MOLECULAR, GENETIC, AND BIOCHEMICAL ANALYSIS OF PHAGE LAMBDA LYSIS

A Dissertation

by

JESSE L. CAHILL

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,	Andreas Holzenburg
	Jennifer Herman
	Jason Gill
Head of Department,	Gregory D. Reinhart

December 2017

Major Subject: Biochemistry

ABSTRACT

Spanins are required for outer membrane disruption, the final step in phage lysis of Gram-negative hosts. Recently spanins were shown to be fusogenic and it was proposed that spanins function by fusing the inner and outer membrane. This work uses a genetic approach to probe the functional domains of the λ spanins Rz and Rz1 by selecting for lysis-defective alleles. Our selection showed single missense mutations clustered within subdomains essential to other membrane fusion systems, including coiled-coil domains and a proline-rich region. Surprisingly, most products of lysis defective alleles were normal for accumulation and complex formation. This suggests that a majority of the mutations blocked function at a downstream step, e.g. membrane fusion.

To gain insight into spanin function, we selected for spontaneous suppressors that restored plaque formation to lysis defective alleles. Strikingly, regardless of the site of inactivating mutation, the second site rescuing mutations clustered within a coiled-coil domain near the cytoplasmic membrane. These changes encoded polar insertions into the hydrophobic core and were not allele specific. Furthermore, suppressor mutants were defective for Rz accumulation and exhibited a defect in lysis morphology. Instead of identifying point-to-point contacts, a global suppression pattern was indicated. This suggests that destabilization of the membrane-proximal segment of the Rz coiled-coil can rescue function for lysis defective alleles of Rz or Rz1 at cost of normal saltatory function.

Lastly, it is not known how λ causes lysis from the poles of *E. coli*. To address this, we use time-lapse microscopy to monitor the activity and subcellular localization of lysis proteins in the seconds prior to lysis. Results exclude the endolysin and spanin and indicate that the holin, which initiates lysis, also controls the site of lysis.

DEDICATION

To my wife Alyssa, for her constant support, selflessness, and kindness. You always see the best in me…this couldn't have happened without you.

ACKNOWLEDGEMENTS

I would like to thank my PhD mentor, Dr. Ry Young. I am sincerely grateful and privileged to work in your lab. Your example of scholarship and style of mentoring helped me prosper and become passionate about research. Thanks to my committee members, Dr. Holzenburg, Dr. Herman, and Dr. Gill, for their guidance and support throughout the course of this research. I am lucky to have a committee with such a range of expertise that has provided me with both invaluable encouragement and thoughtful critique.

Thanks go to all members of the Young lab (past and present) and everyone in the Center for Phage Technology. I am especially thankful for the hands-on training and mentoring provided by the predecessor of the spanin project, Dr. Manoj Rajaure. I also thank Dr. Karthik Chamakura for helping with my training. Thank you to Rohit Kongari and Adriana Hernandez for their encouragement and help preparing for talks. I am thankful to Daisy Wilbert for going above and beyond to provide assistance with scheduling and administrative work. Thank you to my undergraduate research team: Chandler O'Leary, Ashley Holt, Russell Moreland, Armando Marrufo, Jake Chamblee, and Jordan Sloan. Also thanks to the research technicians that have helped me with my project: Eleni Mijalis and Matthew Theodore.

Thanks to the good friends I've found in graduate school, especially Dr. Ben Kaster, Antonio Reyna, Yongjin Liu, and Anthony Sperber. This journey would have been impossible without your comradery.

I would also like to thank Jill Larson, Dr. Brian Merkel, and Dr. Michael Zorn for helping me find direction in my career. I am grateful to Thomas Tessner, Jill Larson, Dr. Brian Merkel, and Dr. Mike Hencheck for showing a love of the art of teaching, which was very inspirational to me.

Finally, thanks to all my family members for their encouragement and support. Especially my mother Kari Cahill, my late grandfather, and grandmother Susan Cahill. Thank you for raising me to value education and to think independently. Thanks to Alyssa's parents, Burl and Jo Ann Boone for being the best in-laws ever. Thanks once again to my wife Alyssa for her patience and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Ry Young (advisor) and Jen Herman of the Biochemistry and Biophysics Department, Professor Jason Gill of the Department of Animal science and Dr. Andreas Holzenburg of the Department of Biomedical Sciences at the University of Texas Rio Grande Valley.

The work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by Public Health Service grant GM27099, and by the Center for Phage Technology at Texas A&M University, jointly sponsored by Texas AgriLife

NOMENCLATURE

IM	inner membrane
PG	peptidoglycan
OM	outer membrane
LPS	lipopolysaccharide
TMD	transmembrane domain

`

TABLE OF CONTENTS

Page

Page

CHAPTER I

INTRODUCTION

Bacteriophages

Bacteriophages (phages) are viruses of bacteria. They are the most abundant biological entity on the planet (1). One mL of seawater holds between ten and fifty million bacteriophages and estimates suggest 10^{31} phages inhabit the biosphere (2, 3). If arranged end-to-end, 10^{31} phages would span 10^{8} light years, a distance that could circle our galaxy more than 300 times. Phages are found everywhere on the planet and are especially prevalent in biomes enriched with bacteria, such as in the lower gastrointestinal tract and in sewage (4). The most commonly isolated phages are *Caudovirales*, tailed phages with double-stranded DNA enclosed within a capsid. To propagate, phages must infect, replicate and lyse the host cell to release progeny. Temperate phages encode a genetic switch that allows a "choice" whether to proceed toward lysis or maintain their chromosome within the host. When the latter pathway is followed, the phage chromosome is replicated along with the host as a prophage. In this state, the phage is harmless to the host until the phage is induced to resume the lytic pathway. Phage λ is the model temperate phage and has the most well-studied lysis system. Most of the work described in this text will focus on the function of the λ lysis proteins.

Phage Therapy

Bacteriophages are a promising solution to the emerging threat of drug-resistant bacteria. Phage therapy predates antibiotic use, and was most popular from the 1920s to the 1940s. Trials to treat pathogens with phages were conducted virtually everywhere in the developed world, from the United States to Russia. However, in the 1940s, a number of factors lead to a steep decline in the interest of phage therapy; among the contributors that hindered phage therapy's global acceptance were the discovery of antibiotics, the social climate associated with phage therapy, research methods used, and ultimately the founder of phage therapy himself (5).

In 1917, Felix d'Hérelle (1873-1949), a self-trained microbiologist was studying the growth of enteric bacteria. In one of his gel-cultures, d'Hérelle noticed a cleared zone in otherwise normal bacterial growth. Upon mixing a filtrate of this cleared zone with a fresh sample of dysentery culture, the dysentery bacteria were quickly destroyed. He gave the name "bacteriophage" to the entity responsible. Credit for phages' first discovery could be attributed to Frederick Twort (1877-1950), who made similar discoveries in 1915. However, d'Hérelle pursued the investigation of the bacteriophage; whereas Twort believed agent responsible for bacterial clearing was an enzyme secreted by the bacteria themselves (6). Twort was unable to continue further investigation, due to the start of WWI and the loss of a grant (7).

The use of the bacteriophage against bacterial infections was immediately apparent to d'Hérelle, who in 1919 administered a phage preparation to a 12-year old boy with severe dysentery. d'Hérelle and other hospital interns drank the broth first to

demonstrate its harmlessness. The young boy recovered after one dose (7). A promising career in research was to follow. d'Hérelle wrote 110 articles and 6 books and was awarded an honorary doctorate and appointment as Professor of Protobiology at Yale University from 1928-1933 (8). d'Hérelle formed a commercial laboratory in Paris, producing several phage preparations useful against various bacterial infections. These preparations were marketed by what would later become L'Oréal Company. Phage therapy products were even produced in the United States by Eli Lilly, products to treatment of various cocci as well as *E. coli* infections. The effectiveness of these treatments is still disputed. Commercial production in the United States, Europe, and Russia continued up until the advent of antibiotics in 1941. Antibiotic use led to a steep decline for phage therapy in most of the western world, up until present day (7). Almost every country that had established centers of phage research moved away from phage therapy. However, phage therapy continues in certain areas of the world. The Bacteriophage Institute in Tbilisi (Republic of Georgia) still practices phage therapy today.

The origins of phage therapy and the countries that embraced it were tied to communism. Therefore, the western world regarded phage therapy as "Red Medicine". Additionally, phage research was conducted under loose experimental protocols. A self-trained scientist, d'Hérelle was a zealot for his cause, short-cutting the "bench to bed" time that medical research requires. As a consequence, his impatience led to research (by him and other phage therapy researchers) that lacked proper controls that were needed to prove efficacy to other medical communities. While d'Hérelle's popularity may have

diminished in the west, it grew in the Soviet Union, where he was welcomed as a hero. Phage therapy was used by soldiers on both sides in World War II (predominantly by the Soviet Union). However, documentation of these treatments and supplemental research was not translated into English; as a result, western scientists were unaware of some of phage therapy's more noteworthy achievements during that period (7). After World War II, antibiotics became widely available. Antibiotics lacked some of the drawbacks that plagued phage therapy. To western medicine, phage therapy would effectively be abandoned. Felix d'Hérelle did not live to see the demise of his life's work; he died shortly after the war of pancreatic cancer (8).

Lysis

By the numbers, phage lysis is the most common cell fate on the planet. Lysis by marine viruses accounts for the largest carbon flux within the oceans, estimated to be 150 gigatons annually (9). Evolution provides constant pressure to tune the lytic process in order to optimize phage progeny release. This implies phages have mechanism of tuning the timing of lysis to match the changing demands of the milieu. For example, early lysis produces a faster lifecycle at the expense of a reduced phage yield. However, shortening the morphogenesis period would be advantageous in a host-dense environment. Conversely, delaying the onset of lysis results in an increased production of progeny, which is beneficial in a host-depleted environment (10, 11).

Cell envelope

Dimensions and components

Cell envelopes are multilayered structures which offer protection to the prokaryotic cell. This structure must be destroyed in order to release phage progeny. Most bacteria are classified into one of two major groups based off of the type of cell envelope. Gram-positive bacteria are enclosed by a thick PG layer that surrounds the IM. Gram-negative bacteria have a thin PG layer surrounding the IM. The PG of Gram-negative bacteria is enclosed by an OM that asymmetric; the inner leaflet made up of phospholipids and the outer leaflet composed of LPS (Fig. A.1) (12). Since this work focuses on the lysis system of λ, the components of the *E. coli* (Gram-negative) cell envelope will be discussed.

The periplasm is an aqueous compartment enclosed by the two membrane bilayers (Fig. A.1). Each membrane bilayer is about ~3 nm wide and the periplasmic space between the two membranes is ~25 nm (13). The periplasm contains many soluble proteins, including nucleases, chaperones, and proteases (14). The iso-osmotic balance between the cytoplasm and the periplasm is maintained by membrane-derived oligosaccharides (MDOs), which are made up of 8-10 highly branched glucose units. Osmoregulation is controlled by increasing MDO synthesis in low-osmolarity media (15). The PG offers the majority of structural protection to internal osmotic pressure (16). The PG is located in the periplasm and cryoEM studies have shown that PG is ~6 nm thick and located approximately halfway between the IM and OM (17). PG is composed of disaccharide-peptide repeat units linked by glycosidic bonds that form

glycan strands. These strands are cross-linked through pentapeptides by 4-3 and 3-3 linkages (18). The PG is covalently linked to the OM by an OM lipoprotein, Lpp (Fig. A.1). There are nearly a million copies of Lpp per cell and one-third are bound to the PG via the C-terminal Lys residue. Since the periplasm and cytoplasm are iso-osmotic, the OM must be strong enough to resist internal turgor pressure. Therefore, the PG-Lpp linkage is important for the strength of the OM; supporting this idea, Lpp mutants shed OM vesicles (19).

The IM is a semi-permeable phospholipid bilayer composed of phosphatidyl ethanolamine, phosphatidyl glycerol, lysophosphatidyl ethanolamine and cardiolipin (20). The IM encompasses the cytoplasm and is the site of energy production, protein secretion and transport and lipid biosynthesis. The cell maintains a 0.2 V electrochemical gradient across the IM that provides the major source of cellular energy during aerobic metabolism (19).

The most distinguishing feature of the OM is asymmetry; the inner leaflet is made up of phospholipids, whereas the outer leaflet is composed of LPS. In blood infections, LPS is known to cause endotoxic shock, triggered by the immune reaction to Gram-negative bacteria. The LPS molecule is embedded in the membrane by lipid-A, which distally extends a complex structure of hexose and heptose oligosaccharides. Importantly, phosphate groups present on LPS are able to form ionic bridges between LPS molecules in the presence of divalent cations. Treatment of *E. coli* with ethylenediaminetetraacetate (EDTA) results in increased susceptibility to lysozyme, whereas divalent cations prevent this action (21). The stability of the OM is pertinent to

phage lysis. Shearing forces of aerobic growth artifactually complement the lysis defect of spanin mutants. Conversely, media supplemented by divalent cations stabilizes the OM; in these conditions the spanins are required for OM disruption (discussed later) (22).

Membrane protein sorting and modification

The IM is a barrier to ions and proteins; therefore, proteins synthesized by the cell require transport mechanisms for export of cell envelope proteins. With the exception of the soluble endolysins, all known lysis proteins of tailed phages are anchored in the membrane by a TMD(s) or a lipid anchor (23). Therefore, phage must encode lysis proteins to be compatible with host membrane protein transport and processing machinery. Proteins are most often targeted to the cell envelope by either the Sec or Tat (twin arginine translocation) systems, which share similar targeting signals but differ in terms of translocation mechanisms. Another difference is the Sec system secretes unfolded proteins and is involved in inserting membrane proteins into the IM, whereas the Tat system secretes folded proteins. Sec signal sequences are characterized by ~20 hydrophobic residues at the N-terminus which are flanked by short amino terminal segment encoding positively charged residues. Polar residues are typically found within the C-terminal region. Tat signal sequences resemble Sec signal sequences but usually have a longer amino-terminal segment that encodes the Tat motif, Z-R-R-x-*-*, where Z and "*" represent polar and hydrophobic residues, respectively (24). Soluble periplasmic proteins mostly travel through the Sec translocon and are cleaved at the C-terminal end of the signal sequence after translocation. The presence of small

polar residues at the -1 and -3 positions from the cleavage site are required for recognition by signal peptidase I. Some Sec-translocated proteins are further modified to be lipoproteins, which can be found in either membrane (24). Lipoproteins are processed by signal peptidase II at a Cys residue which is near the C-terminal end of its signal sequence. Cleavage results in a newly exposed Cys at the N-terminus that is lipoylated. Lipoproteins are sorted to the IM or OM by the Lol system. Lipoproteins that encode aspartic acid at the +2 or +3 position of the cleavage site are not recognized by the Lol system, meaning the lipoprotein will remain anchored in the outer leaflet of the IM. Conversely, the presence of small polar residues in place of the Asp signal result in Lol-mediated transport of the lipoprotein to the inner leaflet of the OM (25).

Modes of tailed phage release

λ-like lysis systems

Tailed phages encode genes for lysis proteins dedicated to disrupting each component of the cell envelope during lysis. The holin is an IM protein that controls the timing of lysis. Accumulation of the holin and the other lysis proteins is quiescent to the host until, at an allele-specific time, the holin suddenly forms large holes in the IM (Fig. A.2A). This causes the release of active muralytic endolysin from the cytoplasm to the periplasm where it degrades its substrate, PG (Fig. A.2A). It was once thought degradation of the PG was sufficient for lysis, however spanins are required for OM disruption (22, 23). Spanins form complexes that connect the IM and OM. Spanins have been shown to cause efficient membrane fusion of spheroplasts. Based on these

data, and similarities of spanins to viral fusion proteins, spanins have been proposed to cause OM disruption by fusing the IM and OM, completing the lysis process (Fig. A.2A) (26). Although each step of lysis depends on the previous step, lysis proteins affect different compartments of the cell without interaction with each other.

SAR endolysin

The SAR (Signal Anchor Release)-endolysin system represents an evolutionarily distinct mode of lysis. The prototypical SAR-carrying phage is lambdoid phage 21. In this system, the pinholin $S^{21}68$ forms ~2 nm holes at the time of lysis (Fig. A.2B) (27, 28). This creates a sudden depolarization that releases the SAR (Signal Anchor Release)-endolysin from the inactive membrane-anchored form into the periplasm (Fig. A.2B). Once the SAR-endolysin is released, it is in a soluble form that will attack the PG. As with λ, lysis is complete upon OM disruption by the spanins of phage 21 (23).

Lambda lysis

The lysis proteins of λ have been studied for more than 40 years using biochemical and genetic means, making it the paradigm lysis system. As a temperate phage, λ lysogens can be induced to lyse in synchrony. Most lab λ variants encode a temperature sensitive repressor allele (*cI_{857}*), which can be induced for lysis by shifting growth from 30 °C to 42 °C. This action de-represses the lysogenic state and initiates the lytic genetic circuit. Lysis genes and other late genes are transcribed ~8 min after induction. The λ lysis cassette is under control of the late promoter P_R' (23). During lysogenic growth, transcription from P_R' is terminated by a downstream terminator.

However, during the lytic cycle, λ produces the anti-terminator Q, which causes RNA polymerase to ignore the termination signal (29). Thus late genes are transcribed ~8 minutes after induction. Much of the work described in the following chapters uses medium copy plasmids (denoted pR' or pRE) that encode alleles of lysis genes downstream of P_R'. Q can be supplied by induced λ lysogens or from a compatible vector under IPTG control (pQkan or pQcam). Thus the pRE plasmid system provides a convenient and flexible platform that replicates in vivo expression levels.

Holin

Early work with the λ holin demonstrated that lysogens expressing the holin without endolysin were deficient for lysis. At the time of lysis, cells expressing the holin alone were physiologically dead, lost cytoplasmic contents, and could not be converted to spheroplasts. This led to a model in which during lysis the holin makes holes in the IM large enough to release endolysin from the cytoplasm to the periplasm (30).

The *S* gene of λ encodes two proteins of opposite function. *S105* is the holin, and *S107* is the antiholin (Fig. A.3A) (31). Both genes share the same coding sequence, except the translational start for *S107* is two codons upstream of *S105*. S105 is a 105 aa IM protein with TMDs and an N-out, C-in topology (Fig. A.3B). The antiholin has additional positive charge encoded at the N-terminus that results in the N-in, C-in topology of S107 (Fig. A.3B). The subtle differences in coding sequences have dramatic results; expression of *S105* is lytic in the absence of S107. Conversely, *S107* expressed alone is not lytic (31). The role of the antiholin will be discussed later.

The "lysis clock" function of the holin has been made evident by genetic analysis

(32, 33). Even conservative missense changes within the holin TMDs result in accelerating or retarding the lysis clock (34). The holin is the only determinant of the length of the infective cycle; therefore, from an evolutionary standpoint, the phage can use the holin to optimize the timing of lysis. A conserved feature of holin function is that energy poisons can prematurely activate the holin to disrupt the IM. Based on this, an early model of holin function suggested that holins progressively cause protons to leak across the IM, titrating the pmf until reaching a critical threshold. This model was invalidated by experiments using flagella-tethered lysogens. The speed of rotation of tethered cells is proportional to pmf. Therefore, the proton leakiness model would predict a progressive reduction in the rate of rotation of a tethered lysogen at the time of lysis. In fact, lysogens maintained a constant rate of rotation up until ~19 sec before lysing, demonstrating that holins stop pmf abruptly (35). Since the holin is small and forms holes sufficient in size for endolysin release, high-order oligomerization would be expected. Supporting this model, oligomeric actions have been detected, including mutants blocked in dimerization and oligomerization (36). To account for the evidence of large hole formation and precise timing of the holin, the "death raft" model has been proposed. In this model, holins form dimers and accumulate as large aggregates, or "rafts". Then, at an allele-specific time, rafts would collapse into holes that initiate lysis. This model was refined by a GFP-fusion study, which demonstrated that the holin accumulates in a mobile form (most likely a dimer). Time-lapse microscopy of the holin-GFP chimera demonstrated that rafts form suddenly at the time of lysis. The death raft model was revised to reflect these data, which implied that rafts do not form until a

critical concentration has been reached (Fig. A.3). Furthermore, allele-specific differences in lysis timing most likely reflect changes in the critical concentration required for raft formation (37). This critical concentration model has precedent in halobacteria, where bacteriorhodopsin (BR) forms a two-dimensional lattice (purple membrane). Amino acid substitution within the helix-helix interface resulted in alterations in amounts of latticed-BR versus total BR (38). The critical concentration of S105 mutants is allele-specific, i.e. early lysing alleles have less protein in the membrane at the time of lysis. Therefore, the refined "death raft" model is more plausible, in that the timing of raft formation is genetically programmed. Once rafts form, the model predicts that rafts block energy production. Pmf loss shifts the energy field of the cell which causes rafts to convert to a hole-forming arrangement, which is referred to as "triggering" (Fig. A.3).

The holin forms membrane holes of unprecedented size. This was first made evident by demonstrating that a massive endolysin-β-Gal fusion supports lysis in the place of the endolysin. This indicates that during lysis, holin forms holes large enough to release molecules of ~500 kDa mass into the periplasm (39). Recently, cryoEM has been used to visualize holin lesions in cells expressing S105. Hole size ranged from ~100-1000 nm in diameter, averaging 340 nm. Up to four lesions per cell were detected, and appeared to be randomly positioned in the IM. Estimates suggest there are two holes per cell on average when taking factors such as geometry of viewing, observed diameters and locations of lesions into account. Notably, cell poles were not included in the analysis because the IM appeared discontinuous at the poles of control cells that

were not expressing the holin (40, 41). The perimeter of the holes would be expected to be lined with the holin, since a broken phospholipid bilayer would not be stable in an aqueous solution. This model is supported by cysteine-scanning accessibility data, which suggested the hydrophilic faces of S105 TMD 1 and 3 face the lumen of the hole. Importantly, data suggested that nearly all S105 molecules line the hole perimeter (42). Quantitative western blotting estimates of S105 copy number suggests there are 1000 molecules per cell. If each S105 TMD occupies ~1 nm of area, 1000 S105 molecules would account for 2000 nm of hole perimeter (1000 molecules x 2 TMDs facing the lumen). Since all S105 molecules participate in hole formation, this would account for one ~600 nm hole or two 300 nm holes, consistent with cryoEM measurements.

Although S105 is sufficient for λ lysis, the antiholin we be reviewed briefly. The ratio of S105 to S107 *in vivo* is 2:1, which is determined by the secondary structure of the *S* mRNA (43) Increasing the amount of S107 retards lysis timing (44) and the anti-lytic nature of S107 is dependent on its positively charged N-terminal extension (45). The antiholin is lytic when energy poisons are added, which indicates pmf loss is required for S107-mediated hole formation (31). Furthermore, the N-terminus of S107 must be localized to the periplasm for holin function (46). Based on these data, it was proposed that the role of the antiholin is to keep tight inhibition on S105 by forming inactive heterodimers. The extra positive charge on the N-terminal S107 results in confinement of the S107-TMD1 to interface with the inner leaflet of the IM. Once enough fully active S105 molecules accumulate and disrupt the pmf, the N-terminal TMD of S107 flips into the periplasm (Fig. A.3B) This action converts S107 from an

inhibitory molecule to a functional holin, which would increase the pool of active holin molecules the time of triggering (23).

Endolysin

Endolysins are muralytic enzymes encoded by phages that function at the second step of lysis to degrade the PG. The endolysin of λ is encoded by *R*, which accumulates as a fully active product in the cytoplasm. Endolysins have four different mechanisms of targeting linkages essential to the integrity of the cell wall. Amidases and muramidases break oligopeptide crosslinks, and glycosylases and transglycosylases target glycosidic linkages. The λ endolysin is transglycosylase that cleaves the ß-1,4 glycosidic linkage between MurNAc and GlcNAc. Despite the variety of mechanisms, complementation studies have demonstrated that endolysins will work with any holin (47). It was once thought that destruction of the cell wall was sufficient to disperse phage progeny, however a third step, OM disruption is required for lysis.

Spanins

The λ spanins *Rz* and *Rz1* are unique to biology in that both genes share the same DNA sequence and encode the same function (48). The *Rz* gene was discovered in 1979 by a transposon insertion into the λ chromosome (49). The transposon mutant was deficient for lysis in magnesium-supplemented media; however, the molecular role of Rz was not identified. In 1993, expression of *Rz* resulted in detection of an unexpected 6.5 kDa product. The gene coding this protein was found to be coded from a +1 open reading frame within *Rz* and named *Rz1*. Subsequent experimentation verified

bioinformatics predictions that Rz1 was an OM lipoprotein (50).

In the last decade, it has been shown that Rz and Rz1 belong to a diverse group of proteins called spanins, so named because the products are subunits of a complex that spans the cell envelope (51). In the absence of *Rz* or *Rz1* function, the lysis process is arrested. At the time of lysis, cells convert from rod to sphere shaped as the PG is degraded by the endolysin (22). Studies using cryoEM have shown that spherical cells formed from the failed lysis of spanin mutants are bounded by the OM (40). These cells are physiologically dead, because of IM permeabilization by the holin. These data clearly indicated that the role of spanins was downstream of the holin endolysin steps. Spanins are required to disrupt the OM at the last step of lysis.

Spanin function can be artifactually complemented by shearing forces attendant to aerated growth. In other words, shaking was sufficient to disrupt the OM after cell wall degradation. This practice obscured the role of spanins in lysis for more than 30 years. Supplementation of the media with divalent cations stabilized the cells against the shearing forces (22), most likely by stabilizing the LPS via ionic bridges formed between cations and phosphate groups on the LPS molecules.

Rz

Rz is a 153 aa IM protein of which 130 residues extend into the periplasm. The periplasmic domain of Rz is comprised of two predicted alpha helices that are predicted to form a coiled-coil structure. Rz accumulates as a dimer that is linked by two intermolecular homotypic disulfide bonds (Fig. A.4A), although the disulfide bonds are not required for Rz dimerization *in vitro* (13, 52). Secretion of the Rz periplasmic

domain into the periplasm without a TMD anchor does not support function, which indicates that the TMD of Rz is required. However, replacement of the native TMD with a FtsI did not block function, which suggests the required TMD anchor is non-sequence specific (48).

Rz1

Rz1 encodes a 60 aa prolipoprotein and mature Rz1 is anchored in the inner leaflet of the OM. Like other lipoproteins, the processing of Rz1 is encoded in the lipobox and residues in the +2 and +3 positions downstream of the lipoylated Cys20 are OM localization signals. Rz1 would be expected to undergo the following processing and transport pathway is inferred from other lipoproteins: After Rz1 passes through the Sec translocon, prolipoprotein diacylgyleryl transferase (Lgt) adds a diacyl glycerol moiety to Cys20 by thioether linkage. Then SpaseII cleaves Rz1 between positions 19 and 20 and the free amino group of Cys20 is lipoylated by apolipoprotein N-acyltransferase (Lnt). After lipoylation, mature Rz1 has three fatty acid chains that modify Cys20. Rz1 is transported to the OM by the Lol system (described above). Notably, there are no reliable structural predictions for Rz1 due to its high proline content; 10 of 40 aa in the mature Rz1 sequence are proline. Like Rz, Rz1 accumulates as a dimer, linked by homotypic intermolecular disulfide bonds at position Cy29 (Fig. A.4A and B). Converse to Rz, serine substitution at this position blocks formation of Rz1 dimers *in vitro* (52).

The spanin complex

Both Rz and Rz1 are required for OM disruption and would be expected to form a complex via C-termini. The best evidence of Rz-Rz1 complex formation *in vivo* comes from two sources. The first is the observation that expression of *Rz* in an *Rz1am* background results in instability of the Rz product, detectable by western blotting. Instability is not detected when both Rz and Rz1 are expressed together. This suggests that in the absence of Rz1, there is a protease-sensitive site that is exposed near the C-terminus of Rz. Fluorescence microscopy using a functional GFP-Rz allele provides additional evidence of Rz-Rz1 interaction *in vivo*. Expression of GFP-Rz alone shows uniform labeling of the IM, whereas co-expression of GFP-Rz with Rz1 exhibits a punctate distribution (22).

The length of the Rz and Rz1 periplasmic domains approximate the span of the periplasm: 170 residues (130 Rz and 40 Rz1) x 0.15 nm/residue, assuming alpha helical structure, equals 25.5 nm. These estimations are consistent with TEM single particle analysis of the periplasmic domains of Rz and Rz1 (denoted sRz and sRz1) purified and mixed *in vitro*. These domains formed supramolecular rod-shaped structures that were measured to be ~24 nm long and ~6 nm wide, which is in agreement with measurements of the periplasm and the width of the PG lacunae. The width of these rods is consistent with 8-10 Rz monomers. Whether the oligomers formed by sRz and sRz1 *in vitro* have biological significance has not been determined. However, there is some evidence of agreement between genetic and biochemical data: Using CD analysis, this study demonstrated increase in helicity of sRz when sRz1 was added. This provides a

plausible mechanism for Rz instability *in vivo*: Interaction of Rz with Rz1 confers order to Rz (13).

Recently the spanin complex was shown to be a heterotetramer (Fig. A.4A). Both Rz and Rz1 are homodimers linked by intermolecular disulfide bonds. Two of the disulfide linkages are in Rz (at position 99 and 152) and one is in Rz1 at position 29 (Fig. A.4A). Serine substitution demonstrated that function of the spanin complex requires either Rz_{C152}-Rz-$_{C152}$ or $Rz1_{C29}$-$Rz1_{C29}$. Intermolecular disulfide bonds are rare for bacteria. This study demonstrated that *dsb* system is important for correct registry of these homotypic linkages.

Rz and Rz1 provide a prototype for a two component spanin system. Bioinformatic analysis of 137 genomes has shown that the two-component system is the most common, but not the only form of spanin. Seven of the genomes showed evidence of a unimolecular spanin (u-spanin). Gp11 of phage T1 is predicted to be OM lipoprotein with a C-terminal TMD and five short ß-sheets within the periplasmic domain. Spanin function is complemented by expressing *gp11* in a spanin nonsense mutant background (51). This indicates that a membrane-spanning complex is a key feature for spanin function.

Membrane fusion model for spanin function

Based on similarities to class-I viral fusion systems, it was proposed that the spanins disrupt the OM by fusion to the IM (53). In this model, spanin complexes accumulate in a metastable membrane-spanning conformation (Fig. A.4). A conserved mechanism of all membrane fusion systems is a hinge-like conformational change

(discussed below). Therefore, in the extended form, premature function is prevented by intact PG. When PG is degraded by endolysin, the spanin complex becomes active to undergo a conformational change, fusing the IM and OM (Fig. A.4).

The membrane fusion model is supported by recent data that demonstrates spanins fuse spheroplasts. Spheroplasts are formed when *E. coli* cells are carefully treated with EDTA and lysozyme. This procedure results in a spherical cell bounded by the IM. Authors of this study hypothesized that if both spanin subunits were anchored to the IM, the periplasmic domains of Rz and Rz1 would be in apposition between neighboring spheroplasts. This required using missense allele of *Rz1* (denoted *imRz1*), which is retained to the IM. In this allele, substitution of positions T21 and S22 with Asp prevents Lol-uptake of imRz1 to the OM. Prior to spheroplasting, cells expressing *Rz* were also induced for expression of cytoplasmic GFP, whereas cells expressing *imRz1* co-expressed mCherry (Fig. A.5). After mixing and spheroplasting, fluorescence microscopy showed 10% of fused spheroplasts were false-colored yellow, indicating membrane fusion between spheroplasts. This demonstrates that spanins are fusogenic for spheroplasts (26). That is, fusion between two IM bilayers was demonstrated, however whether IM-OM fusion occurs as part of the normal lysis pathways is still inferred.

Membrane fusion

This review summarizes mechanisms used by virus and cell fusion proteins. Out of the three classes of viral fusion proteins, this review outlines the function of class I viral protein, emphasizing on the influenza virus model system. The rationale behind

this was largely due to similarities between these proteins and the spanins of λ. Class I viral proteins are alpha helical coiled coils, whereas class II and III exhibit beta sheet and beta/alpha secondary structure, respectively (54). Additionally, this review includes a summary of SNARE-dependent fusion; endosome fusion and neuronal vesicle fusion are chosen to be representative models for the SNARE system.

The hemifusion intermediate model is the current model for membrane fusion. This model describes the mixing of only the closest monolayers, and full fusion is complete upon the complete mixing of both bilayers (Fig. A.6A). Membrane fusion is caused by proteins, many of which are embedded in the membranes that will be fused. An alternate model for membrane fusion is the protein-pore model. This model features a protein pore which is formed and lined by the TMDs of fusion proteins (Fig. A.6B). This model is supported by the requirement of transmembrane anchors. However, the protein-pore model would predict sequence specificity for the TMD, however a wide range of TMD sequences have been shown to support fusion. Furthermore, replacement of the TMD with lipid anchors supports function in some cases (55-57). Taken together, these data challenge the validity of the protein-pore model. However, some viral proteins can fuse in the absence of TMDs (58, 59).

There are energy barriers associated with fusion (58). It is energetically unfavorable for two opposing bilayers to approach a distance of less than 30 Å due to polar head group repulsion and the hydration barrier (60). Supporting this, membrane fusion can be artificially induced by the addition of calcium (to minimizes negative charge repulsion) or PEG, which acts upon the dehydration barrier (61). Another energy

barrier is the clearance of protein from the lipid merger site, a process known as sieving (58). Additionally, energy barrier(s) associated with curvature, stalk, and pore formation must be overcome (58).

There is a large variety of fusion proteins that have distinct and structural and functional motifs. The following is a description of conserved processes required for fusion to occur: The proteins that directly work in the fusion process must tether, dock, and bring opposing bilayers close together. Fusion proteins have structure that is believed to contribute to the induction of membrane curvature. These proteins achieve this by folding and forming complexes. At some point, curvature stress is relieved by fusion-pore formation. Formation of the pore is a committed step in membrane fusion, in which aqueous contents separated by the two bilayers become mixed (58).

General viral membrane fusion model

Viral fusion proteins must be triggered to initiate the fusion cascade. Triggers include pH change, receptor binding, receptor plus co-receptor binding, receptor binding to attachment protein, and receptor plus pH change. The outcome of triggering is conserved; all triggering events initiate a thermodynamically favored event which refolds the protein into a stable post fusion confirmation (60).

Triggering leads to changes in fusion protein oligomeric state and conformation. Frequently a hydrophobic fusion peptide is exposed, which re-orients to interact with the outer leaflet of the target membrane (Fig. A.6C). Some viral proteins insert the fusion peptide in the target membrane as trimers, and others insert as monomers, which trimerize later (Fig. A.6C). Flaviviruses and Alphaviruses (class II fusion protein) insert

a single fusion peptide into the target membrane, and trimerize at a later step (62, 63). Once extended into the target membrane, the fusion protein is in a "prehairpin" conformation (Fig. A.6C). The prehairpin structure then folds into a full-hairpin conformation, which brings the target membrane and virus membranes into close proximity (Fig. A.6C). Distortions occur in the opposing membranes which deform the membranes toward each other. The source of membrane distortions is most likely from the insertion of fusion peptides or loops which exert force on the outer leaflet. Another source of distortion may be due to formation of the hairpin or interactions of the membrane-proximal region. Next, a hemifusion intermediate (also termed lipid stalk) is created by the folding back of the hairpin. A pore forms within the stalk and an irreversible step occurs, in which the viral interior is mixed with the cytoplasm. The outcome of this is delivery of the viral genome. The expansion of the pore is predicted to be the most energetically-demanding step, and likely involves several fusion proteins. While prefusion structures have some degree of structural variability, all post fusion structures are trimeric hairpins (60).

Class-I viral fusion protein

HIV fusion protein hemagglutinin (HA) is one of the best understood fusion proteins, this is in part due to the many structures for the pre and post fusion forms. HA is synthesized in an inactive form (HA0) and rearranges into a post-fusion trimer which has a central alpha helical coiled coil. HA0 is cleaved by cellular proteases. N-terminal HA1 and C-terminal HA2 (anchored in the membrane) are disulfide-linked. The HA trimer is oriented vertically on the surface of the virus, with globular HA1 "head" at the

top and HA2 connecting to the TMD. The site of HA which binds the sialic acid receptor is HA1. The HA2 stalk is triple stranded, featuring alpha helical coiled-coils. Proteolytic cleavage of HA0 frees a ~20-25 aa fusion peptide at the N terminus of HA2. The fusion peptide is then buried in HA. After the HIV nucleocapsid enters an endosome, the HA1 heads separate. This is triggered by a pH drop within the endosome. Despite separation of the HA1 heads, HA2 and HA1 remain linked by a disulfide bond. Protonation of charged residues causes the HA2 fusion peptide to be released. The fusion peptide is extended about 100 Å toward the top of the molecule, interacting with the target membrane. The HA protein then undergoes hairpin conformational change, which allows coiled-coil packing of HA2 (C-terminal region). This brings the TMD and fusion peptide to the same end as the trimer. There is evidence that suggests several HA trimers are required to mediate fusion. Notably, trimers that are not inside the contact region may have a role in expansion of the pore (60).

SNARE-dependent fusion

SNARE proteins mediate many intracellular fusion events. The SNARE complex is a stable tetra-helix bundle. The primary sequence which contributes to the SNARE motif leads to the parallel helical arrangement, the folding of which is believed to engage the fusion reaction. Most researchers agree that at least one TMD must be present in each fusing membrane, however this has recently been disputed. The SNARE complex first assumes a trans configuration, which describes SNAREs interacting from separate membranes. Upon fusion, lipids mix, and all SNAREs are localized to the same membrane (cis complex). The cis complex coincides with membrane fusion. A notable

difference from viral fusion is that SNAREs must be recycled for future fusion events. The ATPase NSF and alpha-SNAP disassemble this cis complex which frees it for the next round of fusion. Also it should be noted that SNAREs never function alone. Other factors, which bind both the SNAREs and membrane are required for efficient fusion (58).

SNARE-dependent fusion: endosome fusion

Endosomes fuse with other endosomes as well as lysosomes. Endosome SNAREs that have been identified are syntaxin 6, 13, VTI1A and VAMP4. These SNAREs are able to fuse liposomes, however are not sufficient for membrane fusion *in vivo*. Fusion of endosomes requires GTPase regulatory proteins called Rab5. Another factor which is required is EEA1, a Rab5 effector, which is a very long coiled coil dimer. EEA1 is necessary for homotypic *in vivo* endosome fusion. Key residues on EEA1 insert into the membrane which may cause curvature stress, promoting fusion. There are other factors which bind to both SNAREs and membranes which may have a role beyond tethering and targeting (58).

SNARE-dependent fusion: synaptic vesicle fusion

Vesicles released in the synapse are used for the transmission of neurotransmitters and intercellular communication. Fusion events are tied to influx of calcium ions. This is a very rapid event, occurring within 1 ms after influx. Three neuronal SNAREs are involved in synaptic vesicle fusion at the synapse. The SNARE synaptobrevin, is located on the vesicle, and SNAP25, and syntaxin-1 are localized to

the plasma membrane. These SNAREs form a tight 4-helix bundled complex (58). Complex formation brings the opposing membranes into close proximity; modeling suggests this complex would fit into a 2-4 nm gap between membranes. The mechanism from this docking step to full fusion has not been well established (64). It should be noted that understanding fusion has important implications for human health that go beyond mechanistic understanding; botulinum neurotoxins are proteases which cleave SNAREs resulting in inhibition of exocytosis (58)

Beyond providing a means for the vesicle to come into close proximity with the plasma membrane, it is believed that the energy released upon oligomerization is transduced through the TMDs into the bilayers. The function of this is destabilization of the membrane. There are other proteins (synaptotagmins and complexins) which play a regulatory role, but are not directly involved in neuronal SNARE-mediated fusion (58).

The model system used for studying neuronal SNARE function is chromaffin cells. Chromaffin cells are derived from the adrenal medulla (65). These cells are used because vesicle maturation can be easily triggered (66). Neuronal SNAREs are able to fuse *in vitro* when incorporated into artificial liposomes. This gives researchers the ability to interrogate the role of energy transduction from oligomer to TMD. High SNARE densities (non-physiological) are required to fuse liposomes. Fusion in this case is accompanied by liposome leakage. Lower SNARE densities will not mediate liposome fusion without the addition of reagents such as PEG (58).

One theory suggests that in order to transduce energy from the SNARE motif to the TMD, the linker region between them must be rigid. Mutations in this region which

code for the insertion of Pro or Gly have almost no effect on fusion *in vivo* or *in vitro*. Additionally, the replacement of the TMD with lipids which span the bilayer did not block fusion *in vitro*. It would seem that the role of membrane destabilization is unclear. It may be that other molecules assist fusion after *in vivo* after SNAREs have brought membranes into close proximity (58).

The state of SNAREs prior to fusion is an important question. Perhaps SNAREs are un-assembled prior to fusion, but zippering of the SNARE is triggered by calcium influx. In this model, energy released during the formation of the complex would be concurrent with fusion. In another scenario, other molecules provide energy for the fusion process. Current data seems to conclude a mean between the above scenarios is most likely. Work with chromaffin cells suggests that SNAREs are partially assembled in complexes prior to fusion, and that the C-terminal region of the complex assembles after calcium triggering. It is believed complexin stabilizes the partially-stabilized SNARE complex. What provides the sensor for timing of neuronal vesicle membrane fusion? The current model suggests synaptotagmin is the calcium sensor. Synaptotagmin is the only other molecule (aside from SNAREs) which has a direct effect on exocytosis kinetics (58).

Questions to be addressed

Although Rz and Rz1 are the most-studied spanins to date, there is little mechanistic information about how spanin cause membrane fusion. In chapter II, we sought genetic evidence for the putative fusion reaction during lysis. By selecting for missense mutants that block spanin function we were able to identify functional domains

within the spanin complex. We identified motifs essential in other fusion systems were also essential to Rz-Rz1 function, including coiled-coil domains, proline-rich segments, and linker regions. We also identified a salt bridge contact between Rz and Rz1, and provide evidence that complex formation is overdetermined. In chapter III, we searched for suppressors of mutations that blocked spanin function with the hope that restorative changes would provide evidence of conformational changes predicted by the spanin fusion model. Rather than identifying point-point contacts, we found that most second site suppressors clustered within a previously undescribed domain of the spanin complex. These data indicated that inactivating mutations made throughout the complex could be globally corrected by destabilizing changes within the juxtamembrane region. This suggested the importance of conformational flexibility during the fusion reaction. In chapter IV, we investigated the molecular basis of how λ lysis from the poles.

CHAPTER II

GENETIC ANALYSIS OF THE LAMBDA SPANINS RZ AND RZ1: IDENTIFICATION OF FUNCTIONAL DOMAINS*

Introduction

For Caudovirales, host lysis is essential for the release of phage progeny. For Gram-negative hosts, lysis is a three step process requiring the function of three types of phage-encoded proteins: holins, endolysins and spanins (41, 48). The first two steps, in which the holin permeabilizes the IM and the endolysin degrades the PG have been studied intensively for decades and were considered necessary and sufficient for host lysis. However, recently we have found that lysis requires the disruption of the OM by phage-encoded spanins (30) (22). In lambda the spanin complex is composed of two subunits: Rz, an IM protein, or i-spanin and Rz1, an OM lipoprotein, or o-spanin. The Rz1 gene is encoded in a bizarre genetic architecture, embedded in the +1 reading frame of *Rz* (Fig. B.1A-C). Rz has type-II IM topology (N-in, C-out), with a single N-terminal TMD and a periplasmic domain (13, 48). Rz1 is a lipoprotein, anchored in the inner leaflet of the OM (48, 67, 68). Rz and Rz1 interact via their periplasmic domains, forming a complex that spans the entire periplasm (thus the term spanin) (48). A defect in either spanin subunit blocks lysis, leaving spherical cells bounded by the intact OM (22). Two-component spanin equivalents are present in nearly all Caudovirales of Gram-negative hosts. In a minority of cases, spanin function is fulfilled by a single

protein, the unimolecular spanin or u-spanin. The u-spanin is anchored to the inner leaflet of the OM by a lipoylated N-terminal Cys residue and embedded in the IM by a C-terminal TMD anchor (51).

Molecular characterization of lambda spanins revealed that both Rz and Rz1 accumulate as homodimers linked by homotypic intermolecular disulfide bonds: both of the Cys residues (C99 and C152) in Rz are involved in these bonds, as is the single Cys residue (C29) in Rz1 (Fig. B.1B) (52). Although the Rz disulfide linkage at C99 is irrelevant, spanin function requires either the Rz C152 or the Rz1 C29 homotypic disulfide linkage. In the absence of Rz1 or in the $Rz_{C152S}Rz1_{C29S}$ double mutant, Rz undergoes substantial proteolytic cleavage. This indicates complex formation is required for stabilization of Rz in the periplasm and suggests covalent homodimerization of the spanin subunits has a role in complex formation.

Rz1 was predicted to be largely unstructured due to its high proline content (10 prolines in the 40 aa mature Rz1). By contrast, the periplasmic domain of Rz was predicted to be highly structured, dominated by two coiled-coil helical domains (Fig. B.1C) (13). Circular-dichroism (CD) studies of the purified periplasmic domains supported these predictions (13).

Recently, it was demonstrated that Rz and Rz1 can mediate fusion between two membrane bilayers, suggesting a general model in which the disruption of the OM is topological, i.e., reflecting the fusion of the IM and OM (26). It has been proposed that after PG degradation, spanin complexes undergo a conformational change that brings the opposing membrane bilayers into close proximity for fusion (26). Evidence for major

conformational dynamics was provided by CD analysis which showed that mixing the periplasmic domains of Rz and Rz1 in vitro resulted in the formation of rod-like bundles and a large increase in helical content (13). We have suggested that the coiled-coil domains participate in this conformational change, as has been shown for well-studied membrane fusion systems in eukaryotes, including the influenza virus HA2 fusion protein (69).

Lysis proteins are inherently intractable to biochemical and structural analysis, primarily due to being membrane-embedded and highly oligomeric. Thus exhaustive genetic screens are required to increase the mechanistic understanding of other lysis protein components, such as holins and pinholins (36, 70, 71). We wanted to use phage genetics to address spanin function, with the goal of isolating mutants blocking intermediate steps in the lytic pathway. A comprehensive genetic analysis of the lambda spanins has not been attempted, mostly due to the embedded genetic architecture of *Rz1* within of *Rz* (Fig. B.1A). Here we report the results of a nearly saturating mutagenesis of the lambda spanins. The results are discussed in terms of spanin function, postulated intermediate steps of the Rz-Rz1 membrane fusion pathway, and evolution of the embedded gene architecture.

Methods

Bacterial strains, plasmids, bacteriophages, growth and induction conditions

The bacterial strains, bacteriophages, and plasmids used in this study are described in Table B.1. Bacterial cultures were grown in standard LB media or, as

appropriate, in LBM, which is supplemented with $MgCl_2$ (10 mM). When appropriate ampicillin (Amp, 100 μg ml^{-1}) and kanamycin (Kan, 40 μg ml^{-1}) were also added. Growth and lysis of cultures were monitored by A_{550} as a function of time, as described previously (22). Lysogenic cultures were thermally induced at A_{550} ~0.2 by a shift to 42˚C for 15 min, followed by continued growth at 37˚C. For inductions of non-lysogenic cultures, isopropyl β-D-thiogalactopyranoside (IPTG) was added to the final concentration of 1 mM for plasmid induction. The pRE plasmid is a pBR322 derivative, which has the lambda late promoter pR' located upstream of the *Rz* or *Rz1* start site. To activate pR', the antiterminator Q is supplied *in trans* by either the induced prophage or the pQ plasmid.

Error-prone PCR mutagenesis and selection for lysis-defective Rz and Rz1

Error-prone PCR mutagenesis was performed using the GeneMorph II random mutagenesis kit without any modification to the manufacturer's instructions. To maximize the single nucleotide changes, pRz or pRz1 template DNA of higher concentration (~5 μg) was used. The Rz1 gene is inactivated on the pRz plasmid by a nonsense mutation that is silent in *Rz* (Table B.1). Oligonucleotides were obtained from Integrated DNA technologies (Coralville, IA). Mutagenized PCR products were digested with KpnI and BamHI for *Rz* and BamHI and HindIII for *Rz1*. The gel-purified, doubly-digested fragments were ligated into the pRE vector using T4 ligase and transformed into XL-1 Blue cells. After overnight incubation at 37˚C, the transformants were pooled by slurrying and plasmid DNA was extracted using the Qiagen spin miniprep kit. An MC4100 (λ) lysogen carrying the nonsense alleles Rz_{Q100am} or $Rz1_{W38am}$

were transformed with the mutagenized plasmid pool of *Rz* or *Rz1*, respectively. To assess the frequency of mutation, 10 random colonies from each library were tested for lysis defects in liquid culture and their spanin genes sequenced. Six *Rz* plasmids and four *Rz1* plasmids did not complement the lysis-deficient phenotype. Of the remaining clones, three of the latter had missense changes that did not abrogate lysis phenotype. To enrich for lysis-defective mutants, a plasmid retention method was used. Colonies on the transformant plate were collected by slurring with LB, diluted and inoculated into 25 ml of LBM and appropriate antibiotics at an initial A_{550} ~0.5, induced for lysis as described above. At 15 min past the normal lysis time (~50 min), the culture was centrifuged at 4000 rpm for five minutes to harvest the non-lysed, Mg^{++}-stabilized spherical cells, the terminal phenotype of spanin-defective lysis. The harvested culture was carefully washed once with LBM before extracting plasmid DNA using a miniprep kit. Plasmids from the enriched mutant pools of *Rz* and *Rz1* plasmids were used to transform λ lysogens of Rz_{am} and $Rz1_{am}$, respectively. Single colonies were picked and individually screened for a lysis defect by thermal induction in five ml LBM. Lysis-defective clones were sequenced by Eton Biosciences (San Diego, CA).

Detection and quantification of spanin proteins

Accumulation of *Rz* or *Rz1* gene products was assessed by Western blotting of TCA precipitates as described previously (52). Briefly, lysogens with *Rzam* or *Rz1am* mutations were transformed with the pRE plasmid carrying an *Rz* or *Rz1* allele. Approximately 50 minutes after induction, a one mL aliquot was precipitated with 10 % TCA (22). Samples were normalized to A550 units and resolved on a 16.5 % SDS-

PAGE gel. When needed, the His-tagged proteins were probed using anti-His antibody from Sigma-Aldrich.

Covariance Analysis

We identified a lambda family of embedded two-component spanin equivalents based on 40% sequence similarity over 40% of sequence length (72). From this family, we selected six representatives of the i- and o-spanin C-terminal domains.

Identification of codons a single base pair change from Pro

Using Python and BioPython (73), we developed the tool One SNP Away to scan a FASTA sequence for codons that are a single nucleotide change from the query amino acid. This program was used here to identify codons that could be mutated to proline via one mutational step. This tool is available on the Center for Phage Technology's Galaxy instance https://cpt.tamu.edu/galaxy-pub/ .

Results and discussion

Design and implementation of the mutagenesis system

Plasmids carrying Rz or Rz1 (pRz and pRz1) under the native pR' late promoter were used to effect independent mutational analysis on each gene. Both pRz and pRz1, complemented the lysis defect of lambda Rz_{am} or lambda $Rz1_{am}$, respectively (See Methods). The spanin genes were then subjected to random PCR-based mutagenesis in the context of these complementing plasmids, subcloned into a fresh vector, and transformed into a host carrying the Rz_{am} or $Rz1_{am}$ lysis-defective prophage. The pooled

transformant libraries of pRz and pRz1 were selected for lysis-defective alleles by plasmid retention (see Methods) and then screened individually for lysis after induction. A total of 266 pRz and 115 pRz1 lysis-defective clones were sequenced, of which 131 *Rz* and 79 *Rz1* genes had single point mutations. The remaining lysis-defective *Rz* and *Rz1* mutants had either multiple point mutations (48 *Rz* & 19 *Rz1*) or frameshifts, including both base insertions and deletions (87 *Rz* and 17 Rz1). Overall, the nucleotide changes were 42% transition and 58% transversion mutations, as expected for this type of mutagenesis, which indicates that the full range of possible mutations at each codon was equally accessible (74).

Mutational analysis of Rz

For convenience and to emphasize structural features, the Rz sequence (Fig. B.1C) is subdivided into six domains: the N-terminal TMD domain, NTMD (residues 1-24); the two parallel coiled-coil domains, as predicted by COILS software (75) CC1 (residues 27 - 87) and CC2 (residues 121 - 150), linkers, L1 and L2, separating the aforementioned domains and an extreme C-terminal domain (CTD).

Lysis-defective nonsense mutants of Rz

Among the 131 lysis-defective *Rz* alleles obtained by random mutagenesis, 82 alleles were nonsense changes in 34 different positions scattered throughout the periplasmic domain (L1, CC1, L2, CC2, CTD). Estimates of the Rz-Rz1 complex approximate the span of the periplasm: 170 residues (130 Rz and 40 Rz1) x 0.15 nm/residue, assuming alpha helical structure, equals 25.5 nm (76). Therefore, it is

reasonable to assume that most if not all *Rz* nonsense mutations would be lysis-defective. Since there are only 45 positions where nonsense codons can be obtained by a single base change, the random mutagenesis was estimated to be approaching saturation (34 obtained out of 45). The nonsense mutations were distributed across the entire length of Rz, except for the extreme C-terminal region, suggesting that this domain of Rz is not essential (Fig. B.2B, Table B.2).

To test this hypothesis, site-directed mutagenesis was used to introduce nonsense mutations in the last three residues. Phenotypic analysis revealed that only the C-terminal R153 residue is dispensable; nonsense mutations at positions 151 and 152 were lysis-defective (Fig. B.3 and Table B.2). Importantly, the C152S allele, which abrogates one of the two intermolecular disulfides of Rz, is functional if Rz1 retains its intermolecular disulfide linkage at C29 (52). Thus, oddly, Rz can tolerate a Cys to Ser substitution that abrogates the C-terminal disulfide bond, but not a deletion of this residue, suggesting a strict chain length requirement for the Rz periplasmic domain.

Lysis-defective missense alleles of Rz

The remaining 49 alleles were single missense changes in 34 positions in various domains of Rz (Fig. B.2A), some of which were isolated multiple times, another indication that the selection was approaching saturation (Table B.2). Overall, the missense mutations were significantly biased towards the distal half of CC1 (positions 61 through 87), where ~62% of the missense changes were clustered in only 20% (residues 57 to 88) of Rz. The mutations were more prevalent in CC1 (25 alleles in 17 positions) compared to CC2 (six alleles in five positions).

Coiled-coil domains are composed of a heptad repeat motif that typically contains bulky hydrophobic residues alternating at every first and fourth residue (the "*a*" & "*d*" positions of *abcdefg*) (75, 77). Six missense alleles were isolated in the eight "*a*" and "*d*" positions between residues 60 and 86 in CC1. Of these, three were conservative mutations (V61A, A73V, and V86G), which cause changes in the bulk of a hydrophobic side chain. This suggests that these residues are involved in inter-helical packing of two parallel CC1 domains. One stretch of 10 residues, L64 to A73, was the most mutationally sensitive region of Rz, with 13 missense alleles mapping to seven of the 10 positions, including five missense mutations at two "*a*" positions. Within this stretch, L64 and D65 were the most mutationally-sensitive positions, with six missense alleles conferring a lysis defect. Although polar or charged residues are tolerated in *a* and *d* positions it is unusual for Asp to be in the "*a*" position of a parallel helical interface (78, 79). It has been reported that Asp (not Glu) in the "*a*" position of a parallel coiled-coil creates a flexible hinge (80). Detection of three alleles at this position suggests D65 may form a junction in CC1 that is essential for spanin function.

Interestingly, when coding sequences are aligned respective to the embedded architecture of *Rz* and *Rz1*, there was virtually no overlap between the mutationally-sensitive regions of the two genes (Fig. B.2A panel ii., red rectangle), despite the fact that the mutational selections were done separately on each spanin subunit gene, with the cognate subunit supplied in trans. Only one missense change, Y127N, was obtained in this region (Fig. B.2B, positions 94 to 134). This allele product does not accumulate indicating it is unstable (not shown). The part of the *Rz1* reading frame encoding the

mature lipoprotein lies entirely within this ~40 codon region of *Rz* that is mutationally-silent. Presumably this reflects the unique evolutionary pressures extant in the embedded character of these two genes, so that no part of the nested architecture is subject to the functional requirements of both spanin subunits. This would suggest that i-spanin genes from phages with separated architectures may be free to evolve a more structurally-defined L2 region.

In contrast to the rich and diverse mutational profile of the middle region of CC1, the periphery of CC1 and the entire CC2 domain were relatively insensitive to missense changes other than helix-breaking Pro substitutions. Most of the mutations within CC2 were located at the extreme C-terminus, between residues 143 and 151, with four alleles in three positions. This finding, along with the results of site-directed mutagenesis (see below) suggests this segment of CC2 interacts with Rz1.

Phenotypic analysis of proline substitution

11 of the 34 missense mutations in Rz were Pro substitutions, including 10 in the predicted coiled-coil helices, and one in the TMD near the periplasmic interface. Given the degree of saturation, the distribution of Pro substitutions within the set of codons that can be changed to a Pro codon with a single base change (i.e., XCX or CXX) should be a good indicator of essential helical secondary structure. We used the One SNP Away tool (81) to scan for such codons (Fig. B.2A panel iii). In our screen we did not isolate Pro substitutions within 20 accessible codons between positions 23 and 63 (i.e., all of L1 and the proximal half of CC1) identified by the screen. Additionally, no prolines were identified in 22 such codons from positions 92 - 133. Conversely, proline substitutions

were isolated in nine of 13 possible positions between 64 and 93. Similarly, there were two of five possible proline substitutions identified within an eight-residue stretch of CC2. Assuming proline substitutions obtained by this selection serve as an indicator of essential helical structure, the essential Rz helices span from position ~60 to ~90 and ~130 to ~140. Using JPRED4, predictions based on primary structure find longer helices, from 27 to 87 and 121-150 (82). To gain more insight to the potential length of these helices, we selected residues Q36, A50, A62, T107, R125, and Q151 for proline substitution (Fig. B.2A panel iii, compare green and black "P" and Fig. B.4). Among these changes, mutations within the most stringently predicted coiled-coil stretches (Fig. B.2A panel iv: #1 and 2) resulted in lysis-defective alleles. Conversely, residues T107 and Q151 tolerate proline substitution, which would be expected since they are outside of predicted helices. Although Q35, A50 and R125 fall within predicted helices, proline substitutions in these positions do not inactivate Rz function, suggesting that the regions found to be proline-sensitive correspond to coiled-coils. Since these residues sample helical segments of Rz with low scoring coiled-coil prediction, it is apparent that proline substitutions are tolerated only in stretches of Rz without well-defined coiled-coil helical structure. Taken together, these data suggest that CC1 and CC2 are two regions of coiled-coil structure important for spanin function.

Role of the Rz TMD

There were only three mutations isolated in the TMD, suggesting that the TMD serves only as a membrane anchor. We tested this notion by replacing residues 5-24 with an artificial TMD (Fig. B.2B). The resulting allele was fully functional (Fig. B.3).

However, one missense allele, L19P, was isolated in the lysis-defective selection. Proline residues are generally well-tolerated in TMDs and would not be expected to abrogate membrane-anchoring (83, 84). Furthermore, of seven possible changes to proline accessible by a single-base change in the TMD region, only L19 was isolated. The codon nearest to L19 susceptible to Pro substitution with a single base change is S20; S20P was found to be functional (not shown). In another type-II membrane protein system, the position of the proline within the TMD affected integration into the membrane, with drastic differences in protein accumulation and maturation observed between adjacent mutated positions (85). Thus the proline substitution at position 19 may disrupt function by blocking proper maturation of Rz.

The Rz L2 region functions as an unstructured hinge

Rz has a predicted unstructured region between the CC1 and CC2 (L2 region, positions 89-120). The only L2 lysis-defective mutants isolated, R91P and L93S, are within a predicted ß-strand near CC1 (Fig. B.2A panel ii. and B.2B). Based on these data and class-I viral fusion models (54, 60), we hypothesized that this region functions as a flexible linker to connect two helical domains. To test our hypothesis, residues 100 to 115 were replaced by a 16-mer consisting of repeats of the Ser-Gly dipeptide sequence, corresponding to Gly-rich flexible spacers that connect domains of multi-domain proteins (86) (Fig. B.2B). As expected, the synthetic linker replacement did not abrogate spanin function (Fig. B.3). This supports the notion that the linker region of Rz acts as a hinge to bring the two helical domains of Rz into close proximity, thus resembling canonical membrane fusion systems, where two coiled-coil structures bring

the membrane bilayers into close proximity (26, 87).

Mutational analysis of Rz1

Lysis-defective mutants of Rz1

Of a total of 115 lysis-defective *Rz1* mutants, 79 had single point mutations; the rest had two or more mutations or frameshift mutations and were excluded from analysis. Initially, the degree of saturation was thought to be less than that obtained for *Rz*, because the 36 nonsense mutations were found in only 10 of the 19 codons for which a single nucleotide change could yield a stop codon. However, four such codons were in the CTD of *Rz1*, beyond the last nonsense mutation (W46X) that was obtained in the lysis-defect selection. This raised the possibility that the extreme C-terminus of Rz1 is dispensable. This notion was confirmed when each of these six distal sites was converted to nonsense codon by site-directed mutagenesis and tested for their function. None were found to have a lysis defect (Table B.2). Thus 10 of 14 potential nonsense sites were accessed in the selection, indicating the degree of saturation was similar to that obtained for *Rz*. The 43 lysis-defective alleles with single missense mutations mapped to only 14 codons of the 60 codons of *Rz1*, and seven of these mutations mapped to four positions in the signal sequence, all of which would abolish translation or processing of the precursor (88, 89). The 36 missense mutations in the periplasmic domain mapped to only 11 positions, none of which were in the first 10 residues of the periplasmic domain. Taken with the non-essential character of the extreme C-terminus, these results indicate that the central 57% (residues 32 – 54) of the periplasmic domain

comprises the key functional domain of Rz1.

Rz1 also has a periplasmic linker

Since no lysis-defective missense mutations were mapped in the first 11 residues of the periplasmic domain, we hypothesized that the N-terminal segment of the mature periplasmic domain of Rz1 could function as a flexible spacer between the membrane-attached N-terminus and the mutationally-sensitive central domain, like the linker region between the coiled-coil domains in Rz. When we replaced residues 25-30 with three Gly-Ser repeats (Fig. B.2C), the substitution allele was found to retain lytic function, supporting the notion that the role of this region is to link the central domain to the lipid anchor in the inner leaflet of the OM. It should be noted that the linker substitution also abolished the intermolecular disulfide link at position 29, which would disrupt the homodimerization of Rz1. However, this is consistent with previous findings, since spanin function is retained unless homotypic intermolecular disulfide bonds at both $Rz1_{C29}$ and Rz_{C152} are disrupted (52).

The proline rich region of Rz1 is an essential fusion motif

A striking feature of the mutational distribution, in contrast to the frequency of mutations to proline in *Rz*, is the prevalence of mutations in the Pro codons of *Rz1*. Rz1 is proline-rich, with 10 Pro residues occupying 25% of the mature sequence. Most mutants were within the Proline Rich Region (PRR) (Fig. B.1C), especially in four Pro residues in a penta-proline (P_5) stretch (Fig. B.2C), residues 32-36. Interestingly, within P_5, position 34 was not sensitive to alanine replacement, consistent with our previous

finding where an alanine substitution at position 34 did not abrogate spanin function (90).

Another lipid-anchored peptide with proline-rich motifs is the reovirus p15 fusion-associated small transmembrane (FAST) protein (91). Similar to the P_5 stretch of Rz1, p15 has a proline stretch (PPAPPP). Like Rz1, the proline-rich motif in p15 is important for membrane fusion, and the fusion reaction is not sensitive to changes in the third position. Evidence has been presented that the role of polyproline helices in membrane fusion is to promote exposure of hydrophobic side chains of neighboring regions (91).

Mutations in four other positions in the periplasmic domain of Rz1 were lysis-defective: I39V, W46R, W46C, L50P, L50R, and I54N. Of these mutations, the I39V is the most remarkable; Ile and Val side-chains are extremely similar in most contexts except for helix-helix packing (92), suggesting position 39 is involved in an intimate protein-protein contact required for spanin function. The Cys substitution at position 46 (W46C) would be predicted to result in an intramolecular disulfide bond with C29, placing a covalent constraint on the folding of Rz1 (52). Interestingly, although change-from-proline mutations dominate the mutational spectrum of *Rz1*, L50P, which creates Pro-Pro sequence in the distal region of Rz1 blocks function. This suggests a Pro-sensitive secondary structure, presumably alpha-helix, is required at the C-terminus.

To determine whether the lysis-defect of the missense mutations reflected a lack of accumulation of either spanin subunit, we collected whole-cells samples of cells expressing plasmid-borne *Rz* or *Rz1* mutants in the presence of Rz1 or Rz, respectively.

Samples were collected before lysis by TCA precipitation and examined by Western blotting (see Methods). Most of the allele products accumulated to wild type levels, indicating that the lysis defect is not due to protein synthesis or stability (Fig. B.5). For Rz, L72F, G143R, and Q145am appeared to be unstable, marked by reduced accumulation and the presence of apparent degradation products (filled square in Fig. BS.1) or smears. Surprisingly, in the presence of Rz1, Rz_{G143R} is stabilized and Rz_{L72F} is unstable. We interpret this as evidence that Rz and Rz1 form a complex in vivo during the late gene expression period, and mutant products are stabilized, in the case of Rz_{G143R}, or destabilized (in the case of Rz_{L72F}) by conformational changes associated with complex formation. The accumulation of $Rz1_{P33L}$, $Rz1_{P35H}$, and $Rz1_{P36Q}$ appeared diminished, but these mutations fall within the epitope used for immunodetection, so it is unclear if accumulation is actually affected. Importantly almost all *Rz* and *Rz1* mutant products appear to accumulate exclusively as disulfide-linked homodimers (double asterisks) with the exception of $Rz1_{W46C}$, which likely is largely blocked in an internal disulfide-bonded state (single asterisk in Fig. B.5D). Thus, the function of these defective alleles is likely blocked after their dimerization step (52).

Assessing interaction between Rz and Rz1

To test if the various Rz mutants were able to interact with wild type Rz1, we used a pull-down approach with a functional oligohistidine-tagged Rz1 as described before (48). Preliminary data suggested that a majority of Rz mutants co-expressed with Rz1-His were not defective in co-immunoprecipitation. To increase stringency of the assay, each spanin subunit was expressed in separate cultures before interrogating

complex formation in vitro with solubilized samples. Nine mutant alleles mapping to CC1 were tested and, in each case, the Rz product was found to co-purify with Rz1-His (Fig. B.6A and B), suggesting the defect imposed by substitutions at CC1 does not alter Rz-Rz1 interaction. These lysis-defective alleles of *Rz* and *Rz1* were also tested for dominance by expressing *Rz* or *Rz1* mutants from the pRE plasmid in the presence of prophage-borne wild type copy of *Rz* or *Rz1*. All of the mutant alleles tested for complex formation were also unable to block lysis (marked bold in Table B.2). The absence of dominant negative character suggests that 1) there are enough mutant-free spanin complexes present to achieve lysis or 2) hybrid complexes are not poisoned by the presence of mutant product(s).

To address the C-terminal residues involved in Rz-Rz1 interaction we used the pull-down assay to characterize six mutants in CC2, including three (Y147H, Q151R, and Q151K) identified by the screen and three alleles created by site-directed mutagenesis (E150G, Q151X, and C152X). The only CC2 mutant that did not co-purify with Rz1-His was Rz_{E150G}, suggesting this terminal Glu provides an anionic interaction partner with Rz1 (Fig. B.6C and D). The Rz_{Q151} and Rz_{C152} nonsense mutants are defective in accumulation independent of co-expression with Rz1 (Fig. BS.2), suggesting these mutants are defective in complex formation in vivo. Since abrogating negative charge at 150 blocked complex formation in vitro, we examined covariance at the C-termini of Rz and Rz1 equivalents in other lambdoid phages (Fig. B.7). This analysis suggests a linkage between Rz_{E150} and $Rz1_{R59}$; charge-to-polar changes in position E150 are compensated by complementary changes at R59 (Fig. B.7B). To

address whether an electrostatic interaction was required between the residues at this position, we tested whether the *Rz1*$_{R59E}$ allele would suppress the *Rz*$_{E150R}$ defect. Indeed, co-expression of Rz1$_{R59E}$ and Rz$_{E150R}$ complemented the lysis defect of phages carrying *Rzam*/*Rz1am* in vivo (Fig. B.8). This strongly supports the notion that the heterotypic Rz-Rz1 interaction involves a salt bridge between Rz$_{E150R}$ and Rz1$_{R59E}$.

To screen Rz1 mutants for interaction defects, we created nine mutant alleles of *Rz1-His.* Similar to Rz CC1 mutants, all products tested from *Rz1-His* mutants exhibited parental co-immunoprecipitation with Rz (Fig. B.6E and F), suggesting that such mutants are not defective in forming an Rz-Rz1 complex and are presumably defective in a step following complex formation, i.e. the fusion step(s).

As noted above, Rz1 could be truncated to position 55 without loss of function (Table B.2). This is surprising because Rz1 S55X eliminates the C-terminal residues from Rz1, including the salt bridge between Rz1 59 and Rz E150. The simplest explanation is that there are more than one residue pairs involved in Rz-Rz1 complex formation. An over-determined interaction interface between Rz and Rz1 would provide multiple points of contact that may stabilize the spanin complex.

Conclusions: coiled-coils and prolines: a novel fusion matchup

Here we report the first genetic analysis of an embedded gene pair which both genes are required for the same biological function: *Rz* and *Rz1*, which encode the subunits of the two-component spanin of phage lambda. The selection, based on a near-saturation selection for mutants that abrogated lysis, identified mutants that inactivate either Rz, the i-spanin, or Rz1, the o-spanin product encoded by the embedded gene.

The selections were done on artificially dis-embedded genes, but despite this architectural segregation, both genes exhibited mutational clustering in regions that corresponded to mutationally-silent regions of the out-of-frame gene. These mutationally-silent regions were tested by site-directed mutagenesis and found to be replaceable by simple repeated linker sequences, thereby establishing that both Rz and Rz1 have flexible linker domains between the mutationally-sensitive regions. Surprisingly the mutants that were identified by the selection, despite the loss of lytic function, uniformly maintained the ability to form spanin complexes in vitro and in vivo, and most were not defective in the accumulation of gene products. The simplest interpretation is that these mutations blocked a step downstream of periplasm-spanning complex formation. We have proposed that the complex, once liberated from the constraints of the intact PG layer, undergoes oligomerization and then causes fusion between the IM and OM (26). The pattern of disabling missense changes in both Rz and Rz1 are consistent with the notion that most of these mutations block spanin function at this putative fusion step. Importantly, the pattern of single missense mutants highlights mutationally-sensitive subdomains that resemble known fusion motifs, such as domains that are rich in coiled-coils and proline. In class-I viral fusion systems, coiled-coils promote oligomerization and conformational change from extended to hairpin structure, which pulls membranes into apposition. Future studies of mutant alleles of Rz that fall within the coiled-coil domains could determine whether function loss is at the pre-hairpin formation or the subsequent conformational change step. Another unique feature of the spanin fusion array is the PRR in Rz1. As discussed above, polyproline stretches

are key fusion motifs in reovirus FAST fusion proteins. Although there is no robust molecular model for the role of proline-rich stretches in the membrane fusion process, single missense mutants in the PRR region of Rz1 suggest a more specific role than membrane disordering. If the role of the PRR is to force exposure of hydrophobic residues, this may promote fusion by increasing contact between Rz1 and the lipid monolayer. In this way Rz1 could act as a scaffold to promote lipid curvature or to promote stalk radius enlargement, mechanisms which have been proposed in other systems (93, 94). It will be important to test these models against PRR mutants by developing an in vitro fusion system for the spanins.

The Rz-Rz1 spanin system, with its powerful genetics, may be a useful platform for the study of membrane fusion in general. For example, because the spanin-mediated fusion event would have to occur within a 25 nm space between membranes at a precise time in the infection cycle, it may be possible to capture the hemifusion state in vivo by using high resolution cryo-EM and super-resolution microscopy techniques.

CHAPTER III

SUPPRESSOR ANALYSIS OF THE FUSOGENIC LAMBDA SPANINS*

Introduction

Lysis is essential for the release of tailed phage progeny (30). Like other dsDNA phages of Gram-negative hosts, λ produces lysis proteins targeting each major component of the cell envelope (95). λ lysis genes are transcribed from the late promoter pR' (Fig. C.1) from approximately eight minutes after induction and lysis proteins accumulate throughout the morphogenesis period (96, 97). At a genetically programmed time, holins undergo rapid oligomerization in the IM forming a micron-scale hole (30, 40). This allows endolysin to enter the periplasm and degrade PG. The final step of lysis is OM disruption by the spanins, composed of subunits Rz and Rz1. When lysis is monitored on a single-cell level, rod-shape morphology is at least transiently maintained when cellular contents are released explosively (Fig. C.2, A and B) (22). This indicates that the OM is disrupted before complete PG degradation. Inactivating either Rz or Rz1 arrests lysis at a spherical-cell stage where the OM is still intact (Fig. C.2, C-E, G)(22, 98). The spanin complex is a heterotetramer, composed of Rz and Rz1 dimers that span the periplasm (hence the name spanin) and link the IM and OM (48, 52). Both subunit dimers are covalently linked by intermolecular disulfide bonds, two in Rz and one in Rz1 (52).

*Reprinted with permission from "Suppression Analysis of Fusogenic Lambda Spanins" by Cahill et al., 2017 *Journal of Virology*, 413-417, Copyright 2017 ASM journals.

The IM subunit Rz has a periplasmic domain consisting of two coiled-coil helices and forms a complex with OM lipoprotein Rz1 via C-terminal-C-terminal interactions (Fig. C.1B)(13). Based on similarities to class-I viral fusion proteins, it was suggested that spanins function at the final stage of lysis by fusing the IM and OM (48). This notion was supported by experiments demonstrating spanin-mediated fusion of *E. coli* spheroplasts (26). Recent genetic analysis identified mutants that inactivated *Rz* or *Rz1* (99). Mutations clustered within the coiled-coil helices (CC1 and CC2) of Rz and the proline-rich-region (PRR) of Rz1, suggesting these are essential to spanin function (Fig. C.1B). Coiled-coils and proline-rich regions are essential domains in other well-described fusion systems (54). The prevalence of inactivating mutations within these regions suggests that loss of spanin function occurs at the putative membrane fusion step(s), since most point mutants were not defective in complex formation or accumulation.

We sought to use suppression analysis to look for interacting domains within the spanin complex. However, *Rz1* is wholly embedded within the +1 reading frame of *Rz* (Fig. C.1A); therefore, genetic analysis required the construction of a dis-embedded gene architecture. Here we report the construction of synthetic tandem *Rz-Rz1* genes and their exploitation to search for intragenic and extragenic suppressors of spanin defects. The results are discussed in terms of a detailed model for spanin function.

Materials and methods

Bacterial growth

Strains, plasmids, and primers used in this study are listed in Table C.1. Bacterial cultures were grown on Luria broth (LB) as described (99). Phage plating was done on TB-agar plates with molten TB agar as described (98).

Synthesis of "dis-embedded" Rz and Rz1 genes

To separate *Rz1* from *Rz* on the λ chromosome, a plasmid (pAK1) was designed that carries synthetic *Rz* and *Rz1* genes in tandem architecture (^{Syn}*Rz* and ^{Syn}*Rz1)* upstream of a *bor::kan* insertion that is flanked by regions with homology to the λ chromosome (Fig. CS.1 and CS.2). Synthetic *Rz* and *Rz1* genes were synthesized by Genewiz (South Plainfield, NJ). The synthetic sequences and comparisons to the parental *Rz* and *Rz1* genes are provided in Fig. CS.2. The plasmid pAK1 carrying the synthetic dis-embedded genes was transformed into a BW25113 (λ*cI857*) in which the prophage carries the native embedded *RzRz1* arrangement. A Kan^R transformant was grown at 30°C in LB and then thermally induced to generate a recombinant lysate (as described below). The desired recombinant phages have integrated the kanamycin resistance gene from the pAK1 plasmid along with the dis-embedded ^{Syn}*Rz* ^{Syn}*Rz1* in place of the embedded *RzRz1* (Fig. CS.1). The recombinant phage was captured by infecting the MDS12 host and selecting for Kan^R lysogens as described below. Single lysogens were identified by screening lysogens using colony PCR (100), grown in LB

and thermally induced to generate a phage lysate (described in detail below). To prepare template for sequencing, primers 5'-CAGCAATATCTGGGCTTCACTGC-3' and 5'-GCAGTTTCATTTGATGCTCGATGAGTTTTTC-3' were used to amplify the spanin genes using one µL of lysate. The amplicons were purified using a QIAquick PCR purification kit (Qiagen) and sequenced at Eton Biosciences (San Diego, CA) using primers 5'CAAAATTCAAAGAAGCGGGCGGAAC-3' and 5'-GAAGCGCGTGTGTATTGCTCAC-3'. The final construct, designated λ910, carried the synthetic ^{Syn}*Rz* and ^{Syn}*Rz1* spanin genes. As expected, it was indistinguishable from wild type λ in terms of lysis and plaque formation and was used as the parental for the suppressor studies (Fig. CS.1B).

Identifying spanin mutants suitable for suppression analysis

The pAK1 vector was used as template for site-directed mutagenesis. The method described above was used to introduce spanin mutations into the prophage carrying the synthetic *Rz* and *Rz1* architecture described above. Eighteen prophages isogenic to λ910 were generated; of these, 11 were *Rz* mutants (V61A, L64H, D65G, K70E, L72F, V86G, L93S, Y127N, G143R, Y147H, and Q151R) and seven were *Rz1* (P32Q, P33L, P35H, P36Q, W46R, I54N, and E58K). Two *Rz* (K70E and Y127N) and three *Rz1* (P32Q, P36Q and E58K) alleles were not suitable for suppression analysis because the plating defect was too leaky; although these mutants block lysis of bulk culture, plating efficiency of these mutants at six hours was indistinguishable from wt λ for an unknown reason.

Isolation of plaque-forming pseudorevertants

Phage plating was performed according to standard techniques with some exceptions (101). Briefly, a 100 µL aliquot of an overnight MDS12 culture was adjusted to 10 mM $MgCl_2$ and mixed with 100 µL of a phage suspension at 10^8 PFU/ml. After 30' incubation at room temperature, this mixture was mixed with four mL of molten T-top agar supplemented with 50 mM $MgCl_2$ and immediately poured on thick TB plates. Plates were inverted after five minutes and incubated in 37°C for six hours.

Suppressor isolation and lysogen generation

Plaques were isolated by punching out single plaques using a Pasteur pipette and ejecting into 200 µL of λdil. This suspension was sterilized with 10 µL of chloroform, purified by replating twice and used to infect MDS12 as described (101). Single lysogens that formed Kan^R colonies after overnight incubation at 30°C were selected (described above). To sequence putative suppressors, one µL of the chloroform-treated phage suspension was used as template for sequencing of the spanin genes (described above).

Lysis monitoring, phenotype assessment and lysate collection

Fresh overnight cultures from a lysogenic colony were diluted 1:200 in 25 mL of LB supplemented with 10 mM Mg^{++}. This mixture was incubated at 30°C in a water bath shaker with the appropriate antibiotic. Lysis was monitored as described (22). At A550 ~0.25 cultures were thermally induced by transferring to a 42°C bath for 15 minutes and then moved to 37°C for an additional 45-65 minutes. The phenotypes of

lysogens carrying suppressor mutations were categorized by the rate of bulk culture lysis. In the case of wild type λ, lysis occurs within 10 minutes of first drop in turbidity (A550). Suppressors that lysed as rapidly as wild type were given a (+ + +) phenotype (Table C.3 and Fig. C.4). The phenotype (+ +) was denoted for suppressors that lysed with a 5-15 minute delay. Lytic cultures with a more pronounced delay in lysis were denoted by (+). When applicable, 10 mL of lysate was collected in a 15 mL conical tube 65 minutes after induction. The lysate was sterilized by addition of chloroform to a final concentration of 1%. After vortexing for 15 seconds, the lysate was centrifuged for 10 min. at 6000 x g to pellet chloroform and cell debris. The sterilized supernatant was transferred to a 15 mL conical tube and stored at 4°C.

Evaluating lysis morphology using time-lapse phase microscopy

Single cell lysis was monitored by phase time lapse microscopy as described with modifications (22). Briefly, a 1.5 µL sample was taken prior to lysis, applied to a glass microscope slide, covered with a 22 x 25 mm coverslip and imaged immediately using the 10x 0.25 NA phase objective on a Zeiss Axio Observer A1 microscope, with the stage preheated to 37°C. Images were captured every 250 ms for 28 minutes, or until all cells within the field of view lysed. Time lapse series were saved as .MOV files rendered at 4 frames per second with a timestamp. To evaluate the lysis defect of spanin mutants, the timestamp was used as reference to calculate the lysis kinetic of single cells by the difference in time between the first frame showing cell deformation to the first frame showing breach of cell envelope (loss of phase halo). Representative time lapse images in Fig. C.5 were captured with an 40x 0.75 NA objective.

Sample collection for western blotting

A one mL aliquot of whole cell sample was taken 45 minutes after induction, precipitated by trichloroacetic acid (TCA) and prepared for western blotting as described (13, 52). Briefly, TCA pellets were resuspended in 1x SDS-PAGE sample loading buffer in a volume normalized to the A550 units recorded at the time of collection. Samples were boiled for five minutes before resolving on a Novex Wedgewell 4-20% Tris-Glycine SDS-PAGE gel (Thermofisher). Western blotting with the Rz and Rz1 antibody were performed as previously described (48). Densitometry was performed with ImageJ. The loading control band and primary Rz band was used for area measurements. The reported values are normalized to the loading controls.

Results

Separation of spanin genes does not change expression or lysis

In wild type λ, *Rz1* is entirely embedded within the +1 reading frame of *Rz* making genetic analysis of either problematic. To address this, we constructed λ phages with the *Rz1* gene downstream of *Rz* (Fig. CS.1). Moreover, to prevent the separated genes from undergoing recombination within the shared sequence, we altered the codons of *Rz* and *Rz1* to be distinct without changing the coding sequence of either gene (Fig. CS.2). In this arrangement, the DNA sequence of the separated *Rz1* gene is 36% identical to the region of *Rz* that formerly housed *Rz1*. We did not observe any difference in the phenotypes of phages carrying separated spanin genes, i.e. the rate of lysis (Fig. CS.2B), plaque size (Fig. CS.2C), and gene expression (Fig. CS.2D) is

unchanged. This suggests the embedded gene architecture reflects evolutionary pressures to save genomic space.

Spanin-defective λ can be differentiated from wild type λ by plaque assay

For unknown reasons, *Rz* and *Rz1* lysis-defective alleles do not have a strong plating phenotype unless the agar is supplemented with millimolar concentrations of divalent cations. We found that in media supplemented with 50 mM $MgCl_2$, the parental λ910 forms a clear plaque approximately six hours after incubation, whereas no plaques are formed by isogenic missense or nonsense mutants of *Rz* or *Rz1* under these conditions. Therefore, our idea was to construct phages carrying defective missense spanin alleles and select for spontaneous suppressors at six hours. For all the mutants that yielded suppressors, ~10^7 phage particles carrying a single ^{Syn}*Rz* or ^{Syn}*Rz1* mutation resulted in the formation of 1-10 plaques (Fig. C.3B i and ii).

Suppressors of spanin mutants

After purification, the *Rz-Rz1* genes of the plaque-forming phage isolates were sequenced, resulting in the identification of 41 suppressor alleles. Of these, 30 were suppressors of *Rz* mutations and 11 were suppressors of *Rz1* mutations (Table C.2). The majority (33) were second-site suppressors, which approximates the rate observed in other phage systems (102). We created lysogens to test whether second-site suppressors support lysis in liquid media. In order to avoid recombination between λ and other prophages within the *E. coli* genome, we used MDS12, a prophage-free host (103). We found that 19 (58%) of these lysogens were able to lyse cells as rapidly and completely

as wild type (Table C.3). In general, the lytic phenotype of the suppressors of *Rz1* mutations was similar to wild type. Conversely, 14 suppressors of *Rz* or *Rz1* mutations lysed at a slower rate. Of these 14, eight had a lysis rate that was 5-15 minutes slower than wild type. The remaining six had a more severe delay in lysis (Table C.3). Some examples of lysis phenotypes are shown in Fig. C.4.

Alleles with severe accumulation defects require same-site suppressor mutations

We isolated seven same-site suppressors, including true revertants of Rz_{L64H}, Rz_{L72F}, Rz_{L93S} and $Rz1_{L50P}$. Two positions generated pseudorevertants: Rz_{G143R} to Leu and $Rz1_{W46R}$ to Gly and Ile. In both of these cases a mutation involving a substitution with a basic residue was corrected by same-site restoration of an uncharged residue. A simple explanation for positions that did not generate second-site suppressors is that the original inactivating substitution conferred a high degree of proteolytic instability to the allele product, presumably because of a defect in formation of a secondary or tertiary structure. Supporting this idea, Rz_{L72F} and $_{G143R}$ have been shown to be severely defective in accumulation (99). Rz_{L72F} and Rz_{G143R} affect the *a* and *c* positions of coiled-coils within the CC1 and CC2 (Fig. C.1B), respectively, indicating that the stable coiled-coil is critical to avoid periplasmic proteolysis of the Rz subunit.

Second-site suppressors cluster at the juxtamembrane region of Rz

Out of 33 second-site suppressors, 28 were within the juxtamembrane region (residues 20-47) of Rz (Table C.3). This suppressor clustering is interesting in light of the rather disperse location of the lysis-defective mutations which range throughout

every subdomain of the spanin complex, from positions 61-151 in Rz and 33-50 in Rz1 (Fig. C.1B). Seven suppressors (S20P, A30T, A30V, Y33H, R37C, R37H, L44M) were non-allele specific, e.g. Rz_{Y33H} suppresses Rz_{V61A}, Rz_{Y147H}, and $Rz1_{P35H}$ (Table C.3). Additionally, changes at four positions were able to suppress multiple mutants despite different chemical properties of the suppressing residue, e.g. Rz_{R37C} suppresses Rz_{V86G}, $Rz1_{P35H}$, $Rz1_{I54N}$, and Rz_{R37H} suppresses $Rz1_{P35H}$ and $Rz1_{I54N}$ (Table C.3). In the context of extragenic suppressors, a pattern of compensatory changes that is not allele-specific suggests bypass suppression (104).

Lysis by suppressor alleles is preceded by a spherical cell morphology

By phase-contrast video microscopy, striking kinetic defects can be readily observed at the single-cell level for spanin alleles that are normal in terms of bulk lysis and plaque-formation (52). In the wild type context, lysis is complete within a few seconds, whereas cells expressing spanin mutant alleles form a spherical intermediate prior to lysis (Fig. C.2 C-F). In such cases lysis is delayed for more than 10 minutes (52).

We wished to investigate the extent to which suppressors restore the normal kinetics of morphological change, i.e. with or without forming the spherical intermediate. To monitor lysis at the single cell level, we recorded time from deformity (indicating the start of lysis) to overt cell breach (loss of refractility). We focused on three suppressors, the Rz_{Y33H} suppressor of $Rz1_{P35H}$, the $Rz1_{I39F}$ suppressor of Rz_{D65G} and the Rz_{Y26S} suppressors of Rz_{Q151R}, all of which exhibit normal lysis in bulk culture. In each case, suppressor-mediated lysis was delayed and progressed through a spherical

morphology before lysis. The extent of the delay varied between alleles; for example, the suppressor of $Rz1_{P35H}$ lysed within ~60 seconds, whereas the suppressors of Rz_{D65G} and Rz_{Q151R} were slower to lyse (Fig. C.5). Formation of a spherical intermediate prior to lysis indicates that suppressors are functioning after complete PG degradation. This is in contrast to wild type lysis, in which spanin function occurs when the cell is rod-shaped and thus the PG is largely intact (Fig. C.2 A-B) (Fig. C.5B). The delayed lysis of this subset of suppressors implies the majority also function after PG degradation. At a molecular level, this suggests spanin complexes are free from PG confinement and are stalled at a fusion step.

Suppressors cluster within the core of the Rz coiled-coil and reduce stability

Most of the juxtamembrane suppressors, i.e., the suppressors near the membrane interface of Rz, were polar substitutions at *a* and *d* positions of the predicted Rz coiled-coil (Fig. C.6). The substitution of polar residues (including Cys) into the hydrophobic core of a dimeric coiled-coil would be expected to reduce stability (105-107). Allele-specific proteolysis of Rz has been reported previously (52, 99), suggesting that perturbations in the juxtamembrane coil would be detectable by western blotting. We selected a point mutation from each spanin subdomain (Rz_{D65G}, Rz_{Q151R}, and $Rz1_{P35H}$) to investigate stability of eight juxtamembrane suppressors. In each case, the Rz product was decreased or Rz-specific fragments were generated (Fig. C.7). For Rz suppressors, densitometry of the ~36 kDa species showed diminished accumulation compared to the parental (Table C.4). This supports the notion that such suppressors are causing a local disruption in Rz structure within the juxtamembrane region. Further, there was no

correlation between these accumulation characteristics and the rate of lysis of bulk culture (Fig. C.7, compare "Lysis" to band pattern). One of the most frequently isolated suppressors was S20P, which falls within the predicted TMD. Previously, we showed that replacement of the native Rz TMD with an artificial TMD had no effect on Rz function, strongly indicating that the Rz TMD acts as a non-specific membrane anchor (99). It has been demonstrated that proline introduces flexible kinks within TMDs (108, 109). In the context of Rz, flexibility within the TMD may cause improper registry between the dimeric Rz coils. Further, the $Rz_{D65G;\ S20P}$ appears proteolytically unstable, like core-targeting suppressors (Fig. C.7). This suggests S20P suppresses in a manner similar to other juxtamembrane suppressors, presumably by altering the stability of the Rz juxtamembrane coiled-coil.

The only non-juxtamembrane suppressors that support rapid lysis of bulk culture were the intragenic $Rz1_{P56L}$ suppressor of $Rz1_{I54N}$ and the intergenic $Rz1_{I39F}$ suppressor of Rz_{D65G} (Table C.3). For P56L, a compensatory change at this position is surprising, given the fact that truncation of six residues (55-60) from the C-terminus of Rz1 does not block lysis (99). A simple explanation is that a polar substitution at position 54 is compensated by an increase in hydrophobicity at position 56, suggesting that non-polar character at the C-terminus of Rz1 has a role in spanin function. Suppression of Rz_{D65G} by $Rz1_{I39F}$ confers proteolytic instability on the Rz product (Fig. C.7, lanes 3 and 8, Table C.4). The $Rz1_{I39F}$ product was not detected, but the significance of this is uncertain since the I39F substitution lies within the epitope region. The slight reduction in accumulation of Rz_{D65G} in the presence of $Rz1_{I39F}$ suggests instability in Rz is

introduced upon Rz-Rz1 binding. As in the case for $Rz1_{I39F}$, the accumulation of Rz is reduced for other suppressors of D65G (Fig. C.7, compare lane 8 to lanes 5 and 7, Table C.4).

Discussion

We sought to use suppression analysis of single missense mutants that inactivated spanin function to gain insight regarding our operational model for spanin-mediated lysis; i.e., fusion of the IM and OM. Intragenic suppression analysis would be expected to yield three categories of second-site compensatory changes, altering protein interaction, amount, or activity (110). The goal of most suppression analyses is to isolate interaction suppressors aimed at gaining structural information, i.e., point-to-point contact between interaction sites or conformational states, based on allele-specific suppression analysis. This would offer a promising tool to probe intermediate steps inferred by pre- and post-fusion structures. However, suppression analysis is rarely used to study membrane fusion, due to the intractability of most model fusion systems to forward genetic analysis. An example of such analysis has been conducted on the alphavirus fusion protein. In this study, a variety of second-site revertants were isolated that rescue a mutation within the ij loop of E1, implying functional contacts during structural rearrangement (111, 112). To our knowledge, there have been no such investigations for class-I viral fusion proteins.

Reversion analysis of spanins required the construction of phages carrying a separated gene arrangement so *Rz1* was not embedded within *Rz* (Fig. CS.1). In this synthetic arrangement, the *Rz* and *Rz1* codons were altered to prevent recombination;

however, the amino acid sequence of each gene was not changed (Fig. CS.2). Despite changes in the DNA sequence of each gene, there was no consequence to λ carrying the separated genes in terms of lysis phenotype or gene expression. This suggests the embedded arrangement may improve fitness by other means, e.g., by saving genome space or preventing segregation of interacting domains. Since there are many examples of phages with separated spanin genes, we would speculate that phages with larger genomes carry separated spanin genes more frequently.

We conducted suppression analysis with the goal of identifying contacts between Rz and Rz1 and subdomains. Such information could be used to construct a structural model to provide evidence for steps inferred from fusion systems, e.g. a hairpin conformational change. However, almost all suppressors were clustered in a coiled-coil domain near the outer leaflet of the IM (the juxtamembrane region). The pattern of suppression is striking, given these changes compensate for inactivating mutations spanning all the subdomains of the spanin complex. Further, many suppressors exhibit global suppression (i.e., suppressed mutations in multiple sites), which ruled out the notion of direct suppression (110). Within the juxtamembrane region, suppression resulted in mostly polar substitution within the hydrophobic core of a predicted coiled-coil domain, and in each case there is evidence of reduced Rz stability.

Mutation within juxtamembrane segments of other fusion systems

Taken together, these data suggest that disruptive changes in the juxtamembrane region of the proximal coiled-coil domain of Rz can suppress inactivating mutations that map essentially everywhere else in the spanin complex. Insight into the basis for the

global effect of such suppressors can be gleaned from analysis of the membrane fusion proteins of eukaryotic viruses. For example, mutations near transmembrane-adjacent coiled-coils of the paramyxovirus SV5 F protein resulted in destabilization along with hyper-fusogenicity. It was suggested destabilization decreased the energy required to engage a fusion-active conformational change (113, 114). Likewise, mutations within the coiled-coil domain of the F protein near the fusion peptide destabilized the coiled-coil and enhanced fusion (115). The SNARE system has also been used to study the role of juxtamembrane composition. Fusion efficiency was increased by the insertion of two prolines into the juxtamembrane region of the v-SNARE, suggesting release of helical strain promotes fusion (116). In another study, the juxtamembrane region of the t-SNARE was reported to be a fusion switch, which was intrinsically unstable. It was suggested this switch was activated upon conformational change, while forming helical structure via zippering (discussed below).

Suppression model: kinetic penalty for increased conformational freedom

Although there is significant variety in the composition of membrane fusion systems, a consistent theme is that structural rearrangement of the protein complex pulls membranes together (55, 58). It has been suggested that the free energy associated with oligomerization and conformational change overcomes the electrostatic repulsion of apposing membranes. Mutants of Rz and Rz1 that inactivate spanin function without disrupting complex formation were suppressed by substitutions that confer proteolytic instability on Rz, presumably as a result of destabilizing the juxtamembrane segment of the proximal coiled-coil domain (Fig. C.6A). Like other membrane fusion systems, we

suggest that the spanins assume a hairpin conformation that drives the merger of two membranes. For viral and SNARE system(s) a "zippering" model has been proposed to describe the progressive assembly of helical coiled structure proceeding toward membrane anchors, a process that overcomes the energy barrier to fusion (117-119). If a majority of inactivating point mutants block lysis by disrupting the zippering reaction, it is tempting to speculate that suppression results in modification of juxtamembrane coiled-coil structure that allows the zippering process to restart (Fig. C.8). In this model, loss of structure within the juxtamembrane region increases the degrees of freedom sampled by Rz-Rz1 complexes once PG has been degraded. Therefore, mutations that stall zippering could be bypassed by a decrease of restraint within the juxtamembrane coil structure (Fig. C.8D and E). Alternately, suppression may increase membrane scaffolding interactions within the juxtamembrane region. There is a correlation between membrane curvature and formation of fusion intermediates (55). Therefore, structural interruptions that result in increased Rz-membrane contacts (by exposing hydrophobic or positively charged residues) would enhance the curvature required to facilitate membrane fusion. Future biochemical and structural studies will be required to test such models.

CHAPTER IV

THE LAMBDA HOLIN CAUSES POLAR LYSIS

Introduction

Phage lysis is the most frequent cytocidal event in the biosphere (120). The Caudovirales (tailed phages) of Gram-negative hosts require three classes of lysis proteins directed to each component of the cell envelope (23). The holin forms holes in the inner membrane (IM), the endolysin degrades the peptidoglycan (PG) and the spanins disrupt the outer membrane (OM) (Fig. D.1A). In the λ infection cycle, these lysis proteins accumulate during the late or morphogenesis period beginning at about 8 min, resulting in lysis at 50 min, a time genetically programmed into the holin, S105 (11, 23). S105, so named because it is a 105 aa product of the *S* gene, is the prototypical class-I holin, an IM protein with three transmembrane domains (TMDs) disposed in an N-out, C-in topology (Fig. D.1A) (30). Prior to lysis, the S105 holins accumulate as dimers (36), freely mobile and uniformly distributed throughout the IM (37). The timing of lysis is allele-specific to the *S* gene and varies dramatically with single missense changes, especially at positions within the TMDs. It has been proposed that lysis begins when the holin reaches a critical concentration that nucleates oligomerization into two-dimensional aggregates or "rafts" (39). This model has precedent in the switch that results in bacteriorhodopsin forming the purple membrane arrays (121, 122). Holin rafts have been detected by GFP-fusion studies (37). In the current model for lysis, the rafts are thought to mediate a collapse in the membrane potential, which in turn promotes quaternary rearrangement into lethal lesions, or "S-holes" (23). This aggregation/hole-

formation process, unique to holins, had been termed “triggering” (30). Cryo-electron microscopy and tomography studies revealed that the typical infection cycle results in ~2 holes of micron-scale, distributed randomly in the IM (37, 40, 41). Endolysin-mediated degradation of the PG requires holin function, presumably because the endolysin molecules, which accumulate in the cytoplasm with full muralytic activity, gain access to the PG when released through these holes into the periplasm. (Fig. D.1Aii). Triggering can be prematurely engaged by sudden loss of pmf, e.g. by energy poisons or an abrupt shift to anaerobic growth (30, 123). Experiments using flagella-tethered *E. coli* demonstrate that collapse of the pmf precedes lysis by ~20 seconds (35).

After PG degradation, the last step of lysis in λ infections is destruction of the OM (48, 120). This is accomplished by the spanin complex, composed of subunits Rz, an integral IM protein, and Rz1, an OM lipoprotein (Fig. D.1A) (48, 99). Spanin complexes involving C-terminal interactions between Rz and Rz1 bridge the entire periplasm and thus, during accumulation throughout morphogenesis, must be threaded through the PG meshwork. Destruction of the PG is proposed to result in spanin activation, completing the lysis pathway (Fig. D.1A ii-iii, B) (120). In most cells, lysis is observed as explosive events in which the cellular contents escape from a cell pole, resulting in the generation of non-refractile ghosts that at least temporarily retain rod shape (120). Missense mutations in the spanin genes abolish lysis without affecting accumulation of either subunit or spanin complex formation (99, 124). At the end of infection cycle in such spanin mutants, the infected cell undergoes a transition from rod shape to spheres, bounded by an intact OM(Fig. D.1B ii a-c) (125). The spanins are

similar to class-I viral fusion proteins and recent experiments have demonstrated spanins are fusogenic (26). Based on these findings, it has been proposed that spanins disrupt the OM by fusing the IM and OM (Fig. D.1A iii).

Here we report studies of the terminal lytic event at the single-cell level and focus on the kinetics and subcellular localization of the lysis proteins. The results are interpreted as a detailed model for the λ lytic pathway.

Methods

Strains, bacteriophages, plasmids and growth conditions.

The *Escherichia coli* K-12 derivative MG1655 *laci*^{q} Δ *lacY* was used as a host in this study. Bacteriophages, plasmids, and primers used are listed in Table D.1. Overnight cultures were made in LB supplemented with 100 μg/mL ampicillin, 40 μg/mL kanamycin, 10 μg/mL chloramphenicol when appropriate. Growth of cultures and lysis profiles were monitored as described previously (99). Briefly, overnight cultures were diluted 1:200 in LB supplemented with the antibiotics above and 10 mM $MgCl_2$ when appropriate or unless otherwise indicated. All cultures were maintained at 30 °C until induction. Lysogens were thermally induced at A550 ~0.3 by aerating at 42 °C for 15 min and continued growth at 37 °C until lysis. When indicated, isopropyl β-D-thiogalactopyranoside **(**IPTG) was added at the time of induction at a final concentration of 1 mM. For induction of *phoA-R* under araBAD control (pSec-R), arabinose was added at a final concentration of 0.4% at 46 min after induction.

DNA manipulation

Plasmid DNA isolation, PCR, amplification, transformation, subcloning, and sequencing were performed as described previously (99). Synthetic gene fragments (gBlocks) were obtained from IDT (Integrated DNA Technology, Coralville, IA). Restriction modification was performed using enzymes purchased from NEB (New England Biolabs) according to manufacturer's specifications.

Phase contrast time-lapse microscopy

All samples were collected for imaging at 45 min after induction except for Ethylenediaminetetraacetic acid (EDTA) experiments. Subcultures used for the EDTA lysis assay were not supplemented with $MgCl_2$. At 42 min after induction, five mL of 0.5 M EDTA were added to 20 mL of aerating cultures. Shaking was continued for five min and cells were imaged immediately. All samples were prepared and monitored the following way: A 1.5 µL sample was placed on a glass slide and covered with a coverslip. Cells were imaged immediately using a plan-apochromat 20x/0.8 Ph2 objective installed on a Zeiss Axio Observer 7 inverted microscope. Time-lapse video was captured at 0.5-5 frames per second for 10 min or until lysis or shape change was complete. When cells were in different focal planes, Z-stacks were included in the time-lapse program to maintain focus on all cells. To maintain a stable temperature of 37 °C on the mounted slide, a Lab-TekTM S1 heated insert was pre-heated to 42 °C and temperature control was maintained for the duration of the experiment. Video scaling, editing, measurement, and image export was done using the Zen 2.3 software. The

subcellular site of blebbing and lysis were scored based on observed morphological changes by the following criteria: The long axis of cells was divided into five equal compartments. Using this convention, cells have two polar compartments, and three compartments within the sidewall of the cell (two parapolar and one midcell).

Fluorescence microscopy

Samples were collected and prepared for imaging as described above, except cells expressing *S105-GFP* were collected at 49 min after induction. For experiments monitoring Thioflavin-T (ThT) fluorescence, ThT was added from a 1000x stock at the time of thermal induction at a final concentration of 10 µM. ThT was purchased from Sigma-Aldrich (St. Louis, MO); stocks were made in filter-sterilized water and kept at -20 °C prior to use. ThT time-lapses were captured at 2 frames per second using filter cube 91 HE and 20 ms exposure time at 50% light intensity. GFP time-lapses were captured at 1-3 frames per second using filter cube 90 HE and 40 ms exposure time at 50% light source intensity. Brightness and contrast were adjusted with Zen 2.3 software setting the area outside of cells to be the background. Raft number and location was determined by foci apparent in the GFP channel. Rafts were determined to be polar, parapolar, or midcell-localized by dividing the long axis of the cell into five equal compartments (described above, and shown in Fig. DS.1).

Western blotting

Protein samples were collected as described previously (124). Briefly, a 1 mL aloquot of whole-cell sample was taken at 45 min after induction. Samples were precipitated

by tricholoroacetic acid (TCA) and pellets were resuspended in sample loading buffer normalized to the A_{550} units at the time of collection. After 5 min of boiling, 0.3 units of sample were resolved on a Novex Wedgewell 4-to-20% Tris-Glycine SDS-PAGE gel (Thermo Fisher). Gel transfer and immunodetection were done using the iBlot and iBind systems (Thermo Fisher) according to manufacturer's recommended protocol.

Results

Frequency of polar lysis

Previously, lysis morphology of induced λ lysogens in an MC4100 host were monitored by phase time-lapse microscopy. Out of 40 cells, 85% were found to rupture from the poles (120). Here, we repeated this experiment using isogenic lysogens (λ900) within a MG1655 host and generated similar data; 75% of cells lysed from the poles (n = 134) (Fig. D.2A, Movie 1, movies can be found as separate files that accompany this document). To narrow down which lysis protein(s) were directing lysis at the poles, we wanted to investigate what part of the cell the first steps of the lysis reaction occurred. To do this, we eliminated the last step of lysis, which is OM disruption by the spanins, by using λ900$Rz_{am}Rz1_{am}$ lysogens. These lysogens were induced and monitored at the time of lysis. Cells converted from rod- to spherical-shaped as PG support was lost. Notably, shape conversion begins from one pole and progresses to the other in 34 out of 40 cells (Fig. D.2B Movie 2), consistent with previous findings in the MC4100 host (120).

We envision two models to explain the polar basis of shape loss. One model is that the endolysin begins PG degradation nearest the site of release into the periplasm. Another less parsimonious explanation is that the endolysin preferentially degrades PG at the poles. Polar preference for the endolysin was tested by releasing the endolysin into the periplasm using the secretion system. If secreted endolysin shows shape loss beginning from the pole(s), morphologically similar to Fig. D.2B, it would suggest the endolysin prefers polar substrate. On the other hand, if shape loss is uniform, it would suggest endolysin is not biased to the poles and degrades PG nearest the site of release into the periplasm. To address this, we constructed an endolysin allele encoding the *phoA* secretion signal, *SS-phoA-R,* under control of the arabinose promoter (denoted pSec-R). As these cells convert from rod to spherical-shape, the long axis contracts while the short axis expands uniformly (Fig. D.2C, Movie 3). This suggests that the endolysin is degrading PG in a uniform manner. This finding is consistent with an *in vitro* study that used a muramidase to degrade purified *E. coli* sacculi; results demonstrated a homogenous susceptibility to PG degradation (126). Shape transition was polar when the endolysin was released by the holin, and non-polar when endolysin was secreted by the translocon with no holin present. These data suggest that the endolysin and spanins are not the cause of the polar lysis phenotype. Further supporting this model, cells expressing the pinholin-SAR (Signal Anchor Release) endolysin system of lambdoid phage 21 show spheroid morphology during lysis. In contrast to λ, the endolysin of phage 21 accumulates in an inactive form, tethered to the IM with its muralytic domain in the periplasm. Triggering by the pinholin creates nm-scale holes

that lead to the release and activation of the SAR-endolysin. Hybrid λ lysogens carrying the lysis system of phage 21 (λhy21) lyse as prolate shapes, usually from the area between the midcell and the poles. Shape conversion of λhy21 ($Rz_{am}Rz1_{am}$)^{21} was not polar. This supports the notion that where endolysin is released controls the site of lysis. Taken together, this suggests that polar lysis is caused by hole formation at the poles by the holin.

Lysis with EDTA shows sites of hole formation

The holin initiates the start of lysis by forming holes in the IM. In a previous study, images of lysogens expressing *S105-GFP* showed the holin forms foci (rafts) coinciding with the time of lysis (37). After triggering, rafts are assumed to transition into a hole-forming arrangement, with S105 molecules lining the perimeter of the membrane hole. Using cryoEM tomography, holes were found to be randomly distributed throughout the IM (37, 40, 41). The major drawback of this technique was that imaging was limited to the sidewalls of the cells; therefore, any holes formed at the poles of the cell could not be detected. Furthermore, both of the studies above used endolysin-deficient cells, making it impossible to correlate the site of lysis with holin activity. In our model, the holin controls the site of lysis by forming more (or larger) holes at the poles. Therefore, we wanted to correlate where holin lesions form with the site of shape loss in λ900$Rz_{am}/Rz1_{am}$ lysogens. We hypothesized that weakening the OM would allow us to localize where holes are being formed by the holin.

The OM is stabilized by the presence of divalent cations and EDTA reduces OM strength by sequestering cations from the OM. Thus EDTA would be expected to

complement lysis mediated by spanin-deficient phage. Therefore, we wondered if we could localize the site of nascent PG degradation in conditions where the OM is weakened. Would we be able to see blebs corresponding to where the endolysin is released? If distortions in cell shape were detected, could we discern whether they correspond to the holin forming holes in the IM? λ900Rz_{am}/$Rz1_{am}$ lysogens were thermally induced and EDTA was added to the culture five min prior to when the rod-to-sphere shape conversion occurs. After incubation with EDTA, an aliquot of cells was withdrawn and a random field was monitored using phase time lapse microscopy. The prediction is that if holes formed are formed in the sidewall of the IM, we would expect the sidewall proximal to give a mark of initial hole formation (Fig. D.3A). As expected, blebs formed and swelled prior to eruption of cell contents (Fig. D.3B, Movie 4). As a control to test whether bleb formation was a product of holin activity, lysogens (λ900S_{am7}) carrying a nonsense mutation in the holin gene were monitored after the same treatment. In this case, no morphological changes were observed (n = 35), suggesting blebs required hole formation in the IM, caused by the holin. Similarly, we used λR_{am} lysogens, to test whether blebs required endolysin activity. As expected, no blebs were detected from the λR_{am} control (n = 41), suggesting blebs are a product of endolysin release, mediated by the holin.

The data above are consistent with the idea that using EDTA to weaken the OM in infections lacking spanins, blebs form opposite the S-hole in an endolysin-dependent fashion. The simplest notion is that local release of endolysin causes PG lesions that lead to blebs from cytoplasmic extrusion. Is all bleb formation entirely constrained to

the poles, or do blebs form first at other sites in the cell? To address this, we monitored morphological changes over time by λ900Rz_{am}/$Rz1_{am}$ lysogens treated with EDTA as described above. Out of 44 cells, 70% of cells formed blebs at the poles before lysing (Fig. D.3B, Table D.2, Movie 4 and 5). Lysis morphology remained predominantly rod-shaped, as by wt lysogens in normal conditions (Table D.2). Some cells (n=15) showed bleb formation in the sidewall of the cell along with polar blebs. Sidewall blebs were smaller and did not cause overt changes in rod cell shape. Detection of blebs in the sidewall is consistent with previous reports of rafts and IM holes in this compartment (37, 40). Some cells (n=13) converted to spherical shape before lysing, as observed with cells that are not treated with EDTA (Movie 6 and compare the third set in Fig. D.3B to Fig. D.2B). Importantly, these cells lost rod shape uniformly, without inflating from either pole. The simplest explanation of this is that the holin lesions are subject to some amount of stochasticity in terms of number, distribution and size. A single hole would be expected to release endolysin into a single spot on the sacculus, proximal to where the hole formed. Conversely, a polydisperse distribution of holes would be expected to expose endolysin to a greater area of the sacculus at once. The later scenario may allow rod-to-sphere shape conversion even when the OM is weakened by EDTA.

After lysis, cell debris of the ruptured "ghost" cells remain. During lysis in normal conditions, wt λ lysogens rupture from the poles and produce rod-shaped ghosts. We analyzed 122 ghosts of λ900Rz_{am}/$Rz1_{am}$ lysogens treated with EDTA. All ghosts had at least one polar bleb and most ghosts (82%) were rod-shaped (Fig. D.3C). Most cells (71%) had 1-2 blebs and 11% of cells had a bleb in the sidewall in addition to a polar

bleb. In both time-lapse and ghost analyses of λ900$Rz_{am}/Rz1_{am}$ lysogens treated with EDTA, percentage of polar blowouts was consistent with the percentage of polar lysis of wt lysogens under normal conditions. Taken together, it is clear that EDTA-treated lysogens expressing only the holin and endolysin form large blebs at the poles prior to lysis. Blebbing is the result of PG degradation by endolysin and the site of endolysin release is controlled by the holin. Therefore, the simplest explanation for polar blebbing is that the holin is forming larger or more holes at the poles of the cell.

Thioflavin-T labeling indicates IM permeabilization before lytic blowout

It was shown previously that loss of pmf preceded lysis by about 19 seconds (see introduction). The time between pmf loss and lysis can be shortened by increasing the amount of endolysin in the cell (35). This indicates that during a significant portion of this time interval the holin is producing holes large enough to release endolysin (larger than 4 nm). Therefore, it was important to know when IM holes form prior to lysis. To address this, we sought to develop a reporter system for IM permeabilization. Thioflavin-T (ThT) is a stain classically used to label amyloid fibrils. Upon fibril binding, the fluorescence emission of ThT has been shown to increase by several orders of magnitude (127). ThT has been shown to bind *E. coli* RNA and DNA (128). ThT is soluble in water and small enough to enter the periplasm; therefore, we would expect ThT to serve as a reporter of IM permeabilization. To test this, we added ThT to log-phase cultures at the time of thermal induction. Cells were imaged prior to lysis using phase contrast and CFP filter settings. We monitored lysis of 42 cells and detected a continuous increase in ThT fluorescence before lysis (Fig. D.4, Movie 7). The average

time from increase in signal above background to lytic blowout was 13 seconds for λ900 (n =27, standard deviation = 5.3). The average time from ThT signal to the start of shape conversion for λ900Rz_{am}/$Rz1_{am}$ lysogens was 10.3 seconds (Movie 8) (n = 15, standard deviation = 5.2).

To test whether ThT signal was dependent on IM permeabilization, we used lysogens encoding a deletion of the holin and endolysin (λ900Δ*SR*). In this background, the IM will remain intact, since there is no holin activity. As expected, these cells showed no change in fluorescence during the same ThT treatment (not shown). This indicates that during lysis ThT is labeling the cytoplasm as the IM becomes permeabilized by the holin. Therefore, if lysis could be induced independent of holin activity, we would expect ThT labeling to occur upon, but not before, cell breach. To address this, we used MG1655 cells expressing pSec-R to complement the lysis defect of λ900Δ*SR* lysogens (Fig. D.4A). As expected, we did not detect any fluorescence prior to lysis. Upon lytic blowout, ThT signal was transiently detected, quickly diffusing along with released cytoplasmic content (Movie 9). Notably, these cells exhibited a rounding phenotype before lysis despite being normal for rate of lysis in bulk culture (Fig. D.4C). We attribute this defect to reduced spanin expression caused by the Δ*SR* deletion ~100 bp upstream of *Rz* (Fig. D.4D). Taken together, this indicates that ThT reports IM permeabilization, which occurs ~10 seconds before lysis in λ lysogens.

Holin-GFP fusion suggests that the holin determines the site of lysis

Previously, a C-terminal GFP fusion was used to monitor holin triggering in real time (37). When this construct was evaluated for the ability to support lysis in bulk

culture, lytic function of the S105-GFP fusion improved with increasing linker length. The allele encoding a 30 aa linker separating S105 from GFP, designated *S105-GFP*, supported lysis in an *S* nonsense mutant background (S_{am7}) (37). Images of endolysin-deficient lysogens expressing the *S105-GFP* allele were captured every min from 75-105 min after induction. This showed that S105-GFP rearranged from a uniform distribution to form foci (rafts) coinciding with the time of lysis. Notably, endolysin deficient lysogens were used in this study, halting the lytic process before PG degradation. Additionally, no attempt was made to correlate lysis morphology with the site of raft formation. To address these questions, we monitored lysis of λ900S_{am7} lysogens expressing *S105-GFP* from a medium copy plasmid under late gene transcriptional control. Notably, the rate of lysis and triggering time of λ900S_{am7} lysogens expressing the *S105-GFP* allele was improved when expressed within MG1655 (Fig. D.5A, B). The reason for this is not clear, but likely due to higher expression in the MG1655 host; however, differences in the amount of S105-GFP between hosts was not apparent by western blot (Fig. D.5B ii).

We used time-lapse fluorescence microscopy to monitor 50 single lysis events of λ900S_{am7} lysogens expressing *S105-GFP* (listed in order of monitoring time to lysis in Table D.3). Most cells progressed to lysis in the following way: S105-GFP signal was uniformly distributed at the start of monitoring. Approximately 93 sec before lysis S105-GFP suddenly formed foci, or rafts as shown previously (Fig. D.5C, Movie 10, Table D.3) (37). Prior to lysing, raft signal became delocalized (n=41), which we interpret as the holin rearranging to form a hole (discussed later). Importantly, the

frequency of polar lysis by S105-GFP was reduced to 56%. The fact that the GFP fusion is altering lysis morphology further implicates the role of the holin in controlling the site of lysis. To investigate how, we looked at where rafts form in regard to the lytic site. In most cases (65%), raft signal was detected proximal to the site of lysis, even when lysis did not occur from the poles (Fig. D.5D). Overall, rafts form most often at the poles and rafts predict the site of lysis (Table D.3). Taken together, this suggests the holin controls the site of lysis.

Dynamics and features of the holin-GFP raft

On average, we detected 2.3 rafts per cell (n=46) and rafts were most often associated with the poles (Table D.3). This number is less than previously reported (3.3 average), can be explained by rafts forming outside of the focal plane that cannot be detected without confocal/z-stack-based detection. Rafts were previously reported to form within less than a minute (37). Here, we used a much faster frame rate and note that nucleation occurs within 2-3 seconds. Interestingly, rafts were sometimes unstable, forming foci, then delocalizing before forming foci again (Fig. D.6, Movie 11) (n=6). We referred to raft instability as "flickering" (Table D.3). In one such case, raft formation was completely abortive: A raft was initially detected at an area of curvature within the parapolar region. This raft dissolved and then new rafts formed at the poles (Fig. D.6B). Our interpretation for this reversibility is that the GFP attachment to S105 may interfere with formation of a raft, since raft formation would be expected to involve close association of holin molecules. Supporting this idea, S105-GFP function improved as linker length was increased (discussed above).

Out of the 46 cells that formed rafts, 33 cells had phase-light spots associated with at least one raft (Table D.3). These spots appeared at the same time as raft formation, and were morphologically similar to plasmolysis zones (Fig. D.6A, blue arrow). This suggests that rafts frequently cause invagination of the IM from the PG. Also of note is the delocalization of raft signal prior to lysis. On average this was ~35 sec (n=41, with a standard deviation of 20 sec) before lytic blowout (Table D.3). We suspect signal loss reflects the transition of S105-GFP from raft to hole arrangement that is predicted by our model. That is, S105-GFP molecules are tightly packed when in rafts, whereas in the hole-forming arrangement, S105-GFP signal would be expected to be relatively diffuse. The nature of S105 structure after hole formation was probed by cysteine-scanning accessibility. This report presented evidence that virtually all molecules of S105 participate in hole formation and that the hydrophilic faces of two S105 TMDs face the lumen of the hole (42). The number of S105 molecules per cell was determined by quantitative western blotting to be ~1000 (129). Each TMD occupies ~1 nm of area, therefore 1000 copies of S105 could produce 2000 nm of hole-lining perimeter. This would correspond to a single 600 nm-diameter hole or two 300 nm holes, which is consistent with the number and diameter of holes observed by cryoEM of S105-expressing cells (37, 40, 41). Taken together, these data provide the basis that at the endpoint, micron-scale holes are supported by single S105 molecules lining the hole perimeter. Therefore, the delocalization of S105-GFP raft signal prior to lysis supports the model that the hole-forming structure is being formed.

Discussion

The first attempts to systematically document morphological changes during phage lysis can be traced to 1933 when photomicrographic analysis was used to capture lysis of *E. coli* at frame rates as fast as 8 per second (130). Thus, the molecular basis of the morphological changes that occur during phage lysis have been a long standing question. Recently, explosive cell lysis was reported to be crucial for biofilm development in *Pseudomonas.* Super-resolution microscopy was used to show that lysis of a sub-population of cells within a biofilm produced vesicles and eDNA which could be used by other cells (131). Therefore, the study of what happens to cells during phage lysis may have implications for human health.

Phage λ is the most well-studied lysis system to date; however, it is not known how λ directs lysis from the poles. To address this question, we used fluorescence and phase time-lapse microscopy to track the distribution and rearrangement of lysis proteins in the seconds preceding lysis. Results suggest the holin rafts determine the site of lysis.

Based on the findings above, we incorporate our observations of S105-GFP rearrangement into the current model of holin function (Fig. D.7). When holin accumulation reaches a critical concentration, ~2 rafts form about 100 sec prior to lysis. The raft is expected to be a densely packed array of S105 molecules, which largely excludes lipids. Thus, rafts will quickly abrogate the energy production of the cell (Fig. D.7A and B), which is consistent with pmf loss ~19 seconds before lysis, as evidenced by spinning bug experiments (35). Hole formation is triggered by pmf loss, and IM permeabilization is reported by ThT ~13 seconds before lysis. Rafts form most often at

the poles, causing the subsequent steps of lysis (hole formation, endolysin release) to also occur at the poles. Spanins disrupt the OM at the site of PG degradation, completing the last step of lysis. Therefore, the holin controls the site of lysis and directs lysis from the poles.

CHAPTER V

DISCUSSION AND CONCLUSIONS

The work presented here describes a rigorous genetic dissection of the λ spanins and investigates the role of each lysis protein in the polar lysis of *E. coli.* Spanins are required for OM disruption of Gram-negative hosts and the λ spanins, encoded by *Rz* and *Rz1*, are the most well-studied to date. The model for spanin function is that OM disruption occurs by spanin-mediated fusion of the IM and OM at the last step of lysis. This model was substantiated by recent experiments demonstrating that Rz and Rz1 are fusogenic for spheroplast membranes. Based on these data, we sought to understand the mechanism of spanin fusion, and membrane fusion in general, via genetic analysis. The use of genetic analysis in this context is novel for two reasons. 1) Most membrane fusion systems are not genetically tractable and rely on *in vitro* assays of fusion proteins reconstituted into liposomes. Therefore, the researcher is unable to design an experiment to select for alleles that block or restore function of the fusogen. 2) The embedded nature of two genes coding for the same function is unique to biology, although it is common for spanin genes within phage genomes. In the embedded arrangement, genetic analysis would not be practical. Therefore, the first steps of this analysis required removal of *Rz1* from its embedded ORF within *Rz*.

Defining functionally important domains of the spanin complex

Mutational analysis presented in this dissertation have made the following advances toward an understanding of how the spanins function: 1) The coiled-coil

domains within Rz and the proline-rich stretch within Rz1 are the most sensitive to mutation. This suggests these motifs are functionally important, which is in agreement with analogous subdomains shown to be essential in other fusion systems. 2) Most mutations within these regions were normal for accumulation and complex formation. This suggests we have isolated mutants that block function downstream of the formation of the heterotetrameric complex that is essential for lysis. The high frequency in which this type of mutant was isolated indicates that generating a (putative) fusion-blocking mutation is not rare. 3) We have identified a salt bridge interaction between Rz and Rz1 and have provided evidence of an overdetermined C-C-terminal interaction. 4) There were stretches within *Rz* and *Rz1* that were mutationally silent in our selection. We reasoned that such areas might encode linker domains, which have been shown to be a general feature in other fusion systems. By substituting Ser-Gly repeats for native sequence, we identified linker regions within Rz and Rz1. We propose these linkers function analogously to linkers described by paradigm fusion systems, providing a hinge to mediate conformational change. 5) The TMD of Rz was also relatively silent to mutation, and lytic function of Rz was supported after replacement with an artificial sequence of hydrophobic residues. Therefore, the Rz TMD is not sequence specific. This finding agrees with studies of other fusion systems, which showed that the sequence specificity of TMDs is irrelevant (55). 6) The heatmap of inactivating missense mutants of *Rz* and *Rz1* show that mutationally sensitive stretches of *Rz* are silent in *Rz1,* and vice versa. This indicates that a given stretch of DNA shared by embedded genes can encode functionally important sequences for one, but not both

genes. Furthermore, this suggests that the modular structure of Rz and Rz1 might have supported the evolutionary pressure to condense two separate genes into one.

Suppressor selection: spontaneous mutations that rescue spanin function

In Chapter II, we selected for single missense mutants that inactivate spanin function. Most of these mutations block function at an unknown step, which we assumed to be the putative IM-OM fusion reaction. We hypothesized that suppressors of these mutations would allow us to construct a model for spanin-mediated fusion of the IM and OM. The success of suppression analysis for identification of contact points between two macromolecules (conformational suppression) is rare (104), but lends itself well to genetically tractable systems. This information would be useful for the field of membrane fusion as a whole, since the molecular rearrangement of paradigm fusion proteins is inferred from pre- and post-fusion structures.

We found that virtually every inactivating mutation made in Rz or Rz1 was globally suppressed by changes within a membrane proximal segment (juxtamembrane region) of the Rz coiled-coil. Most of these changes were polar insertions within the *a* and *d* positions (the core). Since the core is hydrophobic, polar substitution would be expected to destabilize the coil by disrupting the helix-helix interface. As expected, Rz double mutants, which encode the inactivating mutation and the second-site suppressor, showed evidence of proteolytic degradation in comparison to the parental. Presumably the proteolytic sensitivity is due to the change within the hydrophobic core that produces an unstructured region. The key question is: How does destabilization of Rz rescue function? We proposed a model for suppression based off the data above and the

following: 1) The juxtamembrane region of Rz has a low score for coiled-coil predictions relative to other regions of Rz (Fig. C.1). Therefore, we infer when Rz is in the pre-hairpin conformation, the coiled-coil within the juxtamembrane region is not structured. During the hairpin conformational change step, we envisioned this domain to "zipper", forming coiled-coils that drive membranes together. The zippering model has been proposed in other fusion systems (58, 132). 2) The juxtamembrane was mutationally silent (see Chapter II), suggesting that if this region acts like a zipper, single changes are not debilitating to zipper function. Therefore, our model is that spanin-mediated fusion is likely to involve many intimate helical contacts formed while progressively bringing membranes into apposition. We infer from our data that most of the inactivating mutations block function at some point during this process. In theory, this block could be alleviated by increasing the flexibility of the complex. We suggest insertions within the coiled-coil of the juxtamembrane region cause a localized loss of structure, which allows the zippering process to restart. Phages encoding suppressor alleles showed defects for the rate of lysis. Therefore, whatever the molecular basis for suppression might be, it comes at the cost of normal saltatory function.

Although rare, suppression analysis for membrane fusion proteins has been conducted before (111, 112, 133-136). The majority of these studies have been done in the Kielian lab, using the alphavirus (class-II) E1 (one example is described in Chapter III). Frequently, the molecular defect of the defective input allele is not known, and inferred to be blocking function at the fusion step. In such cases, authors interpret second site suppressors as promoting the conformational change steps, either via

domain-domain packing in the final structure (111), reducing sterics (136), or lowering the pH threshold of triggering (135, 136). In other cases, the molecular defect of the fusogen was known prior to searching for a suppressor. Mutation of alphavirus E1 D188K blocks trimerization and suppressor analysis identified revertant alleles carrying both D188K and K176T. Subsequent modeling supported the authors' claim that intersubunit interaction between these residues is key for homotrimer formation (137). Based on the studies above, it is clear that efforts should be made to elucidate the structure of the spanin complex. This information, paired with genetic data described in this work will aid future efforts to identify functional contacts formed during spanin function.

A novel outer membrane disruptor

Spanin genes can be annotated by searching for signals easily identifiable by bioinformatic analysis, including TMDs, a lipobox, and a signal sequence. These signals are rarely encoded by phage proteins that are not involved with lysis. Using these search parameters, recent bioinformatic analysis indicates that approximately 15% of ~600 dsDNA phages of Gram-negative hosts do not encode identifiable spanins (72). This implies that these phages are encoding an alternate mechanism of OM disruption. Phage phiKT is a phiKMV-like virulent podophage of *E. coli* that appears to have an incomplete lysis cassette. That is, the endolysin (*gp27*) and holin (*gp29*) have been annotated, but there are no candidate spanin genes. To address this, we searched for genes within phiKT that complemented lysis of a λ*Rz*$_{am}$*/Rz1*$_{am}$ lysogen, starting with ORFs flanking known lysis genes. We found that *gp28*, a hypothetical novel gene

located between the endolysin and holin, supported lytic function (Fig. E.1). Gp28 is 56 aa highly cationic protein and secondary structural predictions by JPRED4 indicates *gp28* encodes three alpha helices (82) (Fig. E.1C). Therefore, how Gp28 disrupts the OM will be the next big question. Since Gp28 lacks a TMD or signal sequence, it would be expected to localize to the cytoplasm. Therefore, the simplest notion is Gp28 does not access the OM until the holin forms holes, releasing Gp28 into the periplasm. Once in the periplasm, Gp28 causes OM disruption, possibly by interacting with an OM protein or the membrane itself. The latter model seems plausible based on the similarities of Gp28 to anti-microbial peptides (AMPs). AMPs are small, (often) alpha-helical and encode a high ratio of basic and hydrophobic residues. AMPs are produced by multicellular organisms for defense against prokaryotes. One mechanism of AMP action is electrostatic interaction with bacterial membranes, followed by insertion of the peptide and membrane disruption (138). Future steps should include investigating whether Gp28 is in the cytoplasm before lysis and if Gp28 interacts with membranes. If Gp28 works like an AMP, then it will be important to investigate how accumulation of Gp28 is apparently non-toxic to the host during the morphogenic period (Fig. E.1D). Lastly, it is important to note that whatever the mechanism, preliminary data suggests that there is no morphological penalty for lysis by the plasmid-expressed phiKT lysis cassette, which includes *gp28* (Fig. E.1B and E). That is, phiKT lyses abruptly (Fig. E.1E) as a rod-shaped cell. Rod-shaped lysis morphology indicates that Gp28 functions quickly after the holin triggering step. Future experiments should be designed to incorporate *gp28* into λ in place of the spanins. This would allow direct comparison of

the consequence of this alternate mode of OM disruption, in terms of rate of lysis, morphology, and fitness.

How λ directs lysis from the poles

The first attempts to systematically document morphological changes during phage lysis of *E. coli* can be traced to 1932. In this work, Bayne-Jones used "motion photomicrographic analysis" to capture single lysis events at rates as fast as 8 frames per second. Thus the molecular basis of morphological changes during phage lysis have been a long standing question.

Prior to the work presented in this dissertation, it was reported that λ directed lysis from the cell poles. However, previous localization studies indicated that lysis proteins were distributed throughout the cell without bias. To address this (in Chapter IV), we tested the role of the holin, endolysin, and spanins for evidence of polar bias. Lysogens expressing *S105-GFP* in place of the wt holin formed rafts ~100 seconds before lysing. Rafts were most frequently associated with the poles and rafts frequently predicted the site of lysis. Therefore, in addition to its role as the "lysis clock" and hole-former, the holin controls the site of lysis. Whether polar lysis is a general feature for phages encoding λ-like lysis systems is unknown. If polar lysis morphology is conserved, then experiments should be designed to investigate why. Perhaps polar lysis confers fitness to phage progeny by ejecting phage particles further. Alternatively, ejection from the poles may minimize the surface area of the dead host to newly released phage; i.e., nascent virions would be wasted if they were bound to the dead host's

receptors. If an *S* mutant was identified that was normal for rate of lysis but showed non-polar lysis morphology, it could be used to investigate this question.

Future studies should be directed toward visualizing the S-hole in the IM. Use of super-resolution microscopy techniques such as PALM or immunoSTORM could allow us to generate a 3D image of the holin hole and allow us to test the predictions of the death raft model of holin function. Preliminary efforts to obtain a 3D image using are encouraging. After triggering λΔ*SR* lysogens expressing p*S105-GFP* were imaged by a confocal system. After deconvolution, a 3D image was reconstructed and an apparent ring structure can be seen (Fig. E.2). Measurements of the diameter of the ring are consistent with IM holes observed when cells expressing *S105* were observed using cryoEM (discussed in chapter 4). Furthermore, two-color fluorescence could allow us to follow two lysis proteins at once during lysis, or allow us to investigate membrane integrity with a fluorescent membrane label.

REFERENCES

1. Hendrix RW. 2002. Bacteriophages: evolution of the majority. TheorPopulBiol 61:471-480.

2. Bergh O, Borsheim KY, Bratbak G, Heldal M. 1989. High abundance of viruses found in aquatic environments. Nature 340:467-468.

3. Suttle CA. 2007. Marine viruses--major players in the global ecosystem. Nature reviews Microbiology 5:801-12.

4. Letarov A. 2012. Bacteriophages as a part of the human microbiome, vol 24. CABI Press: Wallingford, Oxfordshire, UK.

5. Carlton RM. 1999. Phage therapy: past history and future prospects. ArchImmunolTherExp(Warsz) 47:267-274.

6. Twort FW. 1915. An investigation on the nature of ultra-microscopic viruses. The Lancet 186:1241-1243.

7. Sulakvelidze A, Morris JG, Jr. 2001. Bacteriophages as therapeutic agents. AnnMed 33:507-509.

8. Summers WC. 1999. Félix d'Herelle and the origins of molecular biology. Yale Universities Press, New Haven.

9. Suttle CA. 2005. Viruses in the sea. Nature 437:356-61.

10. Wang IN, Dykhuizen DE, Slobodkin LB. 1996. The evolution of phage lysis timing. Evol Ecol 10:545-558.

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

12. Bayer M. 1974. Ultrastructure and organization of the bacterial envelope. Annals of the New York Academy of Sciences 235:6-28.

13. Berry J, Savva C, Holzenburg A, Young R. 2010. The lambda spanin components Rz and Rz1 undergo tertiary and quaternary rearrangements upon complex formation. Protein Sci 19:1967-77.

14. Heppel LA. 1967. Selective release of enzymes from bacteria. Science 156:1451-1455.

15. Kennedy EP, Rumley MK. 1988. Osmotic regulation of biosynthesis of membrane-derived oligosaccharides in Escherichia coli. Journal of bacteriology 170:2457-2461.

16. Holtje JV. 1998. Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli* Microbiol Mol Biol Rev 62:181-203.

17. Vollmer W, Blanot D, de Pedro MA. 2008. Peptidoglycan structure and architecture. FEMS Microbiol Rev 32:149-67.

18. Glauner B, Holtje JV, Schwarz U. 1988. The composition of the murein of *Escherichia coli*. J Biol Chem 263:10088-95.

19. Silhavy TJ, Kahne D, Walker S. 2010. The bacterial cell envelope. Cold Spring Harbor perspectives in biology 2:a000414.

20. Ingram L. 1977. Changes in lipid composition of Escherichia coli resulting from growth with organic solvents and with food additives. Applied and environmental microbiology 33:1233-1236.

21. Leive L. 1974. The barrier function of the gram-negative envelope. Ann N Y Acad Sci 235:109-29.

22. Berry JD, Rajaure M, Pang T, Young R. 2012. The spanin complex is essential for lambda lysis. J Bacteriol 194:5667-5674.

23. Young R. 2013. Phage lysis: do we have the hole story yet? Curr Opin Microbiol doi:S1369-5274(13)00150-1 [pii] 10.1016/j.mib.2013.08.008.

24. Natale P, Brüser T, Driessen AJ. 2008. Sec-and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—distinct translocases and mechanisms. Biochimica et Biophysica Acta (BBA)-Biomembranes 1778:1735-1756.

25. Okuda S, Tokuda H. 2011. Lipoprotein sorting in bacteria. Annu Rev Microbiol 65:239-59.

26. Rajaure M, Berry J, Kongari R, Cahill J, Young R. 2015. Membrane fusion during phage lysis. Proceedings of the National Academy of Sciences of the United States of America 112:5497-502.

27. 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.

28. Pang T, Savva CG, Fleming KG, Struck DK, Young R. 2009. Structure of the lethal phage pinhole. Proc Natl Acad Sci U S A 106:18966-71.

29. Court DL, Oppenheim AB, Adhya SL. 2007. A new look at bacteriophage lambda genetic networks. J Bacteriol 189:298-304.

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

31. 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.

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

33. 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.

34. 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.

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

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

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 U S A 108:798-803.

38. Krebs MP, Isenbarger TA. 2000. Structural determinants of purple membrane assembly. Biochimica et Biophysica Acta (BBA)-Bioenergetics 1460:15-26.

39. Wang I-N, Deaton J, Young R. 2003. Sizing the Holin Lesion with an Endolysin-β-Galactosidase Fusion. J Bacteriol 185:779-787.

40. 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 U S A 107:2219-2223.

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

42. To KH, Young R. 2014. Probing the structure of the S105 hole. J Bacteriol doi:10.1128/jb.01673-14.

43. Chang CY, Nam K, Young R. 1995. S gene expression and the timing of lysis by bacteriophage l JBacteriology 177:3283-3294.

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

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

46. Graschopf A, Bläsi U. 1999. Molecular function of the dual-start motif in the λ *S* holin. MolMicrobiol 33:569-582.

47. Young R, Wang IN, Roof WD. 2000. Phages will out: strategies of host cell lysis. Trends Microbiol 8:120-128.

48. 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 70:341-51.

49. Young R, Way S, Yin J, Syvanen M. 1979. Transposition mutagenesis of bacteriophage lambda: a new gene affecting cell lysis. JMolelcular Biology 132:307-322.

50. Taylor A, Kedzierska S, WawrzynÓW A. 1996. Bacteriophage λ Lysis Gene Product Modified and Inserted into Escherichia coli Outer Membrane: Rz1 Lipoprotein. Microbial Drug Resistance 2:147-153.

51. 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.

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

53. 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 Micro 70:341-51.

54. Podbilewicz B. 2014. Virus and Cell Fusion Mechanisms. Annual Review of Cell and Developmental Biology 30.

55. Chernomordik LV, Kozlov MM. 2008. Mechanics of membrane fusion. Nat Struct Mol Biol 15:675-683.

56. Xu H, Zick M, Wickner WT, Jun Y. 2011. A lipid-anchored SNARE supports membrane fusion. Proceedings of the National Academy of Sciences 108:17325-17330.

57. Jun Y, Xu H, Thorngren N, Wickner W. 2007. Sec18p and Vam7p remodel trans‐SNARE complexes to permit a lipid‐anchored R‐SNARE to support yeast vacuole fusion. The EMBO journal 26:4935-4945.

58. Martens S, McMahon HT. 2008. Mechanisms of membrane fusion: disparate players and common principles. Nature Reviews Molecular Cell Biology 9:543-556.

59. Langosch D, Hofmann M, Ungermann C. 2007. The role of transmembrane domains in membrane fusion. Cellular and molecular life sciences 64:850-864.

60. Kielian M. 2014. Mechanisms of Virus Membrane Fusion Proteins. Annual Review of Virology 1.

61. Papahadjopoulos D, Nir S, Düzgünes N. 1990. Molecular mechanisms of calcium-induced membrane fusion. Journal of bioenergetics and biomembranes 22:157-179.

62. Stiasny K, Kössl C, Lepault J, Rey FA, Heinz FX. 2007. Characterization of a structural intermediate of flavivirus membrane fusion. PLoS pathogens 3:e20.

63. Kielian M, Klimjack MR, Ghosh S, Duffus WA. 1996. Mechanisms of mutations inhibiting fusion and infection by Semliki Forest virus. The Journal of cell biology 134:863-872.

64. Marsden HR, Tomatsu I, Kros A. 2011. Model systems for membrane fusion. Chemical Society Reviews 40:1572-1585.

65. BADER M, Holz RW, Kumakura K, Vitale N. 2002. Exocytosis: the chromaffin cell as a model system. Annals of the New York Academy of Sciences 971:178-183.

66. Sørensen JB, Wiederhold K, Müller EM, Milosevic I, Nagy G, de Groot BL, Grubmüller H, Fasshauer D. 2006. Sequential N‐to C‐terminal SNARE complex assembly drives priming and fusion of secretory vesicles. The EMBO journal 25:955-966.

67. Kedzierska S, Wawrzynów A, Taylor A. 1996. The Rz1 gene product of bacteriophage lambda is a lipoprotein localized in the outer membrane of Escherichia coli. Gene 168:1-8.

68. Kedzierska S, Wawrzynow A, Taylor A. 1996. The *Rz1* gene product of bacteriophage lambda is a lipoprotein localized in the outer membrane of *Escherichia coli* Gene 168:1-8.

69. Weber T, Zemelman BV, McNew JA, Westermann B, Gmachl M, Parlati F, Söllner TH, Rothman JE. 1998. SNAREpins: minimal machinery for membrane fusion. Cell 92:759-772.

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

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

72. Kongari R, Young R. 2016. Unpublished data.

73. Chapman B, Chang J. 2000. Biopython: Python tools for computational biology. ACM Sigbio Newsletter 20:15-19.

74. Hanson-Manful P, Patrick WM. 2013. Construction and analysis of randomized protein-encoding libraries using error-prone PCR. Protein Nanotechnology: Protocols, Instrumentation, and Applications, Second Edition:251-267.

75. Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. Science 252:1162-1164.

76. Branden CI. 1999. Introduction to protein structure. Garland Science.

77. Berger B, Wilson DB, Wolf E, Tonchev T, Milla M, Kim PS. 1995. Predicting coiled coils by use of pairwise residue correlations. Proceedings of the National Academy of Sciences 92:8259-8263.

78. Akey DL, Malashkevich VN, Kim PS. 2001. Buried polar residues in coiled-coil interfaces. Biochemistry 40:6352-6360.

79. Parry DA. 1982. Coiled-coils in α-helix-containing proteins: analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. Bioscience reports 2:1017-1024.

80. Straussman R, Ben-Ya'acov A, Woolfson DN, Ravid S. 2007. Kinking the coiled coil–negatively charged residues at the coiled-coil interface. Journal of molecular biology 366:1232-1242.

81. Mijalis E, Holt A. 2016. TAMU-CPT/one_snp_away: Initial Release, v1.0. Texas A&M Center for Phage Technology, https://doi.org/10.5281/zenodo.160195.

82. Drozdetskiy A, Cole C, Procter J, Barton GJ. 2015. JPred4: a protein secondary structure prediction server. Nucleic acids research:gkv332.

83. Brandl CJ, Deber CM. 1986. Hypothesis about the function of membrane-buried proline residues in transport proteins. Proceedings of the National Academy of Sciences 83:917-921.

84. Ulmschneider MB, Sansom MS. 2001. Amino acid distributions in integral membrane protein structures. Biochimica et Biophysica Acta (BBA)-Biomembranes 1512:1-14.

85. Chung K-M, Huang C-H, Cheng J-H, Tsai C-H, Suen C-S, Hwang M-J, Chen X. 2011. Proline in transmembrane domain of type II protein DPP-IV governs its

translocation behavior through endoplasmic reticulum. Biochemistry 50:7909-7918.

86. Reddy Chichili VP, Kumar V, Sivaraman J. 2013. Linkers in the structural biology of protein–protein interactions. Protein Science 22:153-167.

87. Harrison SC. 2008. Viral membrane fusion. Nat Struct Mol Biol 15:690-8.

88. Von Heijne G. 1985. Signal sequences: the limits of variation. Journal of molecular biology 184:99-105.

89. Narita S, Tokuda H. 2010. Sorting of bacterial lipoproteins to the outer membrane by the Lol system. Methods in molecular biology 619:117-29.

90. Berry J. 2011. The final step in phage lysis: The role of of the Rz-Rz1 spanin complex in the disruption of the outer membrane. PhD. Texas A&M University, College Station.

91. Top D, Read JA, Dawe SJ, Syvitski RT, Duncan R. 2012. Cell-cell membrane fusion induced by p15 fusion-associated small transmembrane (FAST) protein requires a novel fusion peptide motif containing a myristoylated polyproline type II helix. Journal of Biological Chemistry 287:3403-3414.

92. Zhu BY, Zhou ME, Kay CM, Hodges RS. 1993. Packing and hydrophobicity effects on protein folding and stability: Effects of β-branched amino acids, valine and isoleucine, on the formation and stability of two-stranded α-helical coiled coils/leucine zippers. Protein Science 2:383-394.

93. Jackson MB, Chapman ER. 2006. Fusion pores and fusion machines in Ca2+-triggered exocytosis. Annu Rev Biophys Biomol Struct 35:135-160.

94. Chernomordik LV, Zimmerberg J, Kozlov MM. 2006. Membranes of the world unite! The Journal of cell biology 175:201-207.

95. Young R. 2013. Phage lysis: do we have the hole story yet? Curr Opin Microbiol 16:790-7.

96. Liu X, Jiang H, Gu Z, Roberts JW. 2013. High-resolution view of bacteriophage lambda gene expression by ribosome profiling. Proc Natl Acad Sci U S A 110:11928-33.

97. Young R. 2014. Phage lysis: three steps, three choices, one outcome. J Microbiol 52:243-58.

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

99. Cahill J, Rajaure M, O'Leary C, Sloan J, Marrufo A, Holt A, Kulkarni A, Hernandez O, Young R. 2016. Genetic Analysis of the Lambda Spanins Rz and Rz1: Identification of Functional Domains. G3: Genes| Genomes| Genetics:g3. 116.037192.

100. 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.

101. 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 Micro 13:495-504.

102. Poon A, Chao L. 2005. The rate of compensatory mutation in the DNA bacteriophage φX174. Genetics 170:989-999.

103. 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.

104. Manson MD. 2000. Allele-specific suppression as a tool to study protein–protein interactions in bacteria. Methods 20:18-34.

105. Zhou NE, Kay CM, Hodges RS. 1993. Disulfide bond contribution to protein stability: Positional effects of substitution in the hydrophobic core of the two-stranded. alpha.-helical coiled-coil. Biochemistry 32:3178-3187.

106. Kwok SC, Hodges RS. 2004. Stabilizing and destabilizing clusters in the hydrophobic core of long two-stranded α-helical coiled-coils. Journal of Biological Chemistry 279:21576-21588.

107. Hillar A, Tripet B, Zoetewey D, Wood JM, Hodges RS, Boggs JM. 2003. Detection of α-helical coiled-coil dimer formation by spin-labeled synthetic peptides: A model parallel coiled-coil peptide and the antiparallel coiled coil formed by a replica of the ProP C-terminus. Biochemistry 42:15170-15178.

108. Kim C, Schmidt T, Cho E-G, Ye F, Ulmer TS, Ginsberg MH. 2012. Basic amino-acid side chains regulate transmembrane integrin signalling. Nature 481:209-213.

109. Bright JN, Sansom MSP. 2003. The flexing/twirling helix: Exploring the flexibility about molecular hinges formed by proline and glycine motifs in transmembrane helices. Journal of Physical Chemistry B 107:627-636.

110. Prelich G. 1999. Suppression mechanisms: themes from variations. Trends in Genetics 15:261-266.

111. Chanel-Vos C, Kielian M. 2006. Second-site revertants of a Semliki Forest virus fusion-block mutation reveal the dynamics of a class II membrane fusion protein. J Virol 80:6115-6122.

112. Kielian M, Chanel-Vos C, Liao M. 2010. Alphavirus entry and membrane fusion. Viruses 2:796-825.

113. Russell CJ, Kantor KL, Jardetzky TS, Lamb RA. 2003. A dual-functional paramyxovirus F protein regulatory switch segment activation and membrane fusion. The Journal of Cell Biology 163:363-374.

114. Paterson RG, Russell CJ, Lamb RA. 2000. Fusion protein of the paramyxovirus SV5: destabilizing and stabilizing mutants of fusion activation. Virology 270:17-30.

115. West DS, Sheehan MS, Segeleon PK, Dutch RE. 2005. Role of the simian virus 5 fusion protein N-terminal coiled-coil domain in folding and promotion of membrane fusion. J Virol 79:1543-1551.

116. McNew JA, Weber T, Engelman DM, Söllner TH, Rothman JE. 1999. The length of the flexible SNAREpin juxtamembrane region is a critical determinant of SNARE-dependent fusion. Molecular cell 4:415-421.

117. Pobbati AV, Stein A, Fasshauer D. 2006. N-to C-terminal SNARE complex assembly promotes rapid membrane fusion. Science 313:673-676.

118. Gao Y, Zorman S, Gundersen G, Xi Z, Ma L, Sirinakis G, Rothman JE, Zhang Y. 2012. Single reconstituted neuronal SNARE complexes zipper in three distinct stages. Science 337:1340-1343.

119. Söllner TH. 2004. Intracellular and viral membrane fusion: a uniting mechanism. Current Opinion in Cell Biology 16:429-435.

120. Berry J, Rajaure M, Pang T, Young R. 2012. The spanin complex is essential for lambda lysis. J Bacteriol 194:5667-5674.

121. Isenbarger TA, Krebs MP. 2001. Thermodynamic stability of the bacteriorhodopsin lattice as measured by lipid dilution. Biochemistry 40:11923-11931.

122. Isenbarger TA, Krebs MP. 1999. Role of helix-helix interactions in assembly of the bacteriorhodopsin lattice. Biochemistry 38:9023-9030.

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

124. Cahill J, Rajaure M, Holt A, Moreland R, O'Leary C, Kulkarni A, Sloan J, Young R. 2017. Suppressor Analysis of the Fusogenic Lambda Spanins. J Virol:JVI. 00413-17.

125. Zhang N, Young R. 1999. Complementation and characterization of the nested *Rz* and *Rz1* reading frames in the genome of bacteriophage lambda. Molecular and General Genetics 262:659-667.

126. de Pedro MA, Quintela JC, H"ltje JV, Schwarz H. 1997. Murein segregation in Escherichia coli JBacteriology 179:2823-2834.

127. Biancalana M, Koide S. 2010. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1804:1405-1412.

128. Sugimoto S, Arita-Morioka K-i, Mizunoe Y, Yamanaka K, Ogura T. 2015. Thioflavin T as a fluorescence probe for monitoring RNA metabolism at molecular and cellular levels. Nucleic acids research 43:e92-e92.

129. Chang CY, Nam K, Young R. 1995. S gene expression and the timing of lysis by bacteriophage lambda. J Bacteriol 177:3283-94.

130. Bayne-Jones S, Sandholzer LA. 1933. Changes in the shape and size of Bacterium coli and Bacillus megatherium under the influence of bacteriophage—A motion photomicrographic analysis of the mechanism of lysis. Journal of Experimental Medicine 57:279-303.

131. Turnbull L, Toyofuku M, Hynen AL, Kurosawa M, Pessi G, Petty NK, Osvath SR, Cárcamo-Oyarce G, Gloag ES, Shimoni R. 2016. Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. Nature communications 7.

132. Südhof TC, Rothman JE. 2009. Membrane fusion: grappling with SNARE and SM proteins. Science 323:474-477.

133. Zheng Y, Sánchez-San Martín C, Qin Z-l, Kielian M. 2011. The domain I-domain III linker plays an important role in the fusogenic conformational change of the alphavirus membrane fusion protein. J Virol 85:6334-6342.

134. Melder DC, Yin X, Delos SE, Federspiel MJ. 2009. A charged second-site mutation in the fusion peptide rescues replication of a mutant avian sarcoma and

leukosis virus lacking critical cysteine residues flanking the internal fusion domain. J Virol 83:8575-8586.

135. Vlaycheva L, Nickells M, Droll DA, Chambers TJ. 2004. Yellow fever 17D virus: pseudo-revertant suppression of defective virus penetration and spread by mutations in domains II and III of the E protein. Virology 327:41-49.

136. Stanifer ML, Cureton DK, Whelan SP. 2011. A recombinant vesicular stomatitis virus bearing a lethal mutation in the glycoprotein gene uncovers a second site suppressor that restores fusion. J Virol 85:8105-8115.

137. Liu CY, Besanceney C, Song Y, Kielian M. 2010. Pseudorevertants of a Semliki Forest virus fusion-blocking mutation reveal a critical interchain interaction in the core trimer. J Virol 84:11624-11633.

138. Zhang L-j, Gallo RL. 2016. Antimicrobial peptides. Current Biology 26:R14-R19.

139. 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.

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

141. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K‐12 in‐frame, single‐gene knockout mutants: the Keio collection. Molecular Systems Biology 2.

142. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. JBacteriology 177:4121-4130.

APPENDIX A

FIGURES AND TABLES

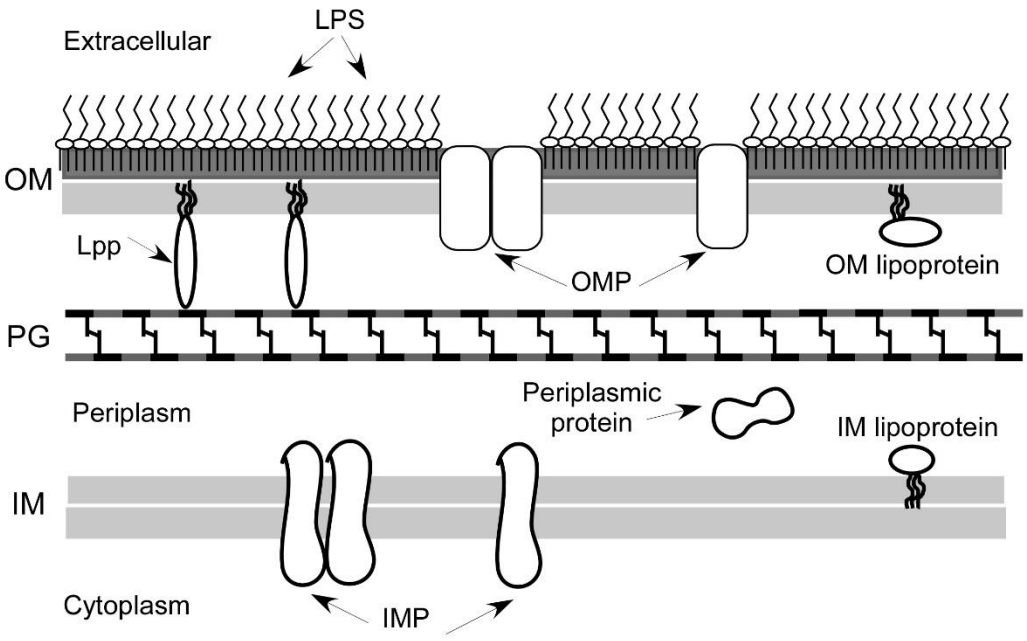


Figure A.1. The cell envelope of *E. coli.*

The outer membrane (OM), peptidoglycan (PG), and inner membrane (IM) are labeled to the left. Inner and outer membrane proteins (IMP and OMP), and lipoproteins, periplasmic proteins and lipopolysaccharide (LPS) are denoted with arrows.

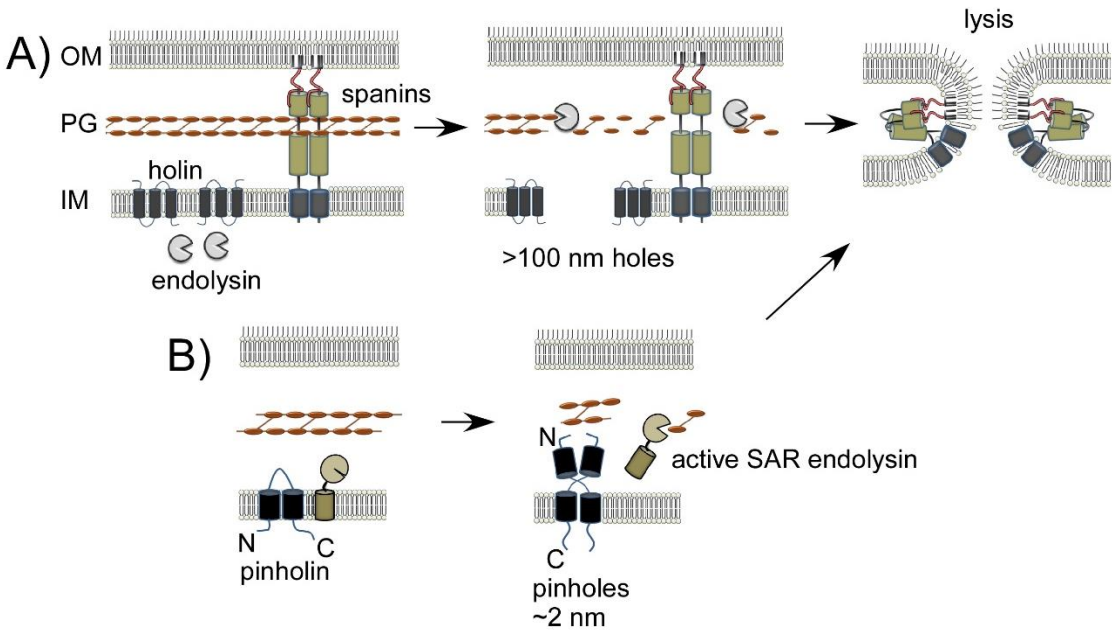


Figure A.2. Tailed phage lysis models.

(A) λ-like lysis model. Endolysin accumulates within the cytoplasm and is released to the periplasm when the holin forms holes in the IM. Spanins function after PG degradation to disrupt the OM by fusing the IM and OM. (B) SAR-endolyin model. The SAR endolysin accumulates in an inactive form, tethered to the IM. After the pinholin triggers, pmf loss releases the SAR endolysin from the IM in an active form. Lysis is complete after OM disruption by the spanins.

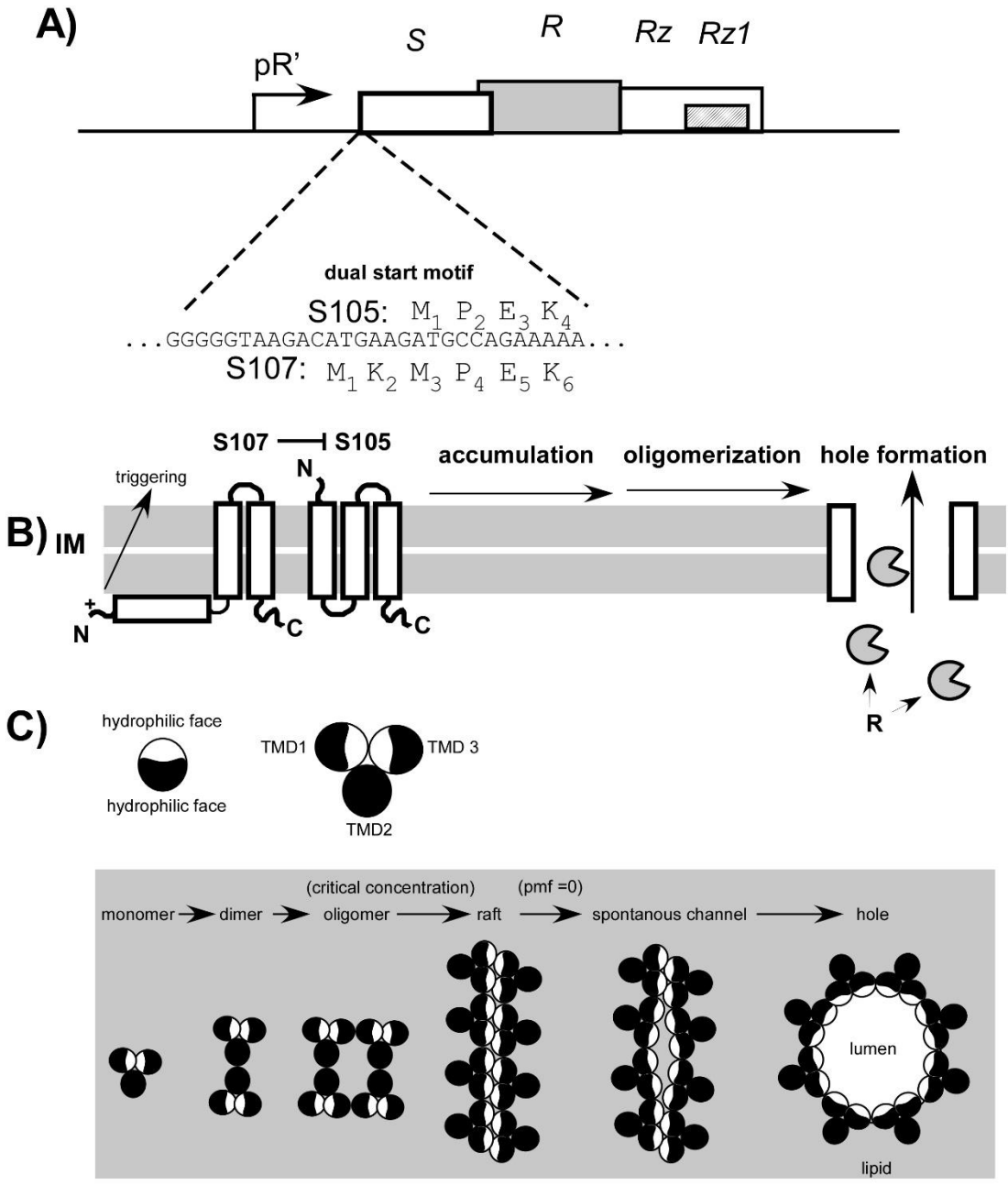


Figure A.3. S105 and S107

(A) The lysis cassette of λ and the dual start motif of the *S* gene is shown. (B) The topology of S107 and S105. In this configuration, S107 negatively regulates S105. At

the time of triggering the S107 TMD1 flips its N-terminus into the periplasm. (C) Top-down view of the holin TMDs. The hydrophobic and hydrophilic faces are shown. Below the lipid bilayer is shown in grey. The death raft model for oligomerization and rearrangement of the holin molecule is shown. The holin accumulates as dimers. At a critical concentration, the holin forms rafts, which get large enough to disrupt energy production of the cell. When pmf is lost, the holin rearranges to form a hole. The hydrophilic faces of TMD1 and 3 face the lumen.

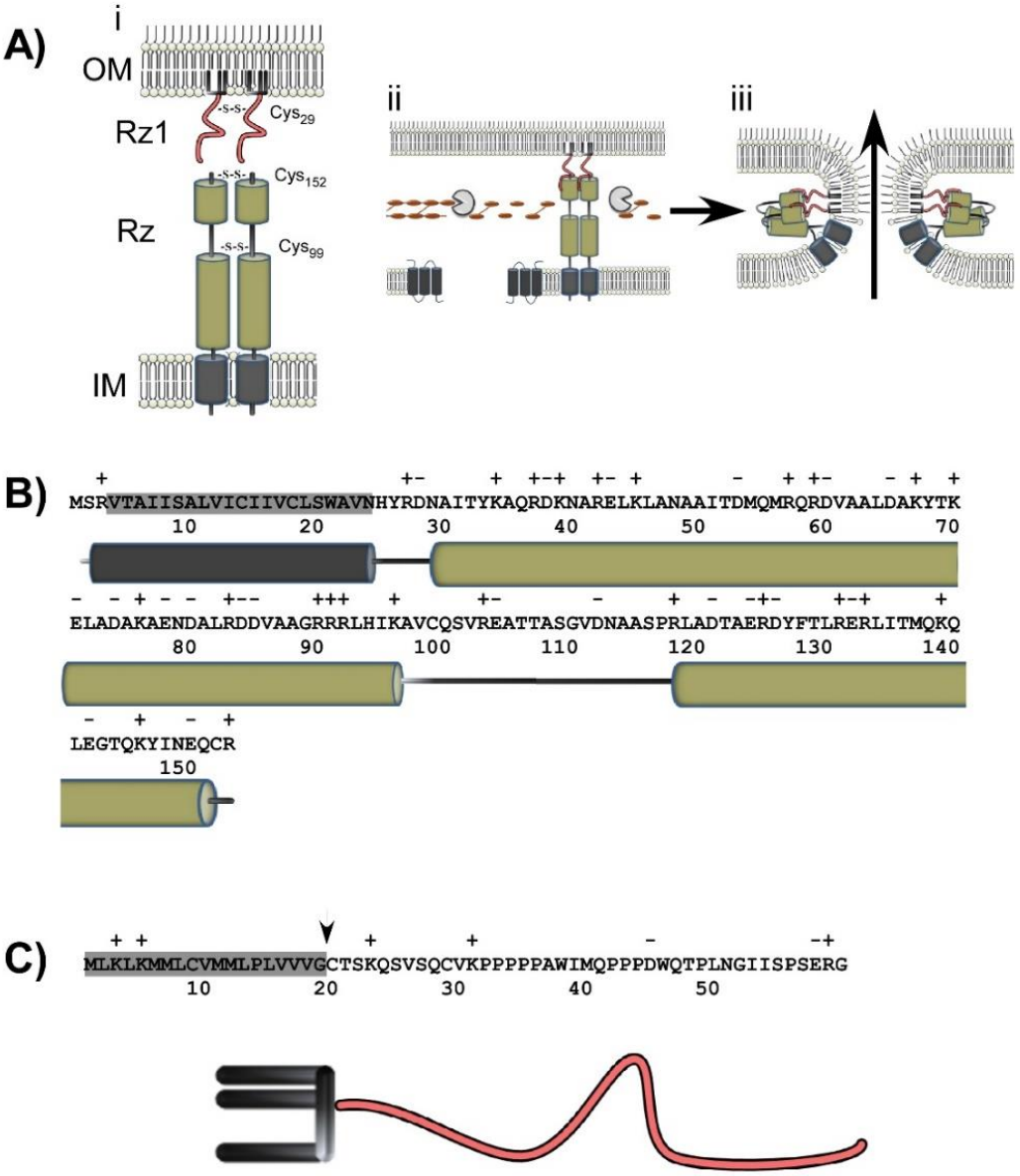


Figure A.4. λ spanins Rz and Rz1.

(A)(i) Cartoon of Rz and Rz1 shown within the cell envelope. The homotypic intermolecular disulfide linkages are denoted. (ii) Model for function: The spanins

activate after PG degradation. (iii) Conformational change that causes fusion of the IM and OM. (B) The amino acid sequence of Rz is shown with predicted secondary structural features below. (C) The amino acid sequence of Rz1. The lipoylated Cys20 is denoted with an arrow. A cartoon of the mature Rz1 lipoprotein is below.

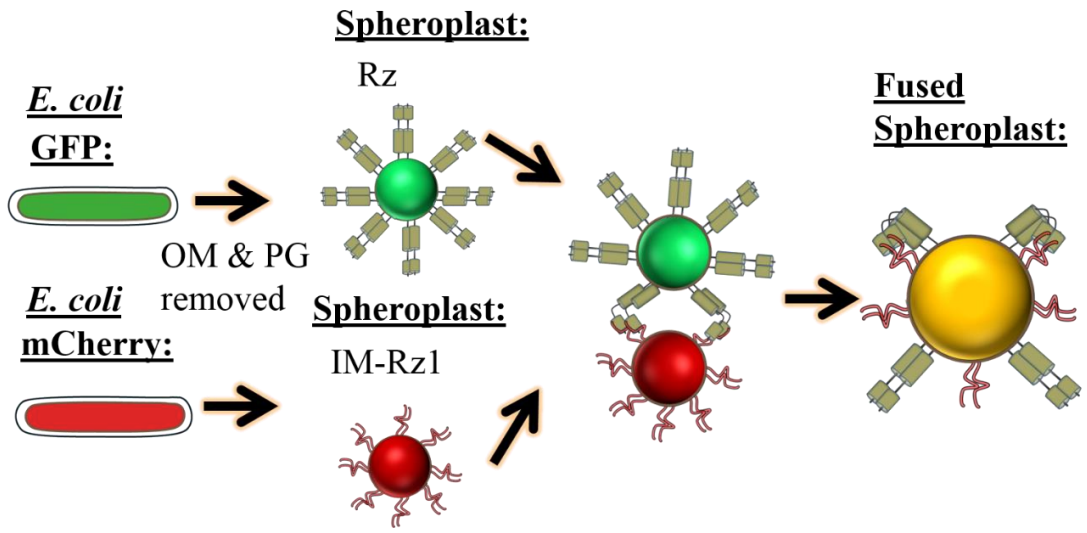


Figure A.5. Spheroplast fusion assay.

E. coli cells are induced for coexpression of GFP and Rz (green cell) or mCherry and Rz1 (red cell). Spheroplasts are formed by treating the cells with EDTA and lysozyme. Cells are now bound by the IM. To retain Rz1 at the IM, a mutant of Rz1 (IM-Rz1) is used. Fused spheroplasts will retain both GFP and mCherry (yellow).

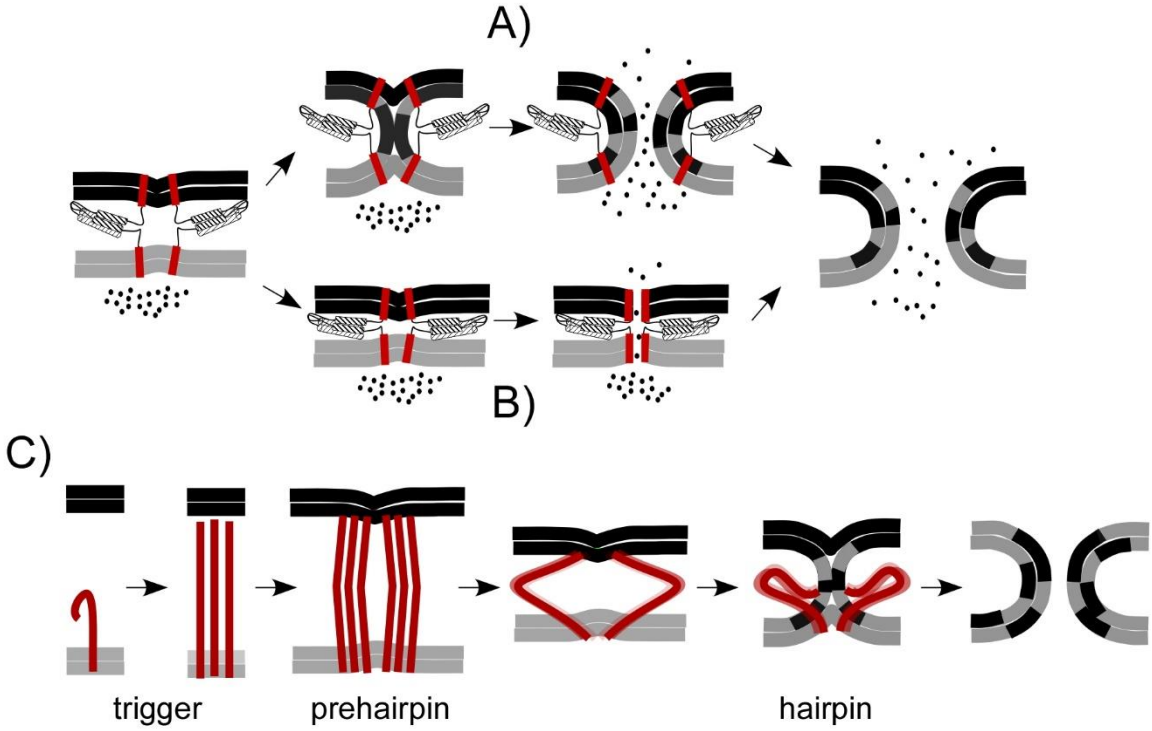


Figure A.6. Models of membrane fusion

(A) The hemifusion model; a lipid pore is formed. (B) The protein pore model; the TMDs (red) are localized in the pore. (C) Viral fusion model. The target membrane is black and the viral envelope is grey. The viral fusion protein is red.

APPENDIX B

FIGURES AND TABLES

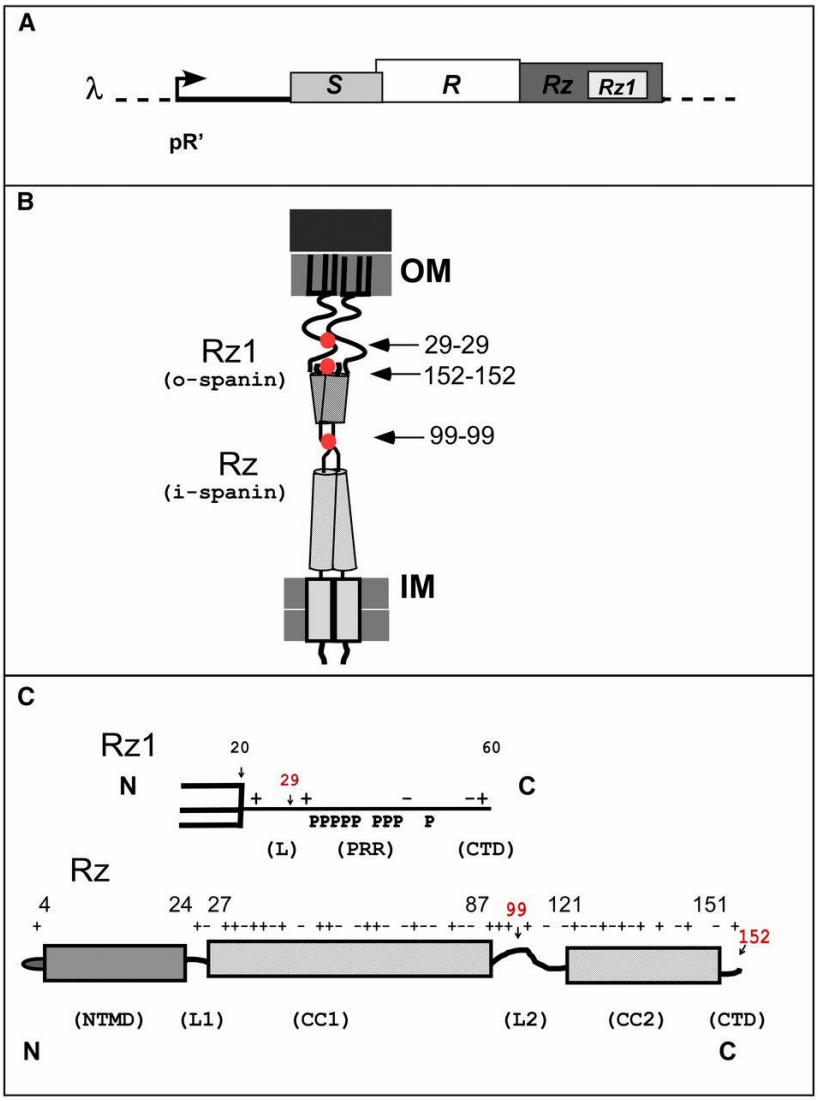


Figure B.1. Features of the lambda spanins.

(A) The lambda lysis cassette. The lysis genes of lambda are S, R, and Rz/Rz1, and are under control of the late promoter pR'. (B) Cartoon of the spanin complex in prehairpin conformation within the cell envelope. The lambda o-spanin, Rz1, is a dimer that is anchored to the inner leaflet of the OM by a lipid anchor. The lambda i-spanin, Rz, is a dimer anchored to the IM by a TMD. Rz extends two helices into the periplasm. Red dots and arrows denote the location of the homotypic intermolecular disulfide bonds. (C)

Cartoon schematic of Rz1 and Rz, highlighting charge distribution, predicted secondary structure features, and domains. The positions of Cys residues are indicated in red.

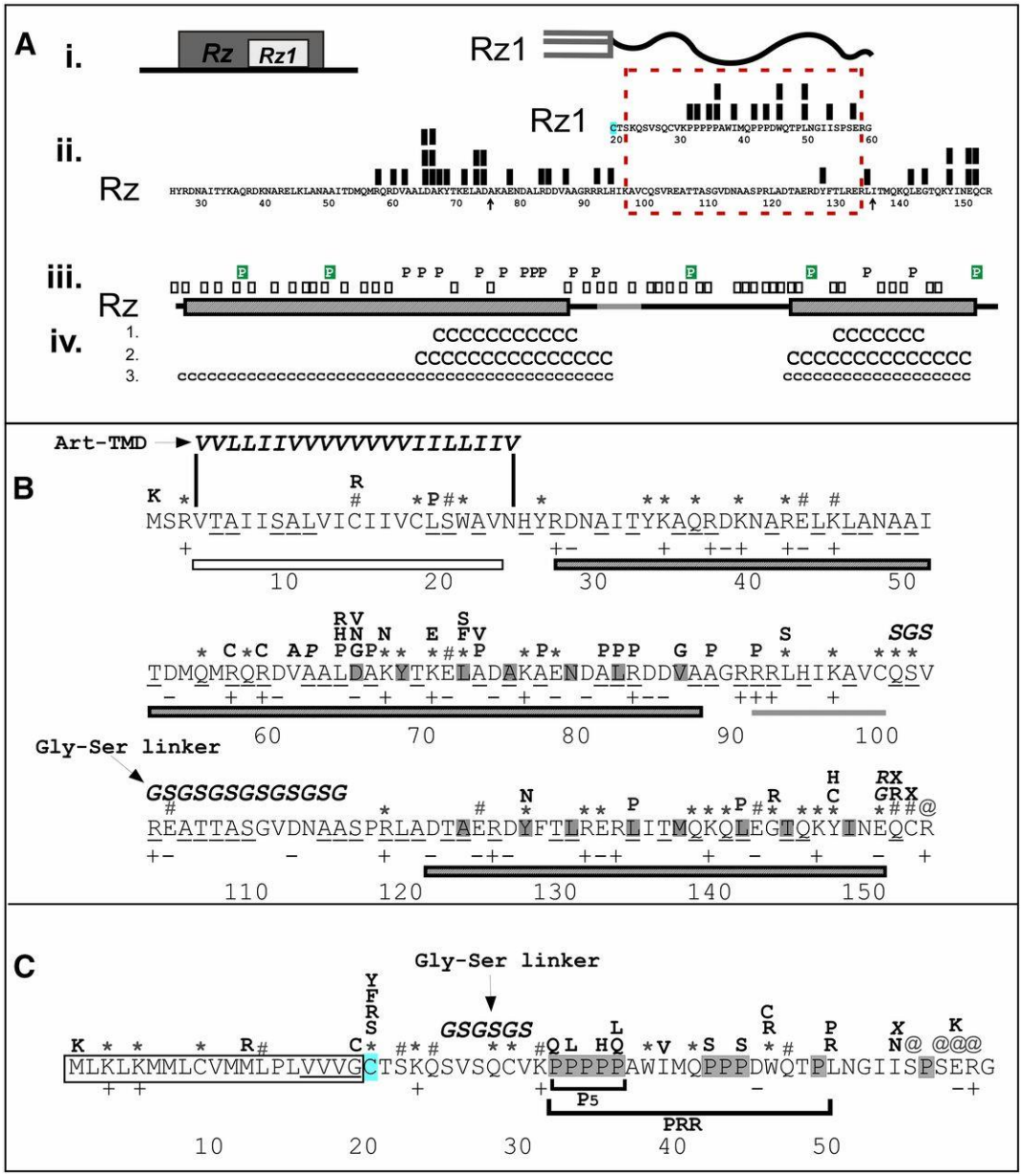


Figure B.2. The location of single missense mutants aligned within Rz and Rz1.

(A) (i) Rz1 is embedded within Rz. (ii) The coding sequences and cartoon products of *Rz* and *Rz1* are aligned according to embedded architecture. Arrows denote where *Rz1* is embedded within the *Rz* coding sequence. The Rz TMD (1–24) and Rz1 signal sequence (1–19) are removed for clarity. The location of single missense mutants that abrogate function are represented with black boxes above the Rz and Rz1 aa

sequence. The lipoylated cysteine of Rz1 (Cys20) is highlighted in blue. The dashed red box highlights a region in the spanin coding sequence where single mutants signal was high in *Rz1*, and low in *Rz*. (iii) Predicted Rz secondary structural features aligned with the primary structure features above. Black line: regions with no predicted secondary structure, Striped rectangles, Rz α-helices; Gray line, β-sheet. Boxes represent *Rz* codons that can be mutated to encode proline in one mutational step. Black "P" denotes positions where a proline blocked function. Green-highlighted "P" denotes positions where a proline mutation did not block function. (iv) Coiled-coil score predictions aligned respective to the secondary structure features of Rz above. (1). "C" represents coiled-coil motif predictions by Coils (score of >0.74, window 14). (2). "C" represents coiled-coil predictions by Coils (score of >0.5, window 21) (3). "c" represents coiled-coil predictions by Paircoil2 (*p*-scores <0.04). (B) Single mutants of Rz shown with primary structure. Single missense mutants are aligned above the amino acid sequence in their respective positions. "*" indicate nonsense codons identified by the screen that were generated by a single base pair change. "#" indicates positions which nonsense codons could be generated with a single base pair change, but were not obtained in the selection. "@" indicates codons of the latter type (#) that do not block Rz function. Underlined residues in the Rz sequence can be changed to Pro with a single base change. Residue charges are identified below the amino acid sequence. Italicized letters indicate mutants identified by site-directed mutagenesis. Relative positions of the Art-TMD, and the Gly-Ser linker are identified with an arrow. The open rectangle, striped rectangle, and gray line represent predictions of the Rz TMD, α-helices, and β-sheet, respectively. The *a* and *d* positions of the Rz coiled-coil as predicted by COILS (score >0.5; window = 21) that fall within the range of predicted Rz α-helix are indicated by gray highlight. (C) The primary structure of Rz1. Single missense mutants are aligned above the amino acid sequence in their respective positions. "*" indicate nonsense codons identified by the screen that were generated by a single base pair change. "#" indicates positions at which nonsense codons could be generated with a single base pair change, but were not detected. "@" indicates codons of the latter type (#) that do not block Rz1 function. Italicized letters indicate mutants identified by site-directed mutagenesis. Relative positions of the Gly-Ser linker are identified with an arrow. The first 19 residues of Rz1 are boxed to represent the signal sequence. The Rz1 lipobox is highlighted at the −1 to −3 positions relative to the lipoylated Cys (20), which is highlighted in blue. Proline residues are highlighted in gray. P_5 and PRR represent the pentaproline stretch and proline-rich region, respectively.

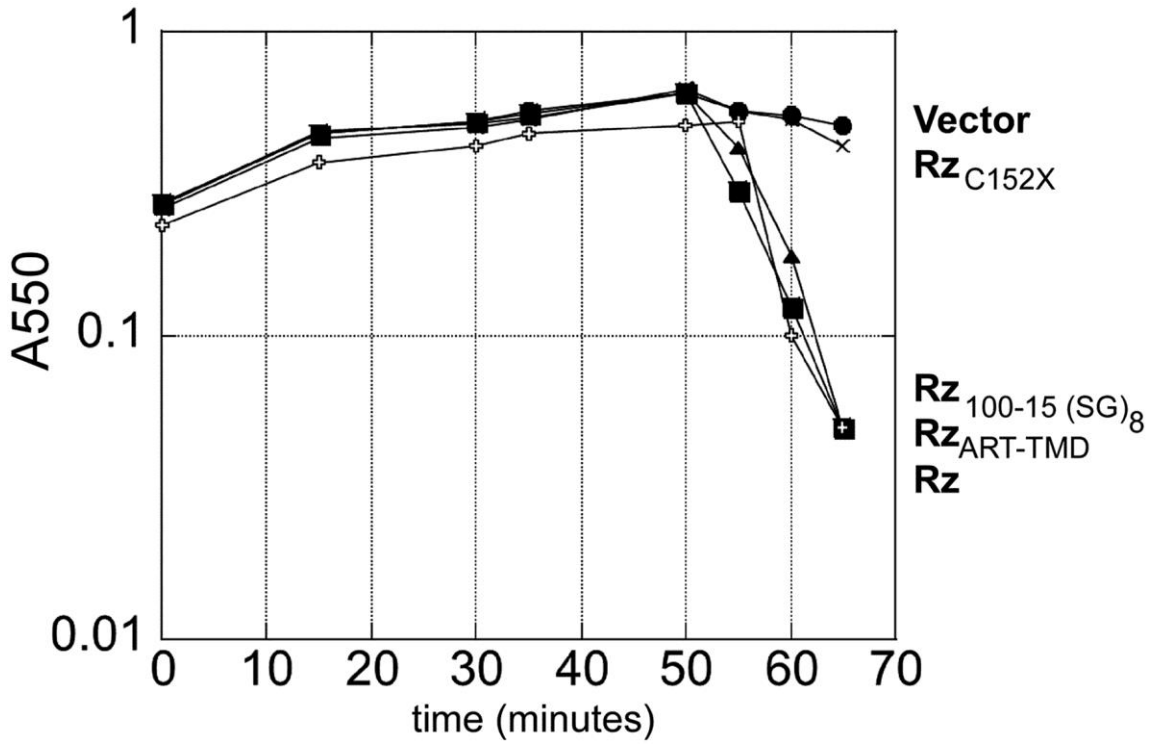


Figure B.3. Lysis profile of Rz artificial TMD, artificial linker, and C152X.

The following lysogens were induced at time = 0 and monitored at A550: MC4100 (λ900 *Rz*$_{am}$) lysogens carrying either an empty pRE plasmid (vector), or pRE with following *Rz* alleles: pRz (WT), or pRz with an artificial transmembrane domain (ART-TMD), pRz 100–115 (Ser-Gly)$_8$, and Rz C152X.

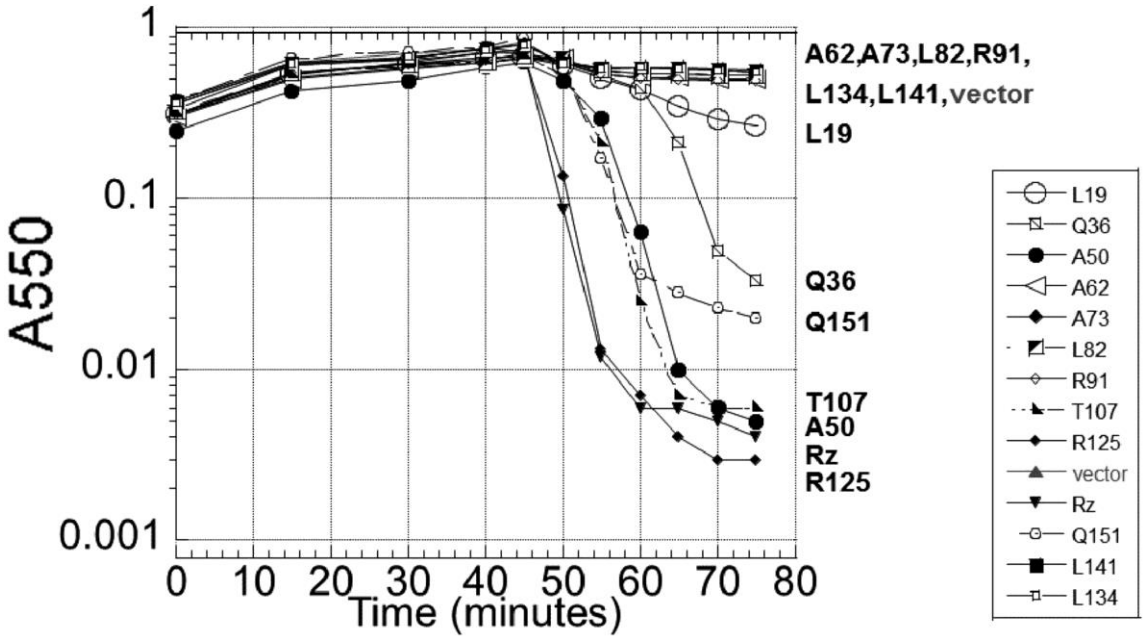


Figure B.4. Lysis profile of Rz proline substitutions.

The following lysogens were induced at time = 0 and monitored at A550: MC4100 (λ900Rz_{am}) lysogens carrying either an empty pRE plasmid (vector), or pRE with following *Rz* alleles: pRz (WT), or pRz with the residue and position of proline substitution. The residue and position of the proline substitution within the plasmid-expressed *Rz* allele is identified in the legend.

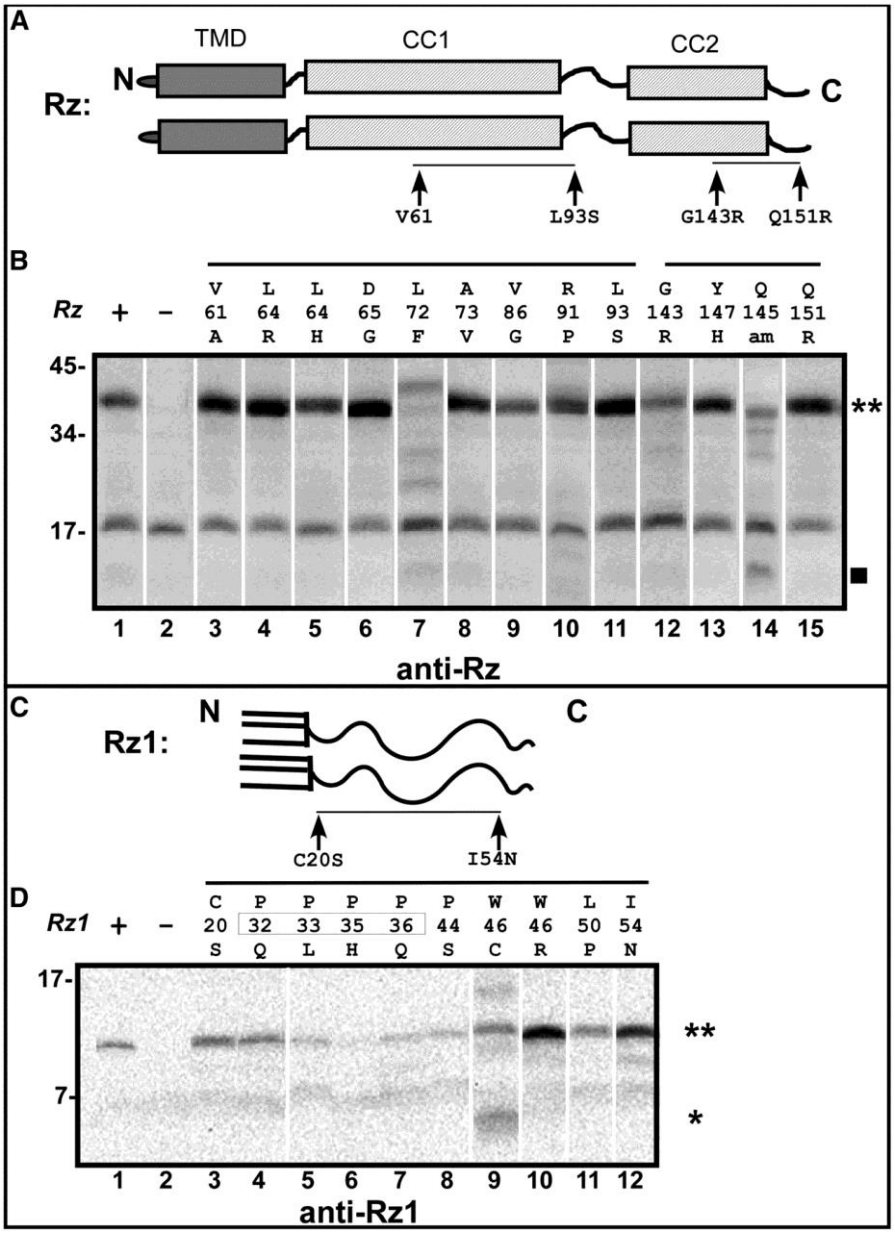


Figure B.5. Accumulation of spanin mutant gene products.

(A) A cartoon of the Rz dimer is shown. Arrows indicate the positions of V61, L93, G143, and Q151 with respect to predicted features of Rz. The black line between the arrows indicates the relative position of mutants in (B) analyzed by the Western Blot. (B) Anti-Rz Western Blot of Rz mutants in the presence of Rz1. Rz mutants are identified above each lane. The Rz dimer band is denoted by "**," and the Rz

breakdown product is denoted by the square symbol. (C) A cartoon of the Rz1 dimer is shown. Arrows indicate the positions of C20 and I54N. The black line between the arrows indicates the relative position of mutants in (D) analyzed by the Western Blot. (D) Anti-Rz1 Western Blot. Rz1 mutants are identified above each lane. The Rz1 dimer band is denoted by "**," and the Rz1 monomer product is denoted by "*."

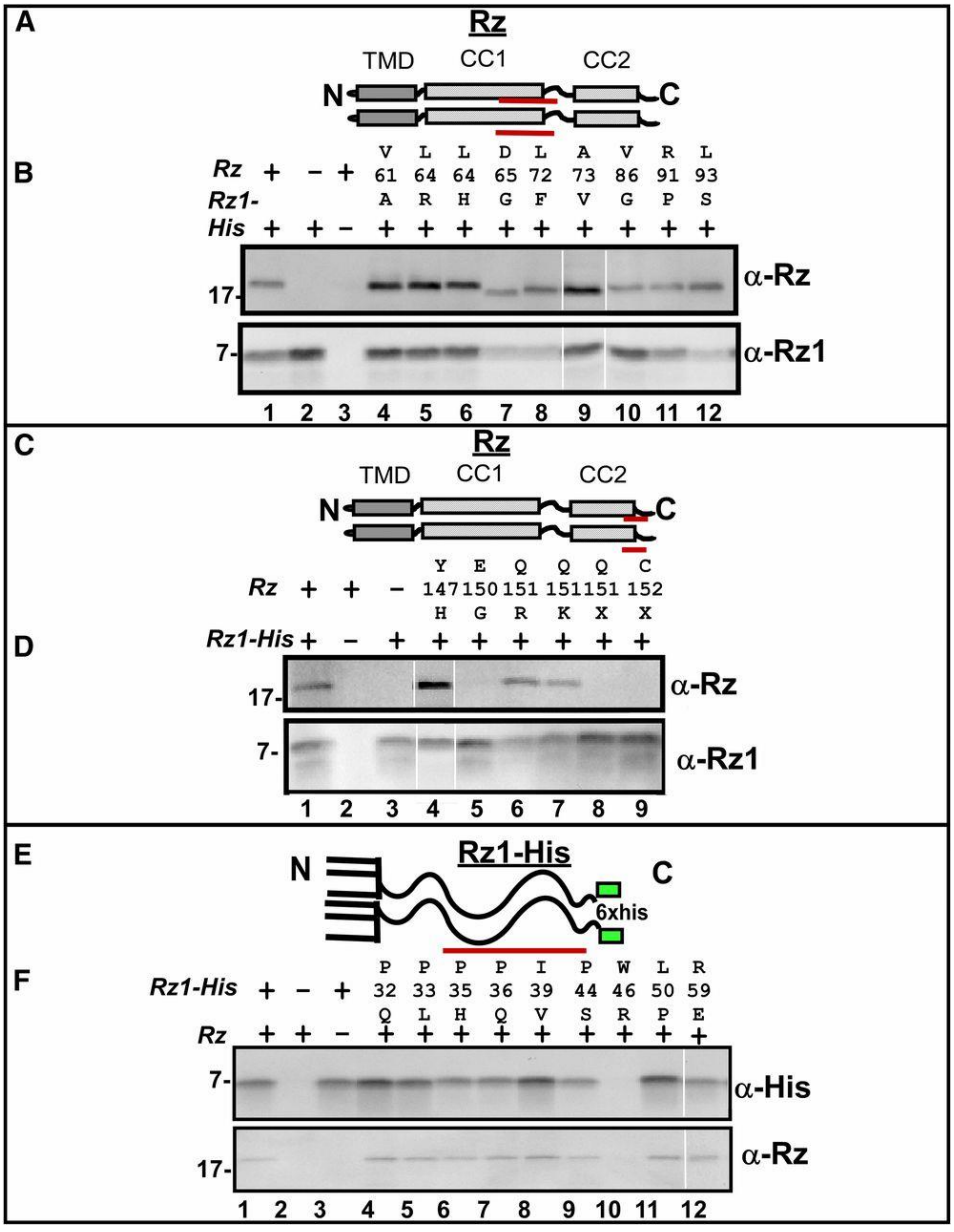
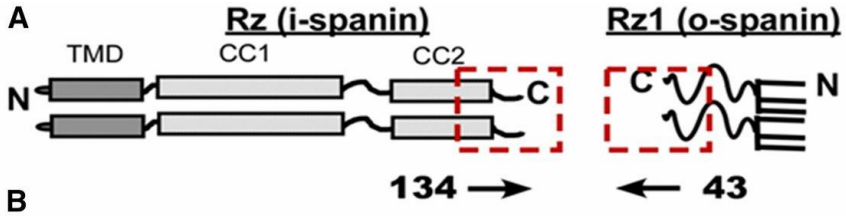


Figure B.6. Oligohistidine pulldown of spanin mutants.

(A) Cartoon of Rz structure showing the relative location of CC1 mutants used in the oligohistidine pulldown. The TMD, and proximal and distal helices are represented with gray and striped boxes, respectively. Red lines below CC1 represent the relative position of mutant residues used in the pulldown assay. (B) Coimmunoprecipitation of Rz CC1 mutants with Rz1-His by oligohistidine pulldown. Pulldown products were analyzed by

Western blot with anti-Rz and anti-Rz1 antibodies. (C) Cartoon of Rz structure showing the relative location of CC2 mutants used in the oligohistidine pulldown. The TMD, and proximal and distal helices are represented with gray, and striped boxes, respectively. Red lines below CC2 represent the relative position of mutant residues used in the pulldown assay. (D) Coimmunoprecipitation of Rz CC2 mutants with Rz1-His by oligohistidine pulldown. Pulldown products were analyzed by Western blot with anti-Rz and anti-Rz1 antibodies. (E) Cartoon of Rz1-His showing relative location of mutants used in the oligohistidine pulldown. Red lines below Rz1-His represent the relative position of mutant residues of Rz1-His used in the pulldown assay. This position of the His tag is represented with a green box. (F) Coimmunoprecipitation of Rz1-His mutants with Rz. Pulldown products were analyzed by Western blot with anti-Rz and anti-Rz1 antibodies.



B

Lambda family	i-spanin; position 134 → C-terminus	o-spanin; position 43 → C-terminus
Lambda	LITMQKQLEGTQKYINEQCR	PPDWQTPLNGIISPSERG
1720a-0	LMTMQKQLEGAQEYIRTQCLN	APDWQTPLNGIISSSENG
933W	LITMQKQLEGTQKYINEQCR	APDWQTPLNGIISSSERG
VT2-Sakai	LMTMQKQLEGAQDYIRTQCLN	APDWQTPLNGIISSSENG
TL-2011c	LITMQKQLEGTQKYINEQCR	APDWQTPLNGIISPSERG
Min27	VITMQKQLEGTQKYINEQCR	APDWQTPLNGIISSSERG
BP-4795	LMTMQKQLEGAQDYIRTQCLN	APDWQTPLNGIISSSENG

Figure B.7. Covariance of Rz E150 with Rz1 R59.

(A) A cartoon model of Rz and Rz1 positioned with C-termini in apposition. The red square highlights the relative position of the amino acid sequences used for covariance analysis. (B) Six lambda family spanin equivalents are aligned respective to position 134–153 of Rz and 43–60 of Rz1. To match the cartoon above, the Rz1 sequence is arranged C–N terminus (positions 60–43). Red and Blue letters identify positions with changing positive and negative charge. Bolded letters identify positions with changes to polar residues. The highlighted positions correspond to Rz E150 and Rz1 R59.

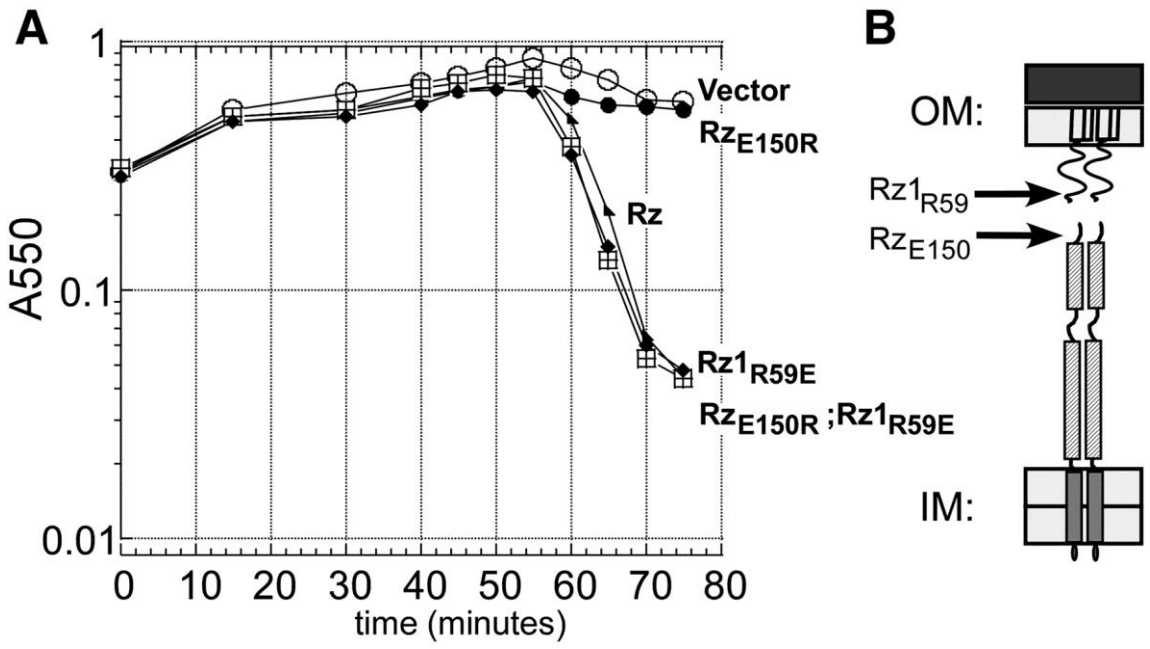


Figure B.8. Lysis profile showing Rz1 R59E suppresses the Rz E150R defect.

(A) The following lysogens were induced at time = 0 and monitored at A550: MC4100 (λ900 *Rz*$_{am}$) carrying the following plasmids pRE (open circle), pRE Rz (triangle), pRE Rz E150R (closed circle). MC4100 (λ900 *Rz*$_{am}$ *Rz1*$_{am}$) carrying pRE Rz E150R Rz1 R59E (square), MC4100 (λ900 *Rz1*$_{am}$) carrying pRE Rz1 R59E (diamond). (B) Cartoon of the spanin complex in the cell envelope. The relative positons of Rz E150 and Rz1 R59 are identified by arrows. IM, inner membrane; OM, outer membrane.

Table B.1. Phages, Strains, Plasmids, and Primers used in this study.

Bacteriophages	**Genotypes and relevant features**	**Sources**
λ900	λ*Δ(stf tfa)::cat cI*$_{857}$ *bor::kan;* carries Cam^R and Kan^R	Lab stock
λ900*Rz*$_{Q100am}$ *Rz1*$^+$	*Rz* gene carries amber codon at position 100 with silent change in embedded *Rz1*	Lab stock
λ900*Rz*$^+$*Rz1*$_{W38am}$	*Rz1* gene carries amber codon at position 38 with silent change in overlapping *Rz*	Lab stock
λ900*Rz*$_{Q100am}$ *Rz1*$_{W38am}$		Lab stock
Strains		
MC4100*tonA::Tn10*	*E. coli K-12 F araD139 Δ(argF-lac)U169 rpsL15 relA1 flbB3501 deo pstF25 rbsR tonA*	Lab stock
MC4100 λ900	MC4100 *tonA::Tn10* lysogenized with λ900	Lab stock
MC4100 λ900*Rz*$_{Q100am}$ *Rz1*$^+$	MC4100 *tonA::Tn10* lysogenized with λ900 *Rz*$_{Q100am}$ *Rz1*$^+$	Lab stock
MC4100 λ900*Rz*$^+$*Rz1*$_{W38am}$	MC4100 *tonA::Tn10* lysogenized with λ900 *Rz*$^+$*Rz1*$_{W38am}$	Lab stock
MC4100 λ900*Rz*$_{Q100am}$ *Rz1*$_{W38am}$	MC4100 *tonA::Tn10* lysogenized with λ900 *Rz*$_{Q100am}$ *Rz1*$_{W38am}$	Lab stock
RY17341	MDS12*ΔtonA;* MG1655 with 12 deletions, totaling 376,180 nt, including cryptic prophages	Lab stock
RY17341 λ	RY17341 lysogenized with temperature sensitive λcI857	This study
RY17299 *lacI*^{q1}	Derived from MG1655 *ΔtonA*	(139)
Plasmids		
pRE	Plasmid with the λ later promoter pR' that is transcriptionally activated by λQ	Lab stock

pRz	pRE carrying Rz alone with Rz1 inactivated by a nonsense mutation	Lab stock
pRz1	pRE carrying Rz1	Lab stock
pRz mutX	pRE carrying denoted mutation of Rz	This study
pRz1 mutX	pRE carrying denoted mutation of Rz1	"
pSynRz	pRE carrying Rz alone. The former region of Rz1 overlap within Rz was codon-optimized to be genetically disparate from WT Rz1 to avoid recombination	"
pSynRz1	pRE carrying Rz1 alone. SynRz1 was codon-optimized to be genetically disparate from WT Rz to avoid recombination	"
pSynRz mutX	pRE carrying denoted mutation within SynRz	"
pSynRz1 mutX	pRE carrying denoted mutation within SynRz1	"
pSynRz Linker 101-115	Residues 101-115 of Rz replaced with Gly-Ser repeats	
pLinkerRz Q100S	Residues 100-115 of Rz replaced with 8 Ser-Gly repeats	"
pRz1 25-30 GS-Linker	Residues 25-30 of Rz1 replaced with 3 Gly-Ser repeats	"
$pRz1_{his}$	Rz1 with His-tag at the C-terminal end	(48)
pQ	pSC101 origin with low-copy mutation; *Q* cloned under $P_{lac/ara-1}$ promoter	(35)

Primer	**Sequence 5' - 3'**
pRz S20P FOR	TCTGCCTGCCATGGGCTGTTAATC
pRz S20P REV	GATTAACAGCCCATGGCAGGCAGA
pRz Q36P	GAGATAACGCCATTACCTACAAAGCCCCGCGCGACAAAAATGCCAGAGAAC
pRz A50P	CTGAAGCTGGCGAACGCGCCAATTACTGACATGCAGATGCGTCAGC
pRz A62P	CAGATGCGTCAGCGTGATGTTCCTGCGCTCGATGCAAAATACACGAAG
pRz T107P FOR	GTGAAGCCACCCCCGCCTCCGGCGTAGATAATG

pRz R125P	CTGGCAGACACCGCTGAACCGGATTATTTCACCCTCAGAGAGAGGC
pRz E150G	CAACTGGAAGGAACCCAGAAGTATATTTAGGAGCAGTGCAGATAGGG ATCC
pRE Rz Q151P	CCAGAAGTATATTAATGAGCCGTGCAGATAGGG
pRz Q151X	AACAACTGGAAGGAACCCAGAAGTATATTAATGAGTAATGCAGATAG GGATCCGTCGAC
pRz C152X	CTGGAAGGAACCCAGAAGTATATTAATGACCAGTAAAGATAGGGATC CGTCGACCTGC
pRz R153X	GGAACCCAGAAGTATATTAATGAGCAGTGCTAATAGGGATCCGTCGA CCTGCAG
pSynRz Linker 101-115 FOR	CGGCAGTGGTAGTGGTAGTGGAAGTCCACGGCTAGCGGAT
pSynRz Linker 101-115 REV	CTGCCGCTGCCACTACCGCTTCCTTGGCAAACCGCCTTAATATG
SynRz E150R SynRz1 R59E*	CCCTTAGGTACCAGAGAGATTGATGTATGAGCAGAGTCACCGCGATT ATCTCCGCTCTGGTTATCTGCATCATCGTCTGCCTGTCATGGGCTGT TAATCATTACCGTGATAACGCCATTACCTACAAAGCCCAGCGCGACA AAAATGCCAGAGAACTGAAGCTGGCGAACGCGGCAATTACTGACATG CAGATGCGTCAGCGTGATGTTGCTGCGCTCGATGCAAAATACACGAA GGAGTTAGCTGACGCCAAGGCGGAGAACGACGCGCTACGGGACGACG TGGCAGCCGGGCGGCGCCGATTACATATTAAGGCGGTTTGCCAATCC GTACGGGAGGCTACTACAGCAAGTGGAGTAGACAACGCGGCAAGTCC ACGGCTAGCGGATACTGCCGAGCGAGACTACTTTACACTTAGGGAAA GACTAATCACTATGCAAAAACAACTGGAAGGAACCCAGAAGTATATT AATAGGCAGTGCAGATAGGGATCCAAGGAGTTAGCTGATGCTTAAAC TCAAGATGATGCTATGTGTAATGATGTTACCACTTGTTGTAGTTGGG TGTACGAGTAAACAATCGGTATCGCAATGTGTAAAACCGCCCCCACC GCCTGCATGGATCATGCAACCGCCACCTGATTGGCAAACGCCACTAA ATGGAATCATATCGCCATCGGAAGAGGGATGAAAGCTTCTGTTTTG

pRz ART-TMD**	ATATGGTACCAGAGAGATTGATGTATGAGCAGAGTGGTGCTGCTGAT
	TATTGTGGTGGTGGTGGTGGTGGTGATTATTCTGCTGATTATTG
	TGCATTACCGTGATAACGCCATTACCTACAAAGCCCAGCGCGACAAA
	AATGCCAGAGAACTGAAGCTGGCGAACGCGGCAATTACTGACATGCA
	GATGCGTCAGCGTGATGTTGCTGCGCTCGATGCAAAATACACGAAGG
	AGTTAGCTGACGCCAAGGCGGAGAACGACGCGCTACGGGACGACGTG
	GCAGCCGGGCGGCGCCGATTACATATTAAGGCGGTTTGCCAATCCGT
	ACGGGAGGCTACTACAGCAAGTGGAGTAGACAACGCGGCAAGTCCAC
	GGCTAGCGGATACTGCCGAGCGAGACTACTTTACACTTAGGGAAAGA
	CTAATCACTATGCAAAAACAACTGGAAGGAACCCAGAAGTATATTAA
	TGAGCAGTGCAGATAGGGATCCGCG
pLinkerRz Q100S	GCCGATTACATATTAAGGCGGTTTGCTCAGGAAGCGGTAGTGGCAGC
pRz1 25-30 GS-Linker FOR	TCCGGCTCCAAGCCACCACCGCCTCCG
pRz1 25-30 GS-Linker REV	GCCGGAGCCCTGCTTTGATGTGCAACCGAC
pSynRz1 D45K FOR	AACCGCCACCTAAATGGCAAACGCC
pSynRz1 D45K REV	GGCGTTTGCCATTTAGGTGGCGGTT
pSynRz1 I54X	GATTGGCAAACGCCACTAAATGGAATCTAATCGCCATCGGAAAGGGG
	ATG
pSynRz1 S55X	GGCAAACGCCACTAAATGGAATCATATAGCCATCGGAAAGGGGATGA
	AAGC
pSynRz1 P56X FOR	GGAATCATATCGTGATCGGAAAGGGG
pSynRz1 P56X REV	CCCCTTTCCGATCACGATATGATTCC
pSynRz1 S57X FOR	CAATCATATCGCCATAAGAAAGGGGATGAAAG
pSynRz1 S57X REV	CTTTCATCCCCTTTCTTATGGCGATATGATC
pSynRz1 E58X	CCACTAAATGGAATCATATCGCCATCGTAAAGGGGATGAAAGCTTCT
	GTTTTG
pRz1-His P32Q FOR	CGTGAAGCAACCACCGCC
pRz1-His P32Q REV	GGCGGTGGTTGCTTCACG
pRz1-His P33L FOR	GTGAAGCCACTGCCGCCTCCGGCG

pRz1-His P33L REV	CGCCGGAGGCGