 39°C. At $\boldsymbol{t} = 0$, expression of the lysis genes was induced by the flow of TB medium containing 1 mM IPTG and 0.5% arabinose. Every 5 min, the flow was stopped for 1 minute to record the rotation of individual cells. The flow was stopped permanently when lysis of the first cells was observed.

The rotational speed of individual cells at each time point was normalized to its speed during the previous interval of stopped flow (from 5 min earlier). This method was implemented due to the fact that many cells failed to spin for the entire experiment. The cells were either stuck, start spinning in the middle of the experiment, or detached from the glass coverslip. This normalization underestimates the overall decrease in speed over the course of the experiment, but it does not affect the comparison between

the control and experimental cells. The primary reason of the progressive slowing of rotation is the elongation of all tethered cells, induced and untreated, as an unavoidable consequence of doing the experiment under conditions that permit protein synthesis. At least twenty cells were examined at each time point, and the mean normalized velocity and standard deviations were calculated.

Gene knockout

The lambda red recombination system is a common approach for making targeted genetic changes and gene knockout. In this study, the red helper plasmid, pKD46, was transformed to the strain of interest, and transformants were selected by plating on ampicillin containing LB plates overnight at 30°C. An overnight culture of the strain containing pKD46 was grown at 30°C in LB media containing 100 µg/ml ampicillin. The overnight culture was sub-cultured 1:100 into fresh LB-ampicillin containing 0.2% arabinose and grown at 30°C for about three hours until the cells reached an A_{550} of 0.6. The cells were washed three times with ice-cold 10% glycerol, and then resuspended in 1/250 of the original volume after the third wash. Linear PCR products were gel-purified, digested with *Dpn*I, repurified, and suspended in elution buffer (10 Mm Tris, pH 8). Electroporation was performed using the MicroPulser™ electroporator from Bio-Rad, and 0.1 cm chambers according to the manufacturer's instructions. I used 50 µl cell aliquots and 100 ng to 500 ng of PCR product. 1 ml LB was added immediately to the shocked cells, then transferred to a culture tube and placed on a 37°C shaker for two hours. One-half was spread onto LB plate containing appropriate antibiotic to select Cam^R (30 µg/ml) or Kan^R (50 µg/ml) transformants. If

none grew within 24 hours, the reminder was spread after standing overnight at room temperature. After primary selection, candidate knock-out mutants were maintained on medium without an antibiotic. They were colony-purified once at 37°C non-selectively and then tested for ampicillin sensitivity to test for loss of the helper plasmid. Finally, mutant candidates were tested by PCR for the correct insertion-deletion event.

Results

Y holins kill without warning

Y begins to accumulate in the cytoplasmic membrane after 10 minutes from the beginning of the morphogenesis period, and continues to accumulate until lysis at about 35 minutes in the lambda-Y hybrid system (100). Biochemical and genetic studies had shown that Y belongs to the canonical holin class I, which form large lesions and allow the passage of ~ 500 kDa β-gal-fused endolysin (100). To examine the integrity of the cytoplasmic membrane and the pmf before Y triggering, I monitored the rotational speed of cells tethered by a single flagellum attached to a glass surface as a function of time after the induction of Y synthesis (106). The rotational velocity of the tethered cells is directly proportional to the pmf up to -160 mV, which covers the entire physiological range for *E. coli,* as it maintains its pmf between -160 to -200 mV when fully energized (107).

As mentioned in previous experiments (29), the lambda *S105* gene was replaced with the P2 *Y* gene in this study. The expression for the lysis genes were under the control of an inducible promoter. Flagella of motile *E. coli* cells harboring these two plasmids were sheared off by a blender and the cell was then tethered to a glass coverslip

by a single flagellum using anti-flagellar antibodies. *CheA* mutant cells were used in these experiments to ensure that the cells rotated exclusively in the counterclockwise direction. Cells tethered on the coverslip were equilibrated in the flow chamber, and the synthesis of Y was induced with IPTG and arabinose. Turbidity plots for the induced bulk cultures indicated that the lysis time for the wild-type *Y* gene under these conditions was ~ 50 minutes (Fig. 28; Table 6). Early lysis alleles Y_{M70I} and Y_{L53M}, holin triggered at 35 minutes after induction (Fig. 28; Table 6). Phase-contrast video microscopy revealed that individual cells maintained their rotational velocity after induction until rotation suddenly stopped less than ~ 30s before lysis (Fig. 29; Table 6), as previously observed for S105 (29). Cells expressing wild-type Y stopped rotating at a specific time and lysed within seconds. The same result was seen in cells with Y_{M70I} and Y_{L53M} early lysis alleles except that it occurred earlier. However, the rotational velocity was maintained with no observable cell death for the null allele (Fig. 29).

Pinholins kill with warning

Identical experiments were performed to examine cells expressing the $S^{21}68$ pinholin from bacteriophage 21. In this case, *S105* was replaced with the $S^{21}68$ gene using the plasmid system mentioned before, in which gene expression was controlled by a separate plasmid containing the *Q* gene. In the cells that synthesized the wild-type $S^{21}68$, rotation of the tethered cells came to a sudden halt ~ 85 minutes after induction (Fig. 30, Table 6). However, the cells began to rotate again for a period of time, then suddenly stopped again a few seconds later. This sudden stop and start pattern of

Table 6. Lysis times and interval between stop of rotation and actual cell lysis.

Plasmids[1]	**Time from induction until lysis in broth culture, (min)**[2]	**Time from induction until lysis for tethered cells, (min)**[3]	**Time from rotation stop until lysis (sec)**[4]
pSam7	No lysis	No lysis	No lysis
pS105	40	65±6	12±5
pY	45	70±4	10±7
pY_{M70I}	35	58±7	8±6
pY_{L53M}	33	55±5	10±4
$pY_{\Delta TMD1}$	No lysis	No lysis	No lysis
pR'rbs $S^{21}68$	40	85±28	138±125
pR'rbs $S^{21}68$V43am	No lysis	No lysis	No lysis
pR'rbs $S^{21}68\Delta$TMD1	45	92±32	124±116
pR'rbs $S^{21}68$ (R^{21}E/Q)	No lysis	No lysis	No lysis

[1] The motile strain RY8797pQ carrying pS105 or isogenic derivatives of this plasmid was used in all experiments.
[2] Cells were cultured in LB medium at 37°C. At an A_{550} of 0.2, expression of the lysis genes was induced by the addition of IPTG and arabinose. The turbidity of the culture was followed until the A_{550} dropped below 0.05. The time when the turbidity of the culture started to decline was defined as the lysis time.
[3] Cells were grown and induced as described in *Materials and Methods* and tethered to glass coverslips. The time until lysis for individual cells was determined by microscopic observation. The number of cells analyzed is shown in parentheses.
[4] Time between the halt of rotation and lysis of tethered cells.

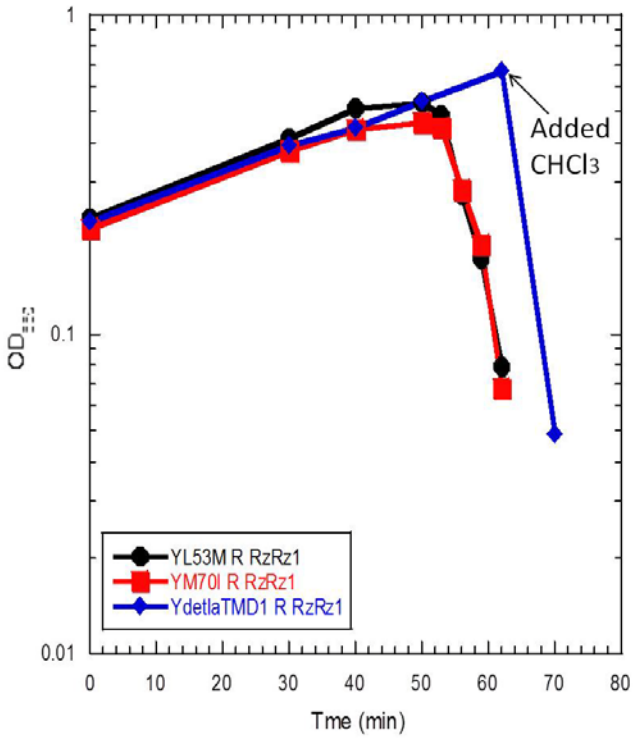

Fig. 28. Lysis curve of Y_{L53M} and Y_{M70I}. Y_{L53M} (black) and Y_{M70I} (red) lyse at 50 minutes. While $Y_{\Delta TMD1}$ (blue) is a non-functional mutant for which lysis does not occurred until the addition of $CHCl_3$.

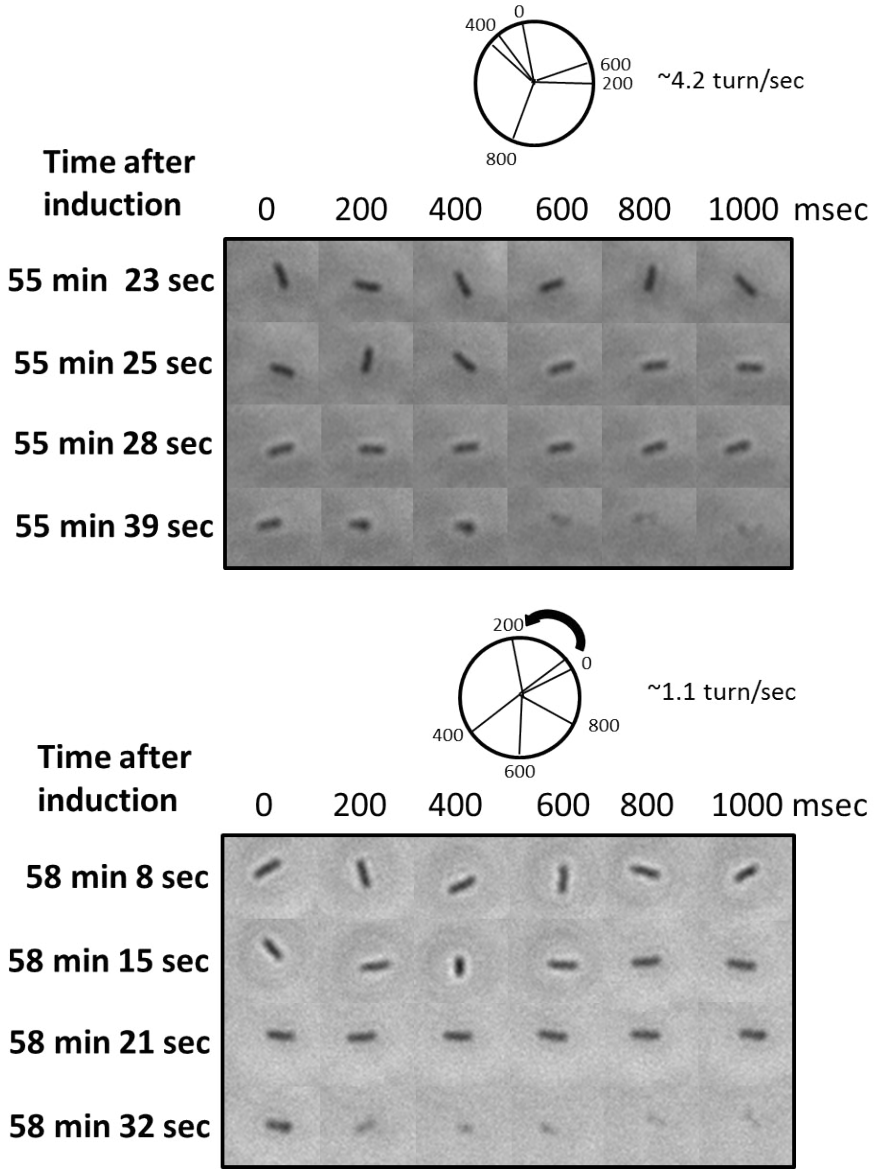

Fig. 29. Lysis of individual cells expressing *Y* alleles. Cells were tethered to coverslips as described in methods. At time = 0, expression of the lysis genes from the plasmids pY_{L53M} (top) and pY_{M70I} (bottom) were induced by IPTG and arabinose. Single frames were chosen from the recordings of two representative cells and are depicted here to illustrate the process of cell lysis. Starting from the time point indicated (Left), single frames were captured every 200 ms. After induction of the lysis genes, the tethered cells rotate at high and constant speed (first row). About 55 or 58 min after induction for each plasmid, rotation of the cell abruptly slows and stops completely within 1 to 5 sec (second row). Cell lysis, because of digestion of the cell wall by the λ R endolysin, occurs within several seconds after the sudden stop in rotation (third and fourth rows).

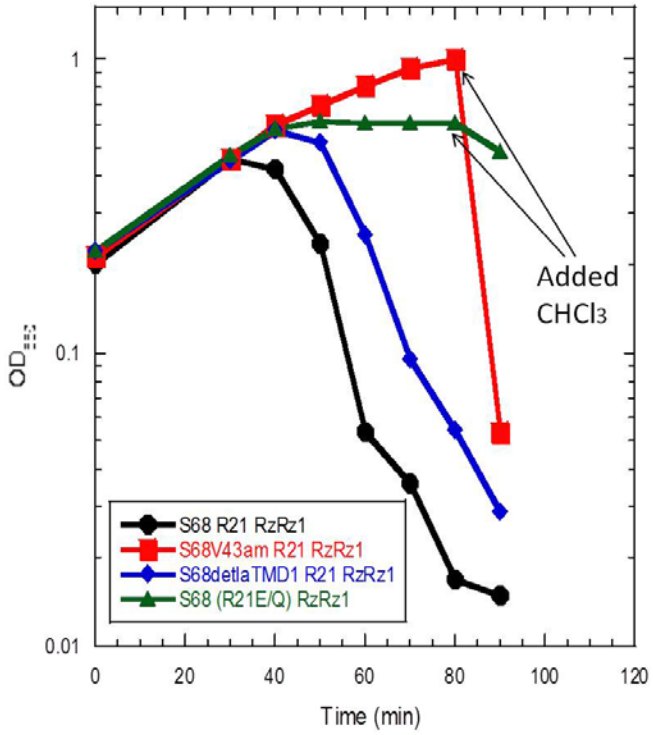

Fig. 30. Lysis curve of $S^{21}68$ and $S^{21}68_{\Delta TMD1}$. $S^{21}68$(black) and $S^{21}68_{\Delta TMD1}$ (blue) lyse at 40 minutes. While $S^{21}68_{V43am}$ (red) is a non-functional mutant for which lysis does not occurred until the addition of $CHCl_3$. R^{21}E/Q (green) is a non-functional endolysin.

rotation repeated itself numerous times until a final stop was observed, and the cell lysed (Fig. 31). The rotational stop-restart phenomenon lasted from 10 to 30 minute on average for a given cell. Furthermore, two methods were used to analyze the spinning patterns of the cells tethered to coverslips. The first method used to evaluate the cells took into account the number of spins (complete rotation) a given cell executed in a two minute interval.

This study focused on three constructs: wild-type ($S^{21}68$ R^{21}), functional mutant ($S^{21}68$ ΔTMD1 R^{21}), and non-lysis mutant ($S^{21}68$ R^{21-}). The results show that a common trend was observed in all three of the constructs – the duration of each stop lasted longer as time increased (Fig. 32). The second method consisted of analyzing the number of stops a cell had within a two minute interval. Twenty cells were chosen from each construct. Overall, no significant difference in p valve was seen in the average number of stops between the constructs (Fig. 33). The number of stops before lysis has a board range (Table 6). Cells expressing the wild-type *$S^{21}68$* allele and cells expressing the truncated allele, *$S^{21}68_{\Delta TMD1}$,* behaved identically with periodic rotational stops from the time of induction until lysis (Fig. 31; Table 6). As expected, no lysis or periodic rotational stop was observed for the null allele of *$S^{21}68$* (Table 6). The presence of the SAR endolysin R^{21} had no effect on the rotational stop-restart pattern except that no lysis was observed after the final halt (Table 6). The repeated stop-restart pattern observed for pinholin system suggested that instead of a concerted "all or nothing" triggering event, the population of pinholin molecules have a different effect on the proton motive force.

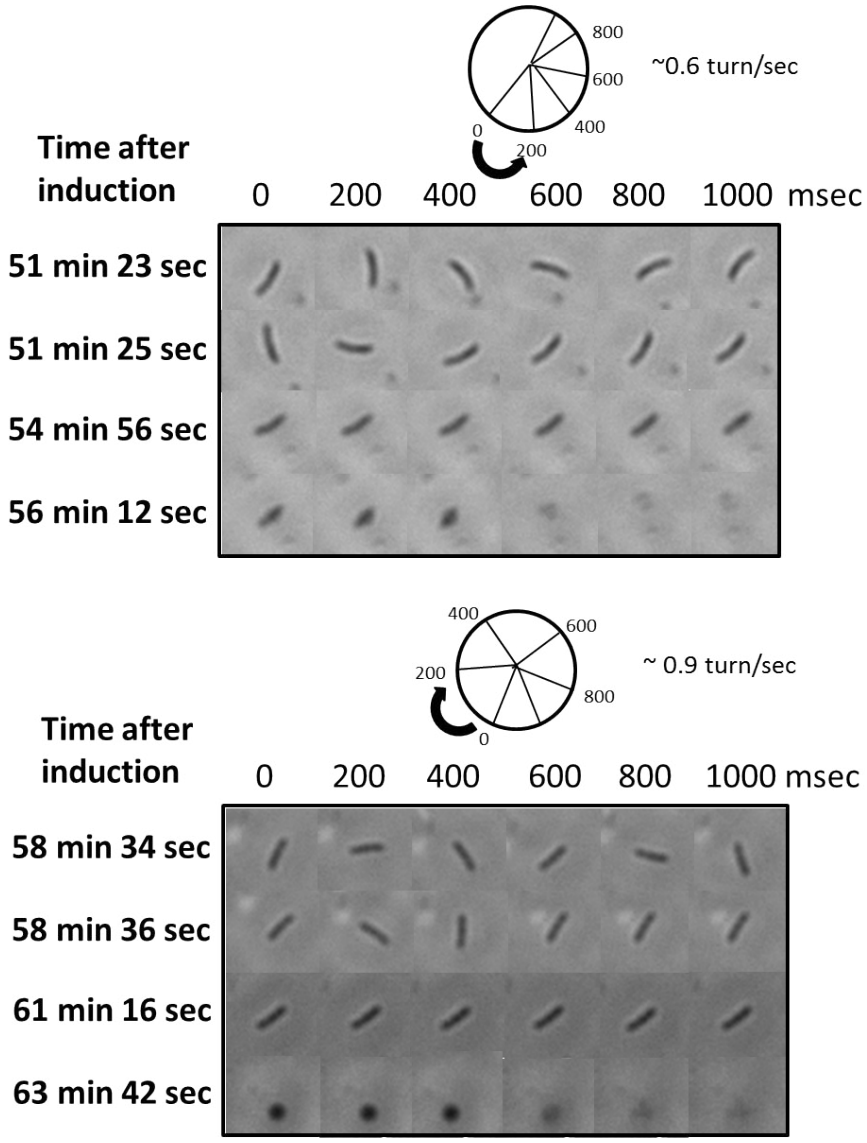

Fig. 31. Lysis of individual cells expressing *S^{21}68* alleles. Cells were tethered to coverslips as described in materials and methods. At time = 0, expression of the lysis genes from the plasmids pS^{21}68 (top) and pS^{21}68$_{\Delta TMD1}$ (bottom) were induced by IPTG and arabinose. Single frames were chosen from the recordings of two representative cells and are depicted here to illustrate the process of cell lysis. Starting from the time point indicated (left), single frames were captured every 200 ms. After induction of the lysis genes, the tethered cells rotate at high and constant speed (first row). About 51 – 58 min after induction for each plasmid, rotation of the cell abruptly slows and stops, then back to rotation repeatedly, until it finally stop completely (second row). Cell lysis, because of digestion of the cell wall by the phage R^{21} SAR endolysin, occurs within several minutes after the final stop in rotation (third and fourth rows).

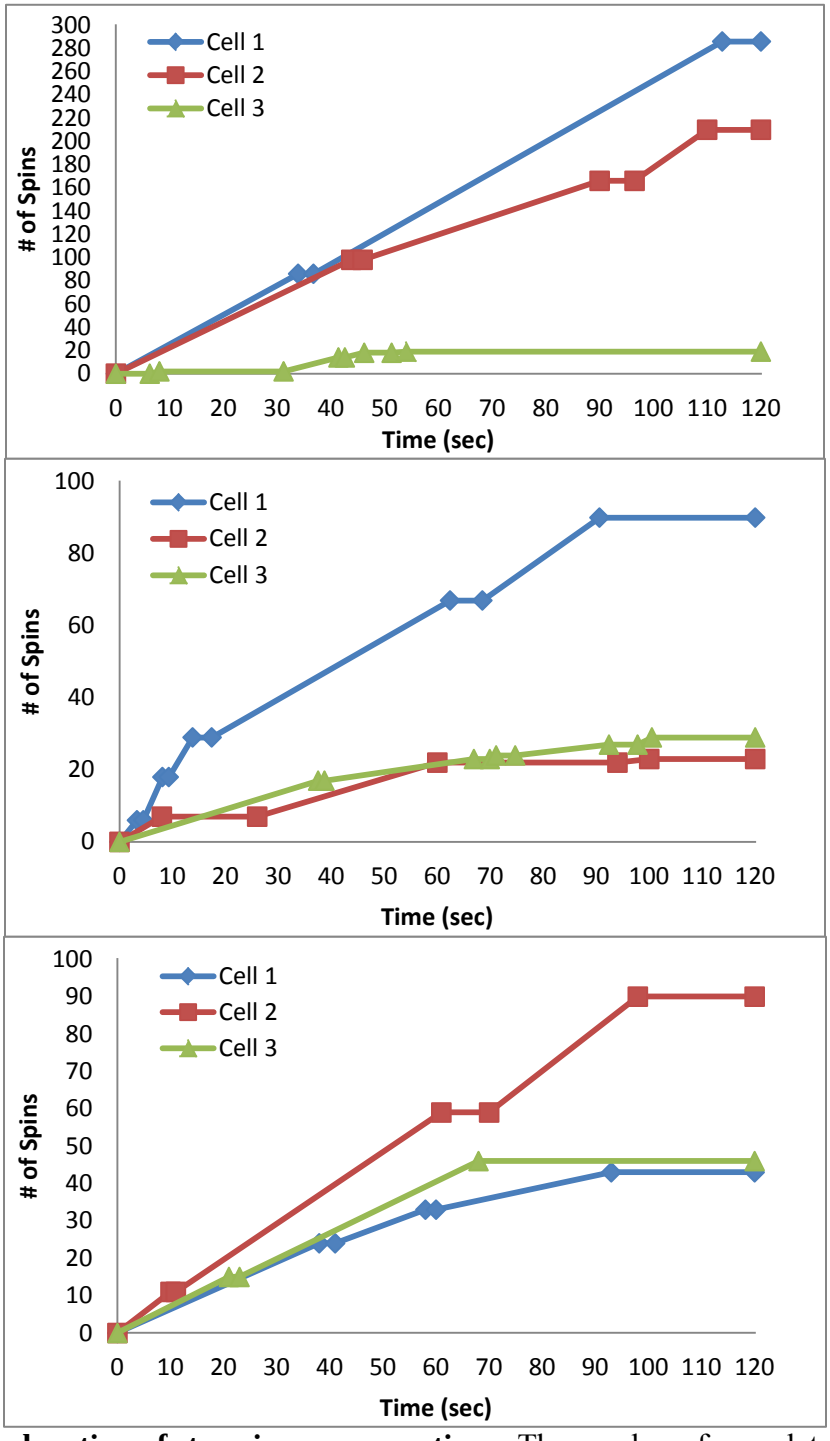

Fig. 32. The duration of stops increase overtime. The number of complete rotations for three representative cells of each construct. $S^{21}68\ R^{21}$ (top); $S^{21}68\ \Delta TMD1\ R^{21}$ (middle); $S^{21}68\ R^{21-}$ (bottom). The duration of stops get longer as the cell spins for a longer period of time.

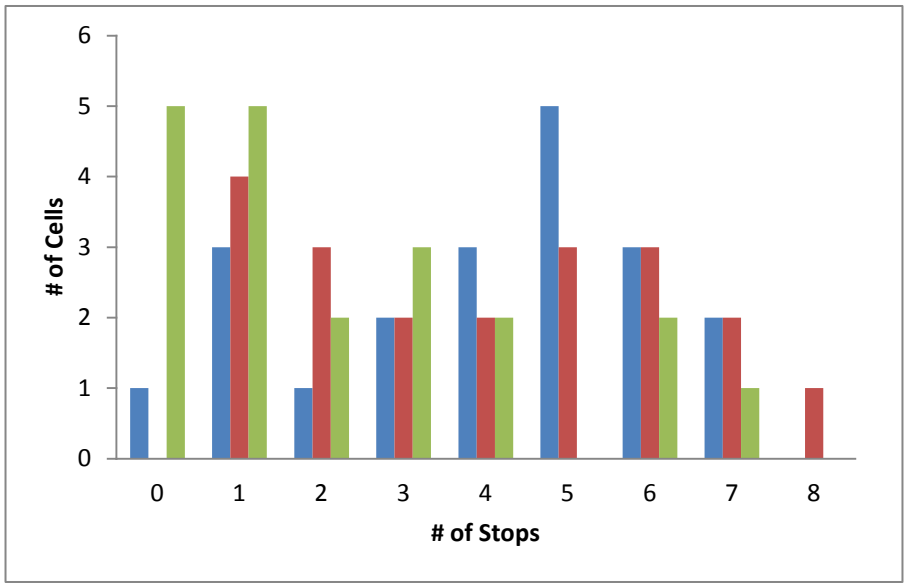

Fig. 33. Number of stops are evenly distributed. Twenty cells were chosen from each construct and the number of stops in a two-minute interval was recorded. $S^{21}68\ R^{21}$ (blue); $S^{21}68\ \Delta TMD1\ R^{21}$ (red); $S^{21}68\ R^{21-}$(green).

Phage-shock-protein independence

If the pinholin triggering pathway is marked by successive waves of membrane depolarization followed by re-establishment of the pmf, the simplest notion is that the cell is mounting a response to the membrane damage. To test whether Psp is responsible for the observed stop-restart phenomenon from cells expressing pinholin, I knocked out the *pspA, pspB,* and *pspC* genes. However, no difference in terms of lysis kinetics was observed between strains with *psp* gene knockout and the parental. The data shows that *pspA, pspB*, and *pspC* knockouts slowed lysis, rather than accelerated it (Fig. 34). Unfortunately, I was not able to test for the presence of Psp protein in these cells due to the lack of antibodies specific to Psp.

Discussion

This study shows that two holins unrelated to lambda S105, P2 Y and the phage 21 $S^{21}68$, accumulate throughout the morphogenesis period without effect on the pmf of the host membrane, exactly as observed for the paradigm S105 holin. I posit that this is a universal feature for holins, presumably a reflection of the imperative that the lysis system should not adversely affect macromolecular synthesis, and thus the yield of progeny virions, during the vegetative cycle (29). Evidence shown in this study implicates the mechanism of raft formation at the molecular level. The current model proposes that only the rafts are inherently ion leaky, so that the critical concentration model controls the transition from a normal membrane to a membrane with ion leaks. Moreover, cells expressing Y holin completely stop rotating at a specific

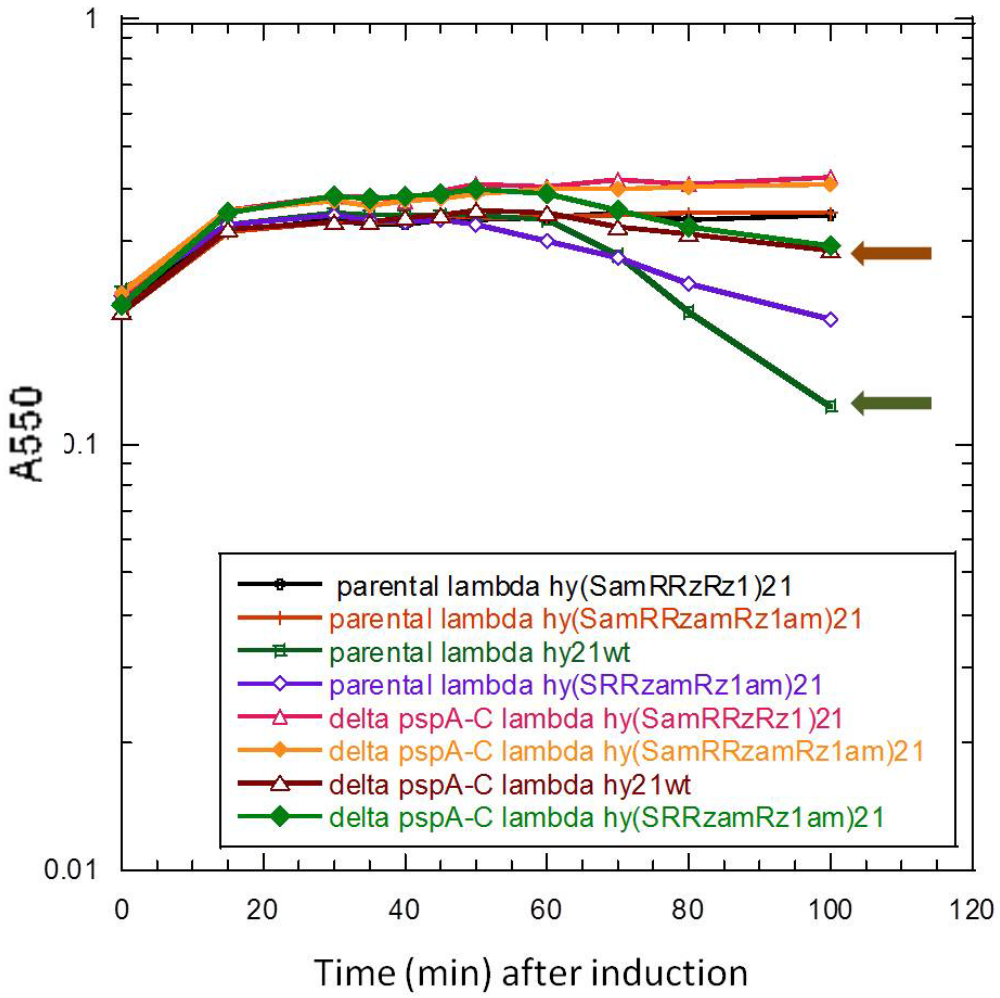

Fig. 34. Parental vs ΔpspA-C phage 21 lysogen. There are no significant differences between the ΔpspA-C strain (brown) and the parental strain (green) in terms of lysis kinetics. The ΔpspA-C strain (brown) slowed lysis, rather than accelerated it.

triggering time, suggesting large hole formation with the notion that marco hole formation, much larger than necessary, is a consequence of the timing.

The ability for holins to accumulate in the cytoplasmic membrane is harmless in terms of the energy state of the cell. The rapid trigger of lysis is a common feature seen in class I holins that form large holes at an allele-specific time. Currently, the simplest working model is composed of two-dimensional aggregates, formed by holin molecules by means of helical packing, that accumulate until the amount of holin molecules reaches a critical level in the lipid bilayer. The mechanical membrane integrity is compromised as the aggregation of holins expands leading to local depolarization. Once proton leakage occurs and the pmf collapses, it then elicits a chain reaction leading to global triggering of holins. As the result, massive membrane lesions derive from large rafts.

In contrast to P2 Y, the triggering pathway appears to be completely different for the S^{21} pinholin. I observed the stop-restart phenomenon for cells expressing pinholin after triggering whether than lysis within seconds like Y and S105. The results shown here indicate no significant difference in p value between the spinning patterns between the three constructs: wild-type ($S^{21}68\ R^{21}$), functional mutant ($S^{21}68\ \Delta TMD1\ R^{21}$), and non-lysis mutant ($S^{21}68\ R^{21-}$). However, within each set there is large variation in the number of stops within a two-minute interval (Fig. 33). Another trend that held true for these three constructs tested was that at longer times, the duration of each stop became longer until the flagellum stop completely (Fig. 32). I hypothesize that perhaps the host has some sort of mechanism to compensate for the small amount of pinholes formed. As

time goes on, an increased number of pinholes form in the membrane, and the host can no longer overcome the decreased proton motive force. This increase in the number of pinholes causes the duration of each stop to become longer as the host attempts to contain ion leakage. The stop-restart phenomenon could be due to a host response, but it would be with pre-existing protein since little protein synthesis can occur when pmf is zero (108). This phenomenon is not seen in the host expressing Y holin because I believe there are only a few large hole forms. The cell is not able to overcome this large disruption to its proton motive force and spinning ceases abruptly and permanently.

I also tested the hypothesis that pinholes may be blocked by phage shock protein complexes as an acute membrane damage response. Although, the Psp system has not been shown to be responsible for the chronic membrane damage in this study, but the results are inconclusive at the moment due to lack of sufficient antibodies. Further experiments are needed to determine if phage-shock-proteins are responsible for the stop -restart phenomenon from cells expressing the pinholin.

Together, these data suggest that there is a universal pathway for holin triggering, and that the holin aggregates may cause proton leakage that lead to lethal function of holin or pinholin. It is possible that the fundamental difference between pinholin and canonical holins can change the progeny yield over time and lead to selective advantage for specific phages in the population. In the future, it would be interesting to examine two unrelated phages, one that uses holin and one that uses pinholin, in levels of competitiveness in a given population and to monitor specific phage yield over time.

CHAPTER V

NEW CLASSES OF HOLIN, PHAGE MU

Introduction

All dsDNA phages of eubacteria and euryarchaea cause lysis of the host at the end of the vegetative cycle (13). Lysis requires the degradation of the peptidoglycan. To date, every phage in which the lytic cycle has been subjected to physiological and analysis has been shown to use the holin-endolysin pathway, in which the holin determines the timing of lysis and causes permeabilization of the cytoplasmic membrane, leading to endolysin attack on the cell wall. Endolysin genes are relatively easy to identify bioinformatically because they fall into a limited number of enzymatic classes: glycosidases, transglycosylases, amidases and endopeptidases. These classes in general have signature motifs that define the catalytic residues and also known motifs or domains that are involved with murein binding (109). Also, endolysin genes are often easy to clone because, with the exception of the relatively rare SAR endolysins, they are not harmful to the cell when expressed constitutively. This is an important consideration since the commonly used cloning plasmids are both high copy number and poorly repressed (103). However, holins are much more difficult to identify, for several reasons. First, they are almost universally toxic when cloned in multicopy plasmids and are thus very difficult to clone and characterize functionally. Second, holins are extremely diverse in terms of both predicted membrane topology and primary structure, as summarized in the most recent comprehensive reviews (5, 12).

Only a few holin genes have been characterized by any type of experimentation, mostly by physiology and genetics; these genes were described in the most recent survey review in 2000 and 2002 (5, 12). Others were found as homologs of these experimentally characterized holins. In a few cases, the holin gene was predicted by positional arguments, the easiest being clustered with other lysis genes, like the endolysin and the spanins (note that spanins were not defined at the time of the last review; instead, the spanin genes were designated as Rz-R1 equivalents or auxiliary lysis genes). Using these criteria, two topological classes, each having multiple unrelated sequence families, were identified in the last major review (5): class I, with three TMDs, N-out, C-in, and class II, with two TMDs, N-in, C-in. A third class, class III, consisted of a single family of holins, homologs of gp*t* of T4, found in the large myophage class of T-even like phages and in T5-like phages. Gp*t* has a single TMD, flanked by a cytoplasmic domain of 34 amino acids and a large periplasmic domain of 163 amino acids.

It is even more challenging to identify antiholin candidates. There are five experimentally confirmed cases where the holin and antiholin are different translational products from the same cistron (12), the most notable being the holin S105 and antiholin S107 from lambda *S* (17, 23). Thus, the presence of a potential dual start motif has been used as evidence for annotating genes encoding small membrane proteins as holin/antiholin genes (12). For separate antiholin genes, there are only three for which there is any experimental evidence for antiholin character: RI from T4 (110), LysA from P2 (100), and LydB from P1 (13). It should be noted that these three proteins have

completely different topologies. RI has a single TMD, with its C-terminus out, whereas LysA has four TMDs with both N and C termini in and LydB has a single TMD, with C-terminus in. Thus, the only thing in common is each has at least one TMD.

In the decade since the last reviews, holin genes have been being annotated based on the topological patterns of the three established holin classes. Ultimately, as genomic information became available, holin genes were thus identified for all, except one, of the paradigm experimental phages, which, for *E. coli*, were: lambda, the "Seven Dwarves" of the T phages (T1 to T7), P1, P2, P22 and Mu (12, 13). However, for one of these phages, phage Mu, no holin gene was identified. Most interest in Mu has been in its role as the most active transposon known in biology, allowing intensive biochemical investigation of the gene transposition process (111). This brought Mu considerable prominence in molecular biology; indeed, besides lambda and T4, Mu is the only phage that has had a dedicated book published about it (112). However, despite intensive genetic and physiological analysis and the availability of the Mu genome sequence, the only lysis gene identified was the endolysin, gene *gp22* or *lys*, based on genetic and physiological evidence (113). Moreover, none of the 56 proteins encoded by Mu have a predicted topology consistent with one of the three canonical holin topologies (see below). This suggested that a holin of a novel topology was encoded by Mu. Here I report a genetic and molecular analysis to identify the holin and antiholin genes in Mu and the exploitation of the results in terms of a bioinformatic assessment of the diversity of holin and antiholin topologies in the current database.

Materials and methods

Bacterial strains, culture growth, and reagents

Experiments to test lysis-defective phenotypes were performed with the temperature sensitive lysogen MC4100 Mu(c^{ts}). All the phage, strains and plasmids used in this study are summarized in Table 7. The conditions for the growth of cultures and the monitoring of lysis kinetics have been described (31). Cultures were grown in LB medium supplemented with ampicillin (Amp; 100 μg/ml) or chloramphenicol (Cam; 10 μg/ml), as appropriate. Construction and phenotypic testing of Mu Δ(*gp19 gp20*) and other mutants were performed in this study. 1 mM L-arabinose (Sigma) was used for induction. Oligonucleotides were ordered from IDT (Coralville, IA). Enzymes were from New England Biolabs unless indicated otherwise. *Taq* polymerase was used in all PCR tests. *Taq* and *Pfu* (Stratagene) polymerases were mixed 10:1 and used per *Taq* instructions to generate DNAs for cloning and mutagenesis. Qiagen kits (Hilden, Germany) were used to isolate plasmid DNAs, gel-purify fragments, or purify PCR products.

Construction and phenotypic analysis of Mu Δ(gp19 gp20)

For gene disruption, transformants carrying a Red helper plasmid (pkD46) were grown in 25-ml LB cultures with ampicillin and L-arabinose at 30°C to an OD_{550} of ~ 0.6 and then made electrocompetent by concentrating 100-fold and washing three times with ice-cold 10% glycerol. PCR products were gel-purified, digested with *Dpn*I, repurified, and resuspended in elution buffer (10 mM Tris, pH 8.0). Electroporation was done by using a Cell-Porator (Bio-Rad) with a voltage booster and 0.1 cm chambers

Table 7. Phage, strains and plasmids used in this study.

Phage	**Genotypes and relevant features**	**Sources, references**
Mu (c^{ts})	*araD139 D(argF-lac)U169 rspL150(Smr) relA1 flbB5301 pstF25 deoC1*	Laboratory stock
$\lambda Q^{21} \Delta(SRRzRz1)^{21}$	cI857 hy($Q^{21}(SRRzRz1)^{21}::Cam^{R}$) *bor::kan*^{r}	This study
Strains	**Genotypes and relevant features**	**Sources, references**
MC4100	*E. coli* K-12 F^{-} *araD139* Δ(*argF-lac*)*U169 ΔfhuA rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR*	(114)
XL1-Blue	*E. coli* K-12 *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI*^{q}*Z*$_{\Delta M15}$*::Tn10 (Tet*^{r}*)]*	Stratagene
MDS12	*MG1655* with 12 deletions, totaling 376,180 nt, including cryptic prophages	(40)
Plasmids	**Genotypes and relevant features**	**Sources, references**
pS^{21}68	Phage 21 lysis cassette under its native promoter pR'^{21} on pBR322 backbone; encodes S68 only	(48)
pMu19	Isogenic to pS^{21}68; *S68* replaced by *gp19* from Mu; (nt 9440-9825 of phage Mu)	This study
pMu20	Isogenic to pS^{21}68; *S68* replaced by *gp20* from Mu; (nt 9815-9955 of phage Mu)	This study
pKD46	*Δ(araD-araB)567 ΔlacZ4787*(::rrnB-3) *λ*$^{-}$ *rph-1 Δ(rhaD-rhaB)568 hsdR514*	(115)
pKD3	*Δ(araD-araB)567 ΔlacZ4787*(::rrnB-3) *Δ(phoB-phoR)580 λ*$^{-}$ *galU95 ΔuidA3::pir*$^{+}$ *recA1 endA9*(del-ins)*::frt rph-1 Δ(rhaD-rhaB)568 hsdR514*	(115)
Primers	**Sequence**	**Sources, references**
c1	5'CTGCTTTACAGGGAGAAACATTATG3'	This study
c2	5'GCTGCATATTTTCCCCCGGTTAAC3'	This study
c3	5'GTGTAGGCTGGAGCTGCTTC3'	This study

according to the manufacturer's instructions by using 100 µl of cells and 10–100 ng of PCR product. Shocked cells were resuspended in 1 ml LB, incubated 1 h at 30°C, and then one-half was spread onto agar to select Cam^R transformants. If no resistant colonies grew within 24 hours, the remainder was spread after standing overnight at room temperature. After primary selection, mutants were maintained on medium without an antibiotic. PCRs were used to confirm that all mutants have the correct *Cam* insertion.

Bioinformatic identification of lambda holin homologues in phage genomes

Sequence comparison searches were carried out against the non-redundant Reference Sequence (RefSeq) database. Starting with a single query sequence, PSI-BLAST (116) was iterated until convergence, using the BLOSUM62 scoring matrix and an *E*-value cutoff of 7×10^{-06}. The lowest scoring sequences from these groups were used as new queries in subsequence rounds of PSI-BLAST. Searches were iterated until no additional hits were found. For annotations based on topological considerations and genome location, the genomic sequences and annotated proteins of the indicated phage were obtained from Genbank. Potential TMD-containing proteins were identified by screening all the annotated open reading frames, irrespective of start codons, with TMHMM (117), TMPred (62) and charged residues. The organization of the holin coding region was determined by manual inspection using the Artemis genome browser (118).

Protein sequence analysis

Protein sequences were aligned with the CLC software suite (CLC bio). Charge distributions and amino acid composition analysis were obtained with phage gene analysis programs at CPT protal (https://cpt.tamu.edu/cpt-software/portal/charges.pl). Holin candidates were "PSI-BLASTed" using an *E*-value cutoff of 0.1. Furthermore, holin candidates were screened for potential TMD(s) using TMHMM (http://www.cbs.dtu.dk/services/TMHMM/), TMpred (http://www.ch.embnet.org/software/TMPRED_form.html), and manual inspection of the amino acid charge distribution.

Complementation assay

The strains and plasmids used in this study are listed and described in Table 7. The plasmids pMu19 and pMu20 are derivatives of the plasmid $pS^{21}68$ (49), a medium-copy-number plasmid that has the entire phage 21 lysis cassette under its native late promoter. The median growth conditions, and the thermal induction of the phage 21 lysis genes from prophages have been described previously (49). Lysis genes from phage 21 were used instead because the *lys* gene in Mu and SAR endolysin in phage 21 were highly homologous. In this study, plasmids were named pMu19 for hybrid lysis cassette $gp19(RRzRz1)^{21}$, and pMu20 for hybrid lysis cassette $gp20(RRzRz1)^{21}$, which has the phage 21 *S68* gene and its Shine-Dalgamo sequence replaced by *gp19* or Mu *gp20* gene from Mu, and their native Shine-Dalgamo sequences respectively. pMu19 or pMu20 was transformed into the lysogenic strain, MDS12Δ*fhuA*($\lambda Q^{21}\Delta(SRRzRz1)^{21}$, and they were used for phenotypic analysis. Cultures were grown in standard Luria–Bertani (LB)

media supplemented with antibiotics where needed: kanamycin(40 μg/ml), and chloramphenicol (10 μg/ml). Plasmids carried by thermo-inducible lysogens were aerating at 42°C for 15 min, followed by aeration at 37°C until the time specified. Lysis profiles were obtained by monitoring A_{550} after thermal inductions, as described previously (30, 119) .

Results

Identifying the lysis gene cassette of phage Mu

The Refseq genome entry for the paradigm coliphage Mu has 55 annotated CDSs, many of which had been identified genetically before the complete genome sequence was available (112). One of these, the *lys* gene, or *gp22*, is located in a long late gene transcriptional unit (Fig. 35A). Inspection revealed that Lys has a "true lysozyme" catalytic triad (Ex8Dx5T) (Fig. 35B), as originally defined in the T4 lysozyme (120). It also has a predicted N-terminal TMD enriched in relatively hydrophilic residues, indicating that it is a SAR endolysin (Fig. 35B). With the location of the endolysin gene, the spanin genes could be identified with confidence. Inspection revealed that gene *gp23* encodes a class II integral membrane protein, and overlapping it is a small reading frame, designated *gp23a*, encoding an OM lipoprotein. Thus, this gene pair is unambiguously identified as the i-spanin/o-spanin gene pair for phage Mu. The finding of the spanin genes adjacent to the genetically confirmed endolysin gene strongly indicated that the Mu holin gene should be nearby as part of a lysis cassette (Fig. 35A). Next, the identification of the holin gene was addressed. Besides the spanin CDSs, only three Mu CDSs were predicted to have at least one TMD, as judged by

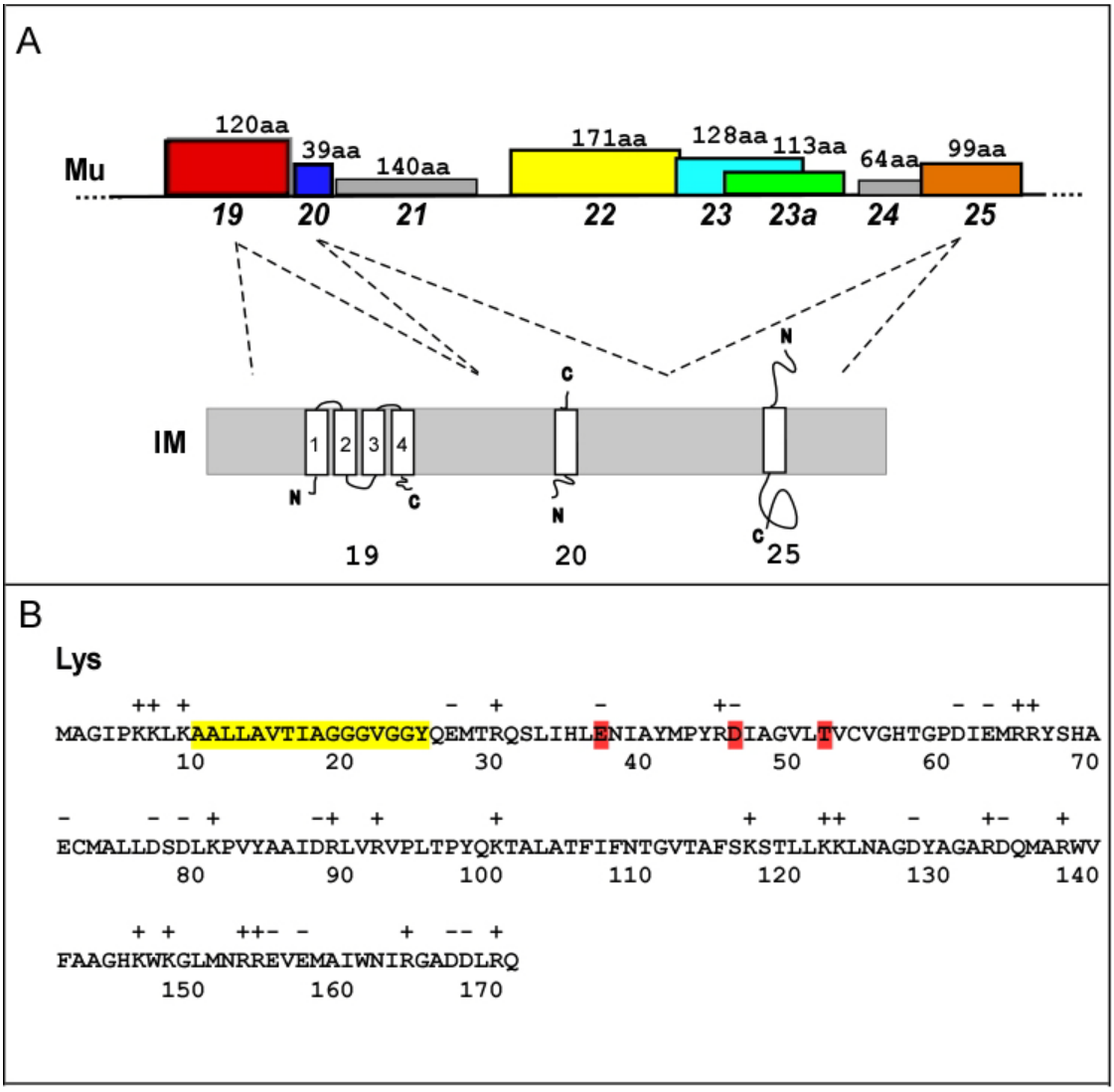

Fig. 35. Lysis cassette of Mu, the holin and the antiholin topologies. (A) Lysis cassette of Mu and topologies. Gene *gp19* (red) encodes a 120 amino acid product that possesses four putative TMDs with cytoplasmic N-terminus and periplasmic C-terminus topology. Gene *gp20* (blue) encodes a 39 amino acid product that possesses one TMD with an N-terminal cytoplasmic domain, and a C-terminal periplasmic domain; *gp22* or *lys* (yellow); RzRz1 spanins (blue and green), respectively; and *gp25* (orange). (B) Primary structure of Lys (gp*22*). Relatively hydrophilic residues (yellow); catalytic triad (red).

TMHMM analysis: gp*19* (120 aa), gp20 (39 aa), and gp*25* (99 aa) (Fig. 35A). However, none of the three proteins have predicted membrane topology consistent with a canonical holin. Both gp*20* and gp*25* are type II integral membrane proteins, with a N-terminal TMD and the rest of the protein predicted to be in the cytoplasm. Gp*19* has four TMDs, with both N and C termini in the cytoplasm. Since holins have at least one TMD, this analysis demonstrated that the Mu holin would be of unprecedented topology.

To address which of these three genes encode a holin, *gp19*, *gp20*, and *gp25* were knocked out using a temperature inducible Mu prophage, and their lysis phenotype *in vivo* was then examined (Fig. 36A). The control for these inductions was the Δ*lys* prophage. As expected for the absence of the endolysin, inductions of the deletion construct did not result in lysis. Instead, the culture mass increased after induction until suddenly reaching a plateau, indicating the holin had triggered. Unexpectedly, however, the same phenotype was observed with Δ*gp25* (data not shown). The molecular basis for this is unknown. The simplest possibility is that Lys, which has been assigned as SAR endolysin, requires gp*25* for its secretion or for its release after holin triggering. In any case, clearly gp*25* does not encode a holin, because the sudden halt in culture growth showed that holin triggering was occurring in both the *lys* and *gp25* mutants. Eliminating *gp25* left only gp*19* and gp*20* as Mu membrane proteins that would be candidates for holin function. The SAR character of the Mu endolysin, Lys, makes the identification of the Mu holin more challenging, mainly because even in the absence of holin triggering, SAR endolysins eventually cause lysis, because SAR endolysins undergo a low level of spontaneous membrane release and activation of the

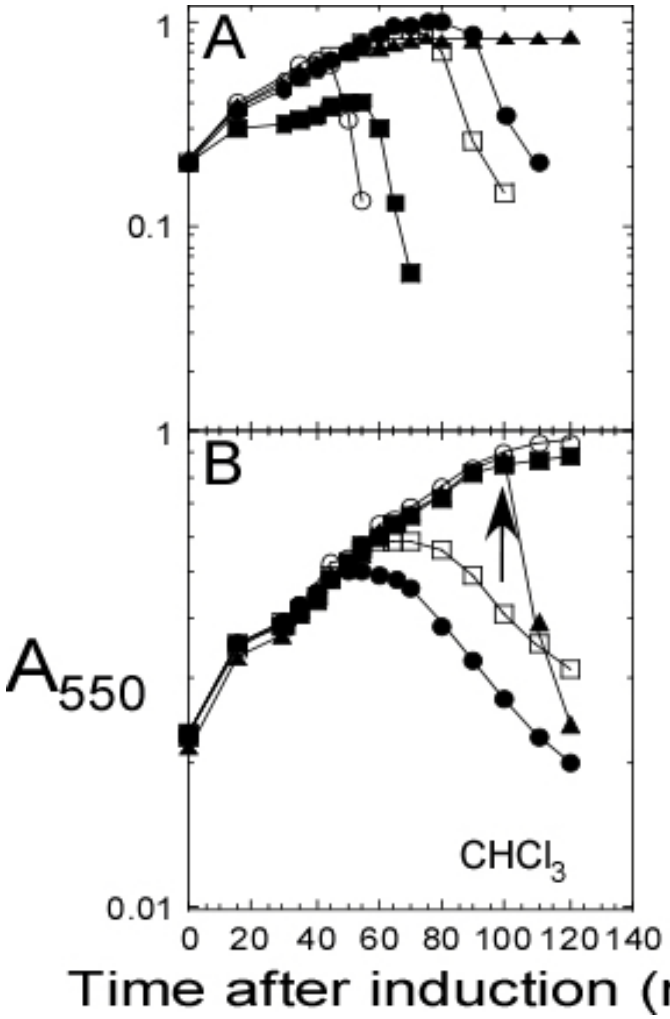

Fig. 36. Mu *gp19* and *gp20* lysis phenotype. (A) Lytic properties of a Mu genes. Inductions: Mu (c^{ts}) (○); *lys::Cam* (▲); *gp19-20::Cam* (□); *gp19::Cam* (●); *gp20::Cam* (■). (B) Lysis for Mu *gp19* but not *gp20*. Inductions: $pS^{21}68$ (●); $pS^{21}68am$ (■); pMu19 (□); pMu20 (○); pMu20, with $CHCl_3$ added at 100 min (▲) indicated by the arrow.

muralytic activity (45). In the characterized SAR endolysin systems, this holin-independent lysis was found to be delayed and less saltatory compared to the cognate holin-triggered lysis (12, 45, 54). As expected, induction of the Δ(*19-20*) double deletion showed this type of delayed lysis phenotype, supporting the idea that one of these genes encodes the holin (Fig. 36A).

To discern which gene had holin function, each was substituted for the S^{21} pinholin gene in a plasmid carrying the phage 21 lysis cassette. Inserts in this plasmid are expressed in the proper temporal and transcriptional context by inducing a lysis-defective prophage, $(\lambda Q^{21} \Delta(SRRzRz1)^{21}$, resulting *in trans*-activation of the plasmid-borne phage 21 late promoter, pR', by the phage-encoded late gene activator, Q^{21}. The *gp19* construct supported lysis with the same general profile as the native S^{21} construct, although delayed by ~ 10 min (Fig. 36B), whereas the *gp20* construct exhibited no lysis. The absence of lysis with the latter construct was not due to lack of sufficient SAR endolysin expression, as addition of $CHCl_3$ immediately caused lysis of the induced culture. Thus, *gp19* was assigned the function of the Mu holin. This made gp*20* the likely Mu antiholin. Some evidence for this idea was gleaned from inductions of the Mu prophages with single deletions, *Δ19* and *Δ20* (Fig. 36A). Lysis with *Δ19* was even more retarded than with the double deletion, whereas with the *Δ20* deletion the timing of lysis was shortened by more than 30 min. In addition, growth of the culture after induction of *Δ20* was significantly retarded, suggesting that the accumulation of the gp*19* holin in the membrane of the induced culture in the absence of the antiholin had premature toxic consequences. Similar premature lethality has been observed with

mutants of the P1 antiholin, LydB, to the extent that *lydB* nonsense mutants fail to form plaques (13).

Taken together these results indicated that gp*19* and gp*20* were the Mu holin and antiholin, respectively. Gp*19* has four TMDs, with the predicted topology of N-in, C-in (Fig. 35A) and thus is the first member of a new topological class of holins, designated as class IV. Gp*20* is the fourth protein, after T4 RI (N-terminal TMD; C-out), P1 LydB (N-terminal TMD, C-in), and P2 LysA (four TMDs, N-in, C-in), to be assigned an exclusively antiholin function. It is predicted to have a single, C-terminal TMD; the distribution of charged residues makes it very likely that the N-terminus is cytoplasmic (Fig. 35A). Thus, each of the four dedicated antiholins, RI, LydB, LysA and gp*20*, has a different topology.

New holin topologies in the phage genome database

The discovery that phage Mu encoded a holin with novel topology was a disruptive finding. Since very few experiments have been done to identify holins by function, most of the holins that have been annotated in the database were so annotated because of having predicted membrane topology corresponding to the three established topology classes: class I (three TMDs, N-out) class II (two TMDs, N-in) and class III, consisting of homologs of T4 gp*t*, which has one TMD and a large C-terminal periplasmic domain (12). This prompted a systematic search for holin genes in the RefSeq database of all complete genomes of phages of Gram-negative hosts. Of 406 genomes, 32 had homologs of T4-like phages, extending the single family of holins that defined class III; these were not examined further. The remaining 374 genomes were

analyzed manually, according to the following "proximity" strategy, relying on the general tendency of phage genomes to cluster lysis genes. Each genome was scanned for genes encoding small integral membrane proteins, defined as proteins of between 30 and 300 aa, with at least a single TMD. If there were such genes within 10 genes of an endolysin, these genes were selected as candidate holins/antiholins. If there was only one such gene within those boundaries, it was designated as the candidate holin. If there were two or more such genes, and if one had a membrane topology of one of the established classes, that gene was annotated as the holin gene. Otherwise, the small membrane protein gene closest to the endolysin gene was chosen. The 19 holins established by some form of experimental evidence in the previous survey were retained in the pool, as a validation test for the system. Table 8 shows the compiled results of this search. By these criteria, of the 374 genomes, 183 are annotated as having holins belonging to class I, II or III (Fig. 37A, Table 8). Including these holin sequences in the three established classes increased the diversity of each class. Compared to the state at the last major review in 2000, there are 190 new families. Many of these putative holin genes were found to possess dual start motifs, strongly suggesting that the genes encoded both a holin and antiholin (Table 8).

In the remaining 125 genomes, the proximity criteria required the annotation of holin genes of novel topologies (Table 8). Four novel topologies were found (Fig. 37B). First, for each of the three known topologies, potential holins were found with the inverted topology, designated in Table 8 as "i' for inverted. Thus, new class I-i has holins with three TMDs but N-in, C-out. Class II-i has holins with two TMDs but N-out,

Table 8. Summary of holin annotations in Gram-negative host.

Class[a]	Families[b]	Sequences	Dual-start motifs[c]
I	28	64	13
II	41	68	18
III	45	51	2
I-i	13	22	5
II-i	15	21	1
III-i	58	76	2
IV	5	6	0
Unknown[d]		66	
Total	205	406	41

[a]Class I to IV, distinctions are based on the number of predicted transmembrane domains and membrane topologies.
[b]Number of holin sequence families
[c]A sequence is recognized as having a dual-start motif if it has the M-X_n-M… structure at its N terminus, where $_n$ is arbitrarily chosen to be <10, and at least one of the intervening residues is arginine (R) or lysine (K)
[d]Unable to identify any holin candidates based on our criteria

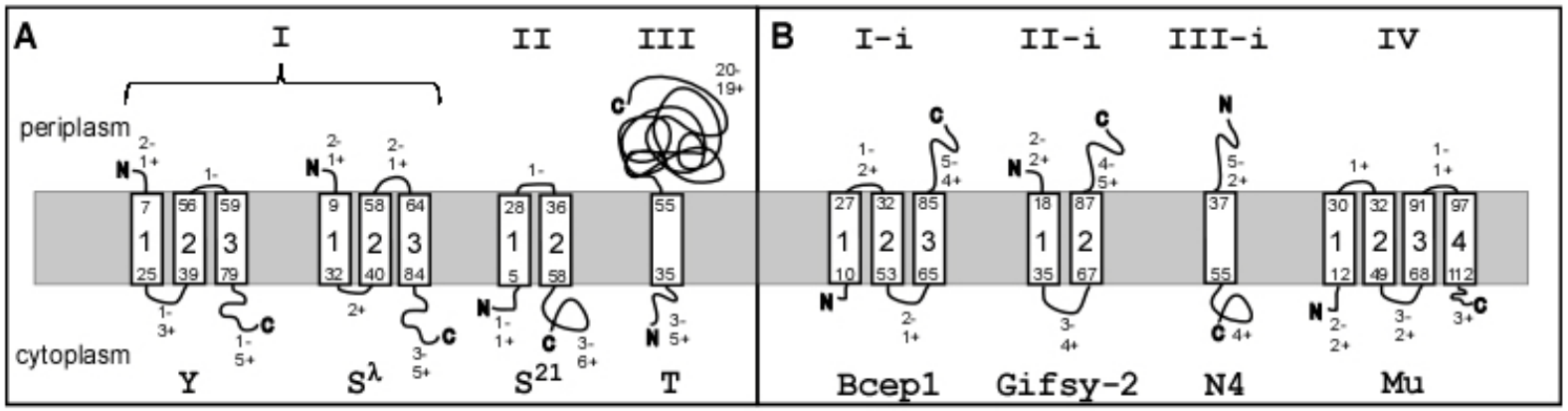

Fig. 37. New holin topologies. (A) Model of the membrane topology of Y, S105, S^{21}, and T. Topology and boundary residues for TMD1, -2 and -3 are based on hydropathicity and charges manual inspection. Three classes of holin including P2 Y and λ S (class I), S^{21} (class II), and T4 T (class III). (B) Model of new membrane topology from phage BcepI, Gifsy-2, T7, and Mu. Topology and boundary residues for TMD1, -2, -3 and -4 are based on hydropathicity and charges. The four classes of new holin including Bcep1 (class I-i), Gifsy-2 (class II-i), N4 (class III-i), and Mu (class IV)

C-out, and class III-i has holins with a single TMD but N-out, C-in. In addition, the criteria identified genes encoding class III holins with a single TMD and N-in but not homologous to T4 gp*t*. Finally, five additional genes were found encoding proteins with the same topology as the Mu gp*19* holin, thus further populating the new holin class IV.

Each of the new topology classes exhibits significant diversity. Overall, including these new holin annotations with the previously annotated holins, there are a total of 347 holins in the seven architectural classes for phages of Gram-negative hosts. Four sequences that were classified as unknown topologies in the previous survey are now each assigned to one of the seven classes. All 19 holins previously assigned with experimental evidence in the previous survey were recovered using this system, which further validated our method. Alignments within each topology class revealed that of the 347 annotated holins, most (53 sequences) were found to be "singletons", unrelated to other holins in the same topological class. For example, five more members of the new class IV established by Mu gp*19*, were identified; none of them is detectably similar to gp*19* or to each other. Table 8 shows the distribution of family sizes in each topological class.

Discussion

The new Mu lysis system

The experimental identification of gp*19* as the Mu holin is significant for several reasons. First, this means that the gene controlling lysis has now been identified for all of the paradigm experimental phages of *E. coli* and *Salmonella*. This includes, among the Caudovirales, the lambdoid phages (lambda, phage 21, P22), P1, the P2-like phages

(P2 and 186), and Mu, as well as the classic microvirus ϕX174 and the ssRNA phages MS2 and Qβ (5, 12, 96, 121, 122). Previously, detailed primary structure analysis of the Mu genomic sequence had revealed only the *lys* endolysin gene (113). The uncertainty about the existence of a Mu holin had left open the possibility that a completely different mode of lysis control existed in one of the major experimental phages. Now the substantial literature about the physiology, molecular biology, and evolution of these classic phages can now be confidently interpreted, at least in terms of the actual lytic event that terminates the infection cycle.

In addition, the analysis of the Mu lysis cassette presented here revealed a new topological class of holin, class IV, with predicted N-in and C-in topology. Prototypes of the established class I (S105, Y), class II ($S^{21}68$) and class III (gp*t*) holins have been subjected to detailed genetic and biochemical analysis. Analogous experiments should be done with Mu gp*19*, to address a number of questions. First, since Mu Lys is clearly a SAR endolysin, it must be established whether or not gp*19* is a pinholin or a canonical holin (Fig. 35B). It should be noted that for the first two SAR endolysins characterized, one (P1 Lyz) was served by a canonical holin. The possibility of pinholin character would be most clearly determined by assessing whether or not gp*19* can cause release of a soluble endolysin, like lambda R; in addition, since each tested canonical holin has been shown to allow lysis with the ~ 0.5 MDa R–βgal hybrid (40), the same test can easily be applied to gp*19*. In the only characterized pinholin, $S^{21}68$, the N-terminal TMD was shown to have SAR character and must exit the membrane during the pinhole formation pathway, leaving only a single TMD to form the heptameric pinhole (47). It is

unlikely for Mu gp*19* to undergo similarly drastic topological dynamics, since it has four TMDs; moreover, none of the TMDs have the characteristic amino acid composition characteristic of a SAR TMD (Fig. 38)

After assignment of gp*19* to either canonical holin or pinholin status, it will be important to confirm the predicted membrane topology. This is most simply done by a combination of PhoA and LacZ fusions at the predicted loops between the TMDs. In addition, experiments probing the lumen-exposed residues of the triggered holin should be informative. In chapter III, cysteine-scanning accessibility studies were used to show that triggered lambda S105 exposed a relatively hydrophilic face of TMD1 and TMD3 to the lumen of the lethal hole. Moreover, primary structure analysis of other canonical class I holins, including P2 Y and P1 LydA, revealed that each had two such potential hydrophilic faces, although in different TMDs. In contrast, similar experiments have shown that the pinholin $S^{21}68$ has only a single face exposed to the lumen of the pinhole (48). Inspection of helical projections of the TMDs of gp*19* revealed that only a single TMD, TMD3, has an analogously hydrophilic face (Fig. 38). This analysis should be tested by the Cys-scanning approach. The presence of only a single TMD with an identifiably hydrophilic surface could be taken as evidence for pinholin character of gp*19*. In any case, a thorough analysis of the topology and lumenal exposure of gp*19* is warranted. In this light, it would be important to conduct a preliminary mutational analysis of *gp19*, as has been done with the other holins. To date, in each case, genetic analysis of holins has been straightforward, taking advantage of the fact that powerful selections and screens can be done for loss of lethal function (survival after induction) or

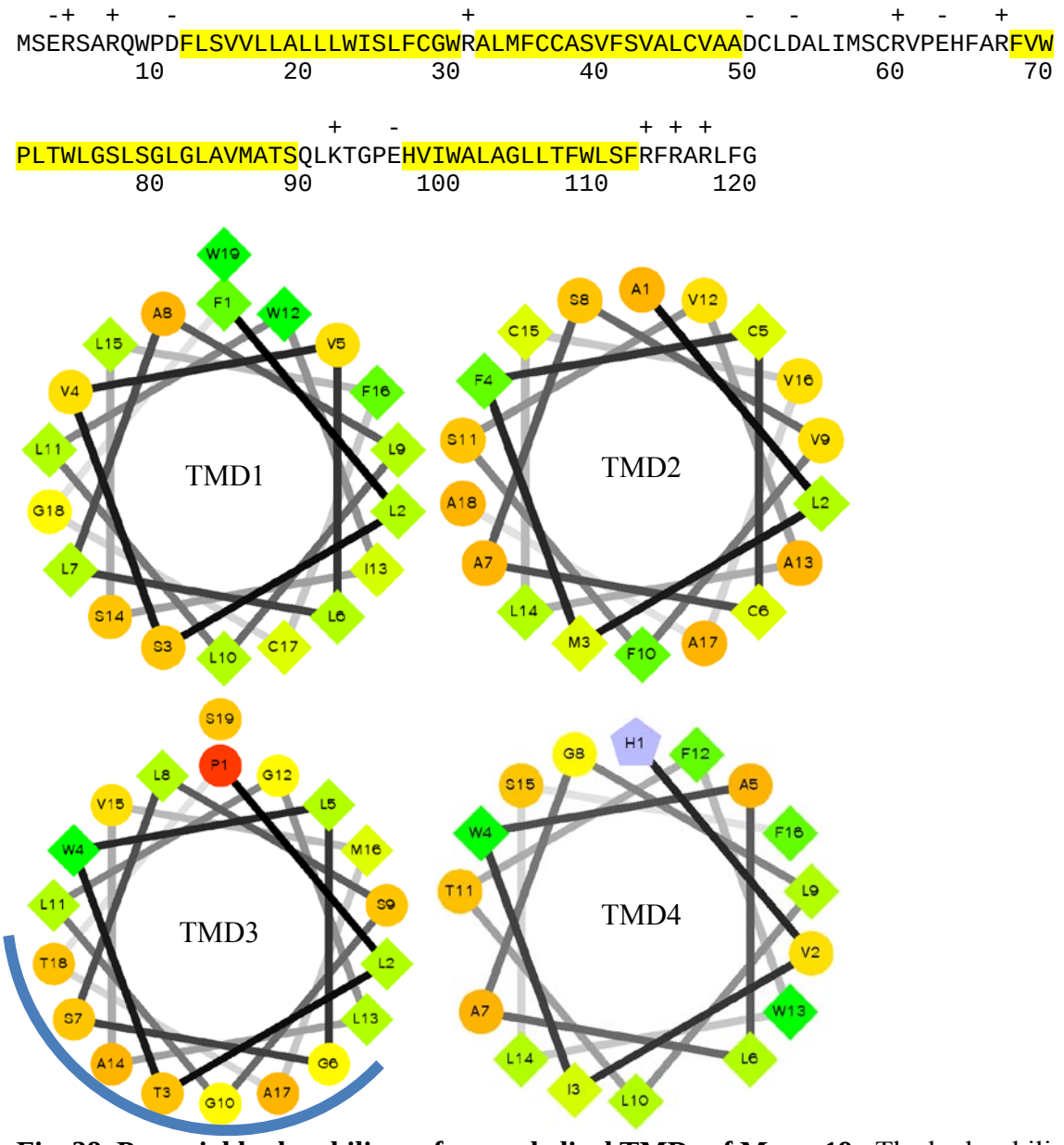

Fig. 38. Potential hydrophilic surface on helical TMDs of Mu gp19. The hydrophilic residues as circles, hydrophobic residues as diamonds, potentially negatively charged as triangles, and potentially positively charged as pentagons. Hydrophobicity is color coded as well: the most hydrophobic residue is 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 potential TMDs are highlighted in yellow in the primary structure (top). The blue arches indicate the predicted lumen face (bottom).

for lytic function (release of virions or plasmids) (49, 100). This would provide important control alleles for studies of hole size and lumenal surfaces. In addition, as shown in chapter II on P2 Y, all holins to date have characteristic mutational patterns, including extreme changes in lysis timing associated with conservative substitutions. It has been suggested that this reflects the evolutionary necessity for rapid changes in lysis timing, so similar patterns should be expected for Mu *gp19*, irrespective of the drastically different topology.

The Mu antiholin, gp*20*, also deserves some scrutiny. First, drawing on similar studies with T4 RI, it should be established whether expression *in trans* to gp*19* can cause lysis delay or inhibition, and if so, is this negative effect specific for gp*19*. In addition, it is difficult to imagine that gp*20* can retard gp*19* triggering without a direct interaction, so the cross-linking experiments used for RI and T, as well as S107 with S105 and $S^{21}71$ with $S^{21}68$, should be informative. The simplest idea is that, like with the previously characterized antiholins, the negative effect is exerted at the level of oligomerization of the holin. The most likely interaction would be via the single C-terminal TMD of gp*20*. Because of the power of the selections available for holin function (noted above), an attractive approach is to select for missense changes in the holin that eliminate the inhibition of gp*20* and then for suppressors in gp*20* that restore the regulation. Such mutation/suppression studies could identify specific interaction sites that define the holin-antiholin complex.

The phenotype of *Δgp25*, in which endolysin function appeared to have been blocked without affecting triggering was unexpected. The most attractive hypothesis is

that gp*25* is required for the release of the SAR domain of Lys. Nothing is known about the requirements for the release of SAR domains, other than the primary structure features. However, at least with P1 Lyz and other SAR endolysins that have been expressed from cloned genes, it is established that no other phage protein is required (54). Each of the characterized SAR endolysins has been shown to eventually cause lysis, even in the absence of holin triggering, as a result of the basal level of spontaneous release from the membrane. Thus, having a phage protein required for this release would confer on Mu an extra level of molecular control of the lytic process. Cloning *lys* and expressing it with and without *gp25* would be the definitive experiment.

Diversity in holin topology

The lack of conserved structure or functional motifs make identification of holins highly problematical and has led to the rather indiscriminate naming of thousands of genes in phage and prophage elements as holins, almost entirely on the basis of having similar topologies to one of the three experimentally established classes of holins (12). As noted above, absent homology with an experimental holin or antiholin, a rare occurrence indeed, there is no basis for choosing between small integral membrane proteins as holins or antiholins, other than clustering with other lysis genes. The "proximity" strategy used here for identifying candidate holin genes was a reasonable first step in seeing whether or not the holin classes established at the beginning of the genomics era represented an accurate sampling. The results were surprising. Although the original set of holins were recovered, which simply reflects the fact that it was the proximity to other lysis genes that led to the discovery of holins in the first place, four

new topology classes were identified (counting the class IV established by Mu gp*19*). If holin character can be established for prototypes of each new class, it would change the perspective on holin function and evolution profoundly. Even with the much more limited topological and sequence diversity available previously, it was already suggested that holin evolution has occurred independently many times and that, in fact, holin character may be much more closely related to the intrinsic molecular nature of integral membrane proteins. It is already known that moderate over-expression of membrane proteins can cause membrane damage and collapse of the membrane potential (5) Starting from this initial propensity, it may not be that challenging to evolve the more specific properties of having a critical concentration-dependent formation of rafts, and then, membrane lesions within the rafts. In this light, it was also surprising that each of the new topology classes, except class IV, represents a topological inversion of one of the established classes. It should be noted that it has already been shown that relatively minor mutational alteration of the charges in the loops connecting TMDs was required to achieve complete inversion of integral membrane proteins (123, 124).

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTION

The bacteriophage lytic cycle is concluded with holin-mediated permeabilization of the bacterial cytoplasmic membrane. This mode of lysis is the most frequent cytocidal event in the biosphere (13). Moreover, since holin function determines the length of the phage infection cycle, holins represent the simplest and most ancient timing mechanism known in biology. Due to advances in sequencing technology and the resultant increase in the amount of available phage genomic information, a huge number of putative holin sequences have been discovered. A search of the GenBank Proteins database with the word "holin" resulted in more than 19,000 hits. Since only ~ 1000 phage genomes have been the vast majority from prophage elements in the > 4000 bacterial genomes in the database. This enormous collection of holin sequences represent a potential "gold mine" of bioinformatic information that could be exploited in terms of understanding holins at the biological and mechanistic levels. However, only a handful of holins have been studied at all, and only three, the prototypes of each of the three established topological classes - S105 from lambda, $S^{21}68$ from phage 21, and T from T4 - have been studied in detail at the genetic, molecular and biochemical level. In this dissertation, I have focused on extending our knowledge about S105 and $S^{21}68$ to gain further insight into the structure of the final holin-mediated membrane hole and its pathway to formation (chapters III and IV), on identifying a new topological class of holin, and on examining a class I holin unrelated to the prototype S105 to see which of the latter's unusual

characteristics are shared and thus likely to be important for holin function (chapter II). In each case, the results and conclusions point towards new lines of investigation that may further illuminate the function, mechanism and structure of these fundamental proteins.

Functional analysis of the class I holin, P2 Y

At the start of the work described here, Y, a 93 aa protein, had been identified as the putative holin of the paradigm coliphage P2. Y was predicted to be an integral membrane protein that adopted an N-out, C-in membrane topology with three TMDs and a highly charged C-terminal cytoplasmic tail. Thus, Y has the same predicted topology and distribution of charged residues as the well-studied, prototype class I lambda holin, S105. Because Y was not related to S105 in primary structure, its characterization was anticipated to be useful in discerning the molecular features responsible for the function and phenotypic characteristics of holins. Here, I used physiological and genetic approaches to show that Y exhibits the essential holin functional criteria, namely allele-specific, delayed-onset lethality and sensitivity to the energization of the membrane. Also, conservative missense mutations in Y exhibited extreme variability in terms of lysis timing, another feature demonstrated for S105. This supports the notion that holins in general are able to evolve quickly, i.e., in a single step, to radically different lysis timing, which would allow the phage to adapt quickly to perturbations in the host growth environment. Also, like lambda S105, purified Y forms high order multimers in detergent when visualized by negative-stain electronic microscopy.

Taken together, these results suggest that class I holins share a set of unusual features that are related to their ability to program the end of the phage infection cycle with precise timing. In particular, these features are consistent with the model that an intermediate in the hole-formation pathway is a two-dimensional, tightly packed array of holin proteins, or holin rafts. In such arrays, conservative changes in hydrophobic side chains in the TMDs could be expected to have large effects on raft structure and stability, as a result of altered helical packing. The parallels between Y and S105 indicate that, at least among class I holins, bioinformatic analysis of the very rich diversity available might be profitably employed to look for mechanistic clues. For example, a database of class I holins could be established to look for co-evolving residue pairs or TMD surfaces that might represent interfaces for protein-protein interactions.

However, there were some differences, also informative, between Y and S105. One of the most intriguing features of the *S* system is that in the wt allele, two proteins, S105 and S107 are produced. As the result of an extra positively charged residue at the N-terminus, the latter, S107, has a different topology, with the N-terminal TMD retained in the cytosol, and as a result has negative dominant inhibitory character, or antiholin character, when complexed as a heterodimer with S105 (12, 25). However, collapse of the membrane potential allows TMD1 of S107 to integrate into the membrane, and thus, to become a functional holin, a functional feature that is necessary if antiholin inhibition is to be reversible. Here it was shown that TMD1 is also essential for Y function, based on the fact that for Y, like for S105, deletion of TMD1 creates a non-functional allele with negative-dominant character. Moreover, when an S107-equivalent mutant of Y was

created by adding positively charged residues to the N-terminus, again negative-dominant character was observed, like for S105. However, collapse of the membrane potential did not activate the mutant Y protein, indicating that there are other features of the N-terminal TMD of S107 that make it able to enter the bilayer and restore holin function to the protein. Because of the genetic power available for working with holins, it should be possible to investigate the molecular basis of this antiholin-holin convertibility. For example, we could select for mutations in the modified *Y* allele that would allow it to become activatable to membrane depolarization.

Finally, instead of encoding a second translational product as an antiholin, like S107 in lambda *S*, the P2 lysis cassette encodes another predicted membrane protein, LysA. I have here shown LysA to have Y-specific antiholin character. Again, the convenient genetics of phage and also for holin genes in particular offer appealing approaches to investigate the mechanism of specific antiholin function. With holin and antiholin genes, it is possible to select for function (i.e., lysis) or loss of function (i.e., survival of host). Thus, it should be possible to select for *Y* mutants that escape *lysA* inhibition and then for *lysA* mutants that suppress these defects, for example. Such extragenic suppressors, if shown to be allele-specific, could map the interaction points between Y and LysA. The first step in such an effort would be to manipulate the expression levels of LysA and Y so that the maximum lysis delay, or perhaps abolition of lysis, could be achieved. In the current constructs, the delay imposed by LysA *in trans* is only ~ 10 minutes. Another interesting line of experimentation would be to see if liposomes containing purified LysA were insensitive to in vitro hole formation by Y,

but not by S105. Of course, this would require establishing a liposome-based in vitro assay, as has been done for S105 (75). However, given the close parallels demonstrated here for Y and S105, the idea that a similar assay could be worked out for Y can be viewed with some confidence. If this can be done, it would perhaps open the way for a detailed mechanistic exploration of both holin and antiholin function.

The structure of the S-hole

Prior to the work described in this dissertation, a general model had been proposed for the S105 hole formation pathway in which the holin accumulated in the membrane as homodimers until reaching a critical concentration, at which point large holin rafts rapidly assembled, leading to localized membrane depolarization and then hole formation. However, the model did not specify any positions on the S105 molecule that were involved in protein-protein interactions at any step of the pathway. In chapter III, a detailed phenotypic analysis of a mutational study involving all possible substitutions for four residues constituting a complete helical turn of TMD2 was reported. The patterns of phenotypic change were found to depend on the helical face where the substitutions were made and were interpreted as identifying interaction surfaces that were important in the pathway. Moreover, Cys-scanning accessibility studies reported here in chapter III revealed that one helical face of TMDs 1 and 3, but not TMD2, were accessible in nearly all of the S105 molecules after triggering and hole-formation, indicating that these helical faces were facing the lumen of the hole. These findings allowed significant refinement of the hole-formation pathway, to the point

where different faces of each of the three TMDs were specified in terms of homotypic and heterotypic interactions in intermediates of the pathway (Fig. 27).

The refined model can be tested in several ways. For example, the Cys-substitution approach can be extended to mapping interfaces. By putting pairs of Cys residues at different positions predicted to undergo homotypic or heterotypic interactions in the pathway, and then treating the cells with the oxidant copper-phenanthroline after triggering, disulfide bonds would be expected to form if the Cys residues are located close together at an interface. This approach was used by Kaback and colleagues to work out the arrangement of TMDs in the lactose permease (125). This would allow more detail to be incorporated into the structure of the final hole (see hole model in Fig. 27). Moreover, since there are mutants available that block the pathway at several different steps, the same approach could be used to interrogate the TMD-TMD interactions in the various intermediates. It may even be possible to stabilize one or more of the intermediates by treatment of induced cells with the oxidant, perhaps allowing purification for biochemical analysis.

As long as S105 and other holins remain refractory to high resolution structural analysis by x-ray crystallography or NMR, it will be necessary to employ model-driven mutational and biochemical analysis to learn anything about the structures important in the S-hole formation pathway. The studies reported here at least provide a framework for such studies in the future.

Holin triggering and host response

Prior to the work reported here, it had been established that lethal holin triggering occurred suddenly and at the same time in nearly all the cells induced for *S105* expression (126). In chapter IV, using the "spinning bug" technique of H. Berg and colleagues (105), video microscopy studies showed that the pmf, as reported by the rotation of cells tethered to a surface by flagellar antibodies, was undisturbed throughout the period of holin accumulation before collapsing completely at the instant of triggering. Moreover, nearly all the cells induced for *S105* suddenly triggered, or stopped rotating, at the same time and the time of triggering was allele-specific. Here, I have extended this result to the heterologous class I holin, P2 Y, indicating that the allele-specific, sudden triggering is indeed a common characteristic of class I holins. However, the results obtained using the same technology with the prototype pinholin, $S^{21}68$, were dramatically different. In these inductions, the preservation of membrane integrity, as seen in the rotation of the tethered cells, was retained during the induction period, as in that of the canonical S105 holin. However, the $S^{21}68$ inductions exhibited triggering differently, in that the cells halted rotation, indicating the collapse of the pmf, and then resumed rotating. This start-stop pattern continued for multiple events, with, on average, the interval between start and stop lengthening, until, finally, stopping permanently. These results indicate that the cell can recover from pinholin triggering and permeabilization of the membrane, apparently mounting a response that restored the pmf to its full level, before succumbing again to triggering. The clear difference between the holin and pinholin systems is that in the former, thousands of holin

molecules trigger to form a few very large holes in the membrane, whereas in the latter, ultimately thousands of pinholins trigger to form ~ 10^3 heptameric pinholes (48, 49). The simplest interpretation of the "stuttering" lethality observed with the pinholins is that not all the pinholes form at one time. Instead, some fraction of the pinholes forms. Apparently, the cell is able to close or block those holes, which allows the pmf to be re-established until more pinholes form. A less likely possibility is that pinhole formation may not be irreversible until a certain level of pinholin has accumulated. In this view, the stuttering stops may reflect a population of pinholins triggering to form enough pinholes to collapse the membrane potential, but then disassembling. The difficulty with this sort of model is that during the pmf-collapsed period, there would be no ATP synthesis to support macromolecular metabolism, so no new pinholin would accumulate.

The most interesting future direction to follow up these results would be to identify the host system involved in the restoration of pmf, presumably involving somehow repairing membrane permeability. The only known system that has been reported to repair membrane damage, the psp or "phage-shock protein" (101), was tested and shown to be irrelevant to the stuttering phenomenon. It is reasonable to hypothesize that whatever the system is, if its function was enhanced, it might resist pinhole lethality entirely. A straightforward genetic approach involving selecting for survival after mutagenesis of the host and then induction of a cloned $S^{21}68$ pinholin gene might be the best method, after first reducing pinholin expression to the minimum that would sustain inducible lethality. These selections would have to be done in the absence of SAR endolysin expression, since even in the absence of pinholin function, the basal activation

rate of SAR endolysins leads to lysis of the host. There would probably be a number of false-positives, in that any host mutant that had reduced copy number or expression of the pinholin gene might show up as a survivor. However, most of these mutants could be screened out by incorporating an easily detectable colorigenic reporter gene downstream of the promoter serving the pinholin gene. Recently, a new system for repair of membrane lesions has been identified in eukaryotic systems (127); it would not be surprising to discover that analogous systems exist in prokaryotic cells. In fact, the evolution of the canonical holins, with their inordinately and unnecessarily large membrane holes, might have been driven at least in part by the need to over-ride such repair systems.

Phage Mu lysis

The work described here suggested that the identification of gp*19* as the Mu holin and gp*20* as the antiholin. The analysis of the Mu lysis cassette presented here revealed a new topological class of holin, class IV, with predicted N-in and C-in topology. Future experiments should be performed with Mu gp*19*, to address a number of questions. First, since Mu Lys is obviously a SAR endolysin, it must be established whether or not gp*19* is a pinholin or a canonical holin. It should be noted that for the first two SAR endolysins characterized, one (P1 Lyz) was served by a canonical holin. The possibility of pinholin character would be most clearly determined by examining whether or not gp*19* can cause release of a soluble endolysin, like lambda R; in addition, since each tested canonical holin has been shown to allow lysis with the ~ 0.5 MDa R–βgal hybrid (40), the equivalent experiment can be applied to gp*19*.

Moreover, a combination of PhoA and LacZ fusions at the predicted loops between the TMDs is needed to confirm the predicted membrane topology of gp*19*. Furthermore, experiments probing the lumen-exposed residues of the triggered holin should be informative. Inspection of helical projections of the TMDs of gp*19* revealed that only a single TMD, TMD3, has an analogously hydrophilic face (Fig. 38). In any case, a detailed analysis of the topology and lumenal exposure of gp*19* is needed. Primary structure analysis of other canonical class I holins, including S105, P2 Y and P1 LydA, revealed that each had two such potential hydrophilic faces, although in different TMDs. In contrast, similar experiments have shown that the pinholin $S^{21}68$ has only a single face exposed to the lumen of the pinhole (48). The presence of only a single TMD with an identifiably hydrophilic surface could be taken as evidence for pinholin character of gp*19*.

The Mu antiholin, gp*20*, also deserves more examination. First, drawing on similar studies with T4 RI, it should be established whether expression *in trans* to gp*19* can cause lysis delay or inhibition, and if so, is this negative effect specific for gp*19*. In addition, it is difficult to imagine that gp*20* can retard gp*19* triggering without a direct interaction, so the cross-linking experiments used for RI and T, as well as S107 with S105 and $S^{21}71$ with $S^{21}68$, should be informative. The simplest model is that, similar to the previously characterized antiholins, the negative effect is exerted at the level of oligomerization of the holin.

Based on the gp*25* result described in chapter V, the current hypothesis is that gp*25* is required for the release of the SAR domain of Lys. However, nothing is known

about the requirements for the release of SAR domains, other than the primary structure features, and it is established that no other phage protein is required (54). Each of the characterized SAR endolysins has been shown to eventually cause lysis, even in the absence of holin triggering, as a result of the basal level of spontaneous release from the membrane. Thus, having a phage protein required for this release would confer on Mu an extra level of molecular control of the lytic process. Cloning *lys* and expressing it with and without *gp25* would be the definitive experiment. Despite all the above questions, other questions will arise as the study of holin proteins continues.

REFERENCES

1. **Young R, Wang IN.** 2006. Phage Lysis, p. 104-126. *In* Calendar R (ed.), The Bacteriophages, 2nd ed. Oxford University Press, Oxford.

2. **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.

3. **Scheurwater E, Reid CW, Clarke AJ.** 2008. Lytic transglycosylases: bacterial space-making autolysins. International Journal Biochemistry Cell Biology **40:**586-591.

4. **Bienkowska-Szewczyk K, Taylor A.** 1980. Murein transglycosylase from phage lambda lysate: purification and properties. Biochem. Biophys. Acta. **615:**489-496.

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

6. **Dewey JS, Savva CG, White RL, Vitha S, Holzenburg A, Young R.** 2010. Micron-scale holes terminate the phage infection cycle. Proc. Natl. Acad. Sci. USA **107:**2219-2223.

7. **Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ.** 1998. JPred: a consensus secondary structure prediction server. Bioinformatics **14:**892-893.

8. **McDonnell AV, Jiang T, Keating AE, Berger B.** 2006. Paircoil2: improved prediction of coiled coils from sequence. Bioinformatics **22:**356-358.

9. **Berry JD, Rajaure M, Young R.** 2013. Spanin function requires subunit homodimerization through intermolecular disulfide bonds. Mol. Microbiol. **88:**35-47.

10. **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.

11. **Harris AW, Mount DWA, Fuerst CR, Siminovitch L.** 1967. Mutations in bacteriophage lambda affecting host cell lysis. Virology **32:**553-569.

12. **Young R.** 2002. Bacteriophage holins: deadly diversity. J. Mol. Microbiol. Biotechnol. **4:**21-36.

13. **Calendar R.** 1988. The Bacteriophages, Plenum Press, New York.

14. **Ceyssens PJ, Lavigne R, Mattheus W, Chibeu A, Hertveldt K, Mast J, Robben J, Volckaert G.** 2006. Genomic analysis of Pseudomonas aeruginosa phages LKD16 and LKA1: establishment of the phiKMV subgroup within the T7 supergroup. J. Bacteriol. **188:**6924-6931.

15. **Raab R, Neal G, Sohaskey C, Smith J, Young R.** 1988. Dominance in lambda S mutations and evidence for translational control. J. Mol. Biol. **199:**95-105.

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

17. **Bläsi U, Chang CY, Zagotta MT, Nam K, Young R.** 1990. The lethal lambda S gene encodes its own inhibitor. EMBO Journal **9:**981-989.

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

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

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

21. **Alexeyev MF, Winkler HH.** 1999. Membrane topology of the Rickettsia prowazekii ATP/ADP translocase revealed by novel dual pho-lac reporters. J. Mol. Biol. **285:**1503-1513.

22. **Bläsi U, Fraisl P, Chang CY, Zhang N, Young R.** 1999. The C-terminal sequence of the lambda holin constitutes a cytoplasmic regulatory domain. J. Bacteriol. **181:**2922-2929.

23. **Graschopf A, Bläsi U.** 1999. Molecular function of the dual-start motif in the lambda S holin. Mol. Microbiol. **33:**569-582.

24. **Grundling A, Blasi U, Young R.** 2000. Biochemical and genetic evidence for three transmembrane domains in the class I holin, lambda S. J. Biol. Chem. **275:**769-776.

25. **Grundling A, Blasi U, Young R.** 2000. Genetic and biochemical analysis of dimer and oligomer interactions of the lambda S holin. J. Bacteriol. **182:**6082-6090.

26. **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.

27. **Garrett J, Fusselman R, Hise J, Chiou L, Smith-Grillo D, Schulz R, Young R.** 1981. Cell lysis by induction of cloned lambda lysis genes. Molecular and General Genetics **182:**326-331.

28. **Garrett JM, Young R.** 1982. Lethal action of bacteriophage lambda S gene. Journal of Virology **44:**886-892.

29. **Gründling A, Manson MD, Young R.** 2001. Holins kill without warning. Proc. Natl. Acad. Sci. USA **98:**9348-9352.

30. **Smith DL, Chang CY, Young R.** 1998. The λ holin accumulates beyond the lethal triggering concentration under hyper-expression conditions. Gene Expr. **7:**39-52.

31. **Smith DL, Young R.** 1998. Oligohistidine tag mutagenesis of the lambda holin gene. J. Bacteriol. **180:**4199-4211.

32. **Smith DL.** 1998. Purification and biochemical characterization of the bacteriophage λ holin. Ph.D. dissertation. Texas A&M University, College Station.

33. **Huang L, Ming M, Liu J, Liu J, Li QG, Ding JD.** 2003. Preparation of liposome containing bacteriorhodopsin with "natural" preferred orientation and detection of its transient photoresponse. Acta. Biochemical Biophysical Sinica **35:**391-395.

34. **Rigaud JL, Pitard B, Levy D.** 1995. Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. Biochem. Biophys. Acta. **1231:**223-246.

35. **Peisajovich SG, Shai Y.** 2003. Liposomes in identification and characterization of viral fusogenic peptides. Methods Enzymology **372:**361-373.

36. **Deaton J, Sun J, Holzenburg A, Struck DK, Berry J, Young R.** 2004. Functional bacteriorhodopsin is efficiently solubilized and delivered to membranes by the chaperonin GroEL. Proc. Natl. Acad. Sci. USA **101:**2281-2286.

37. **White R, Chiba S, Pang T, Dewey JS, Savva CG, Holzenburg A, Pogliano K, Young R.** 2011. Holin triggering in real time. Proc. Natl. Acad. Sci. USA **108:**798-803.

38. **Krebs MP, Khorana HG.** 1993. Mechanism of light-dependent proton translocation by bacteriohodopsin. J. Bacteriol. **175:**1555-1560.

39. **Mehrbod M, Mofrad MR.** 2013. Localized lipid packing of transmembrane domains impedes integrin clustering. PLoS Computational Biology **9:**e1002948.

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

41. **Savva CG, Dewey JS, Deaton J, White RL, Struck DK, Holzenburg A, Young R.** 2008. The holin of bacteriophage lambda forms rings with large diameter. Mol. Microbiol. **69:**784-793.

42. **Dewey JS, Struck DK, Young R.** 2009. Thiol protection in membrane protein purifications: a study with phage holins. Anal. Biochem. **390:**221-223.

43. **Wishart DS, Sykes BD.** 1994. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. Journal Biomolecular NMR **4:**171-180.

44. **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.

45. **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.

46. **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.

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

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

49. **Pang T, Park T, Young R.** 2010. Mutational analysis of the S21 pinholin. Mol. Microbiol. **76:**68-77.

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

51. **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. USA **101:**6415-6420.

52. **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.

53. **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.

54. **Xu M, Struck DK, Deaton J, Wang IN, Young R.** 2004. The signal arrest-release (SAR) sequence mediates export and control of the phage P1 endolysin. . Proc. Natl. Acad. Sci. USA **101:**6415-6420.

55. **Streisinger G, Mukai F, Dreyer WJ, Miller B, Horiuchi S.** 1961. Mutations affecting the lysozyme of phage T4. Cold Spring Harbor Symposium on Quantitative Biology **26:**25-30.

56. **Lu MJ, Henning U.** 1992. Lysis protein T of bacteriophage T4. Molecular and General Genetics **235:**253-258.

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

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

59. **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.

60. **Wang IN.** 2006. Lysis timing and bacteriophage fitness. Genetics **172:**17-26.

61. **Sonnhammer EL, von Heijne G, Krogh A.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Sixth Int. Conf. Intelligent Systems Molecular Biology **6:**175-182.

62. **Hofmann K, Stoffel W.** 1993. TMbase - A database of membrane spanning proteins segments. Biol. Chem. Hoppe Seyler **374:**166.

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

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

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

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

67. **Savva CG, Dewey JS, Deaton JF, White RL, Struck DK, Holzenburg A, Young R.** 2008. The holin of bacteriophage lambda forms rings with large diameter. Mol. Microbiol. **69:**784-793.

68. **Raab R, Neal G, Sohaskey C, Smith J, Young R.** 1988. Dominance in lambda *S* mutations and evidence for translational control. J. Mol. Biol. **199:**95-105.

69. **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.

70. **Ziermann R, Bartlett B, Calendar R, Christie GE.** 1994. Functions involved in bacteriophage P2-induced host cell lysis and identification of a new tail gene. J. Bacteriol. **176:**4974-4984.

71. **Gründling A, Bläsi U, Young R.** 2000. Genetic and biochemical analysis of dimer and oligomer interactions of the λ *S* holin. J. Bacteriol. **182:**6082-6090.

72. **Schweizer HP.** 1990. The pUC18CM plasmids: a chloramphenicol resistance gene cassette for site-directed insertion and deletion mutagenesis in *Escherichia coli*. Bio. Techniques **8:**612-613, 616.

73. **Silhavy TJ, Berman ML, Enquist LW.** 1984. Bacterial strains, p. xi-xiii. *In* Silhavy TJ, Berman ML, Enquist LW (ed.), Experiments with gene fusions, vol. 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

74. **Smith DL, Struck DK, Scholtz JM, Young R.** 1998. Purification and biochemical characterization of the lambda holin. J. Bacteriol. **180:**2531-2540.

75. **Deaton J, Savva CG, Sun J, Holzenburg A, Berry J, Young R.** 2004. Solubilization and delivery by GroEL of megadalton complexes of the lambda holin. Protein Sci. **13:**1778-1786.

76. **Miroux B, Walker JE.** 1996. Over-production of proteins in *Escherichia coli* : mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J. Molelcular Biology **260:**289-298.

77. **Valentine RC, Shapiro BM, Stadtman ER.** 1968. Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from *Escherichia coli*. Biochemistry **7:**2143-2152.

78. **Gründling A, Bläsi U, Young R.** 2000. Biochemical and genetic evidence for three transmembrane domains in the class I holin, λ S. J. Biol. Chem. **275:**769-776.

79. **Rydman PS, Bamford DH.** 2003. Identification and mutational analysis of bacteriophage PRD1 holin protein P35. J. Bacteriol. **185:**3795-3803.

80. **Kyte J, Doolittle RF.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105-132.

81. **Rietsch A, Fraisl P, Graschopf A, Bläsi U.** 1997. The hydrophilic C-terminal part of the lambda S holin is non-essential for intermolecular interactions. FEMS Microbiology Letters **153:**393-398.

82. **Tao T, Lamkin M, Scheiner CJ.** 1985. The conformation of the C-terminal region of actin: a site-specific photocrosslinking study using benzophenone-4-maleimide. Arch. Biochem. Biophys. **240:**627-634.

83. **Bläsi 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.

84. **Walker JT, Walker DH, Jr.** 1980. Mutations in coliphage P1 affecting host cell lysis. J. Virol. **35:**519-530.

85. **Yarmolinsky MB, Sternberg N, Calendar R.** 1988. Bacteriophage P1, p. 291-438. *In* Fraenkel-Conrat H, Wagner RR (ed.), The Bacteriophages, vol. 1. Plenum Press, New York.

86. **Lobocka M, Rose DJ, Plunkett G, III, Rusin M, Samojedny A, Lehnherr H, Yarmolinsky MB, Blattner FR.** 2004. Genome of bacteriophage P1. J. Bacteriol. **186:**7032-7068.

87. **Nam K, Bläsi U, Zagotta MT, Young R.** 1990. Conservation of a dual-start motif in P22 lysis gene regulation. J. Bacteriol. **72:**204-211.

88. **Smith DL, Young R.** 1998. Oligohistidine tag mutagenesis of the lambda holin gene. J. Bacteriol. **180:**4199-4211.

89. **Kolisnychenko V, Plunkett G, III, Herring CD, Feher T, Posfai J, Blattner FR, Posfai G.** 2002. Engineering a reduced *Escherichia coli* genome. Genome Res. **12:**640-647.

90. **Powell BS, Rivas MP, Court DL, Nakamura Y, Turnbough CL, Jr.** 1994. Rapid confirmation of single copy lambda prophage integration by PCR. Nucleic. Acids Res. **22:**5765-5766.

91. **Lu J, Deutsch C.** 2001. Pegylation: a method for assessing topological accessibilities in Kv1.3. Biochemistry **40:**13288-13301.

92. **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.

93. **Curran AR, Engelman DM.** 2003. Sequence motifs, polar interactions and conformational changes in helical membrane proteins. Curr. Opin. Struct. Biol. **13:**412-417.

94. **Adamian L, Liang J.** 2001. Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. Journal of Molecular Biology **311:**891-907.

95. **Zheng Y, Struck DK, Dankenbring CA, Young R.** 2008. Evolutionary dominance of holin lysis systems derives from superior genetic malleability. Microbiology **154:**1710-1718.

96. **Zheng Y, Struck DK, Bernhardt TG, Young R.** 2008. Genetic analysis of MraY inhibition by the ϕX174 protein E. Genetics **180:**1459-1466.

97. **Wimley WC, Selsted ME, White SH.** 1994. Interactions between human defensins and lipid bilayers: Evidence for formation of multimeric pores. Protein Science **3:**1362-1373.

98. **Akabas MH, Stauffer DA, Xu M, Karlin A.** 1992. Acetylcholine receptor channel structure probed in cysteine-substitution mutants. Science **258:**307-310.

99. **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.

100. **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.

101. **Darwin AJ.** 2005. The phage-shock-protein response. Mol. Microbiol. **57:**621-628.

102. **Parkinson JS, Houts SE.** 1982. Isolation and behavior of *Escherichia coli* deletion mutants lacking chemotaxis functions. J. Bacteriol. **151:**106-113.

103. **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.

104. **Garza AG, Biran R, Wohlschlegel JA, Manson MD.** 1996. Mutations in motB suppressible by changes in stator or rotor components of the bacterial flagellar motor. J. Molelcular Biology **258:**270-285.

105. **Berg HC, Block SM.** 1984. A miniature flow cell designed for rapid exchange of media under high- power microscope objectives. J. Gen. Microbiol. **130 (Pt 11):**2915-2920.

106. **Silverman M, Simon M.** 1974. Flagellar rotation and the mechanism of bacterial motility. Nature **249:**73-74.

107. **Fung DC, Berg HC.** 1995. Powering the flagellar motor of *Escherichia coli* with an external voltage source. Nature **375:**809-812.

108. **Baumgarten T, Sperling S, Seifert J, von Bergen M, Steiniger F, Wick LY, Heipieper HJ.** 2012. Membrane vesicle formation as a multiple-stress response mechanism enhances Pseudomonas putida DOT-T1E cell surface hydrophobicity and biofilm formation. Appl. Environ. Microbiol. **78:**6217-6224.

109. **Bienkowska-Szewczyk K, Lipinska B, Taylor A.** 1981. The R gene product of bacteriophage lambda is the murein transglycosylase. Molecular and General Genetics **184:**111-114.

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

111. **Shapiro JA.** 1979. Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc. Natl. Acad. Sci. USA **76:**1933-1937.

112. **Symonds N, Toussaint A, Van De Putte P, Howe M.** 1987. Phage Mu. *In* Symonds N, Toussaint A, Van De Putte P, Howe M (ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

113. **Siboo IR, Sieder F, Kumar K, Howe MM, DuBow MS.** 2004. Characterization of Plys-proximal morphogenetic genes of transposable bacteriophage Mu. Archives Virology **149:**241-259.

114. **Silhavy TJ, Berman ML, Enquist LW.** 1985. Bacterial strains, p. xi-xiii. *In* Silhavy TJ, Berman ML, Enquist LW (ed.), Experiments with gene fusions, vol. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

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

116. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic. Acids Res. **25:**3389-3402.

117. **Krogh A, Larsson B, von Heijne G, Sonnhammer EL.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Molelcular Biology **305:**567-580.

118. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944-945.

119. **Tran TAT, Struck DK, Young R.** 2005. Periplasmic domains define holin-antiholin interactions in T4 lysis inhibition. J. Bacteriol. **187:**6631-6640.

120. **Arisaka F, Kanamaru S, Leiman P, Rossmann MG.** 2003. The tail lysozyme complex of bacteriophage T4. Int. J. Biochem. Cell Biol. **35:**16-21.

121. **Gschwender HH, Hofschneider PH.** 1969. Lysis inhibition of f X174, MS2, and Qβ-infected *Escherichia coli* bacteria by magnesium ions. Biochem. Biophys. Acta. **190:**454-459.

122. **Bernhardt TG, Wang IN, Struck DK, Young R.** 2001. A protein antibiotic in the phage Qβ virion: diversity in lysis targets. Science **292:**2326-2329.

123. **Von Heijne G, Manoil C.** 1990. Membrane proteins: from sequence to structure. Protein Engineering **4:**109-112.

124. **Von Heijne G.** 1989. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. Nature **341:**456-458.

125. **Wolin CD, Kaback HR.** 2000. Thiol cross-linking of transmembrane domains IV and V in the lactose permease of *Escherichia coli*. Biochemistry **39:**6130-6135.

126. **Gründling A, Manson MD, Young R.** 2001. Holins kill without warning. Proc. Natl. Acad. Sci. USA **98:**9348-9352.

127. **Mellgren RL.** 2011. A new twist on plasma membrane repair. Communicative Integrative Biology **4:**198-200.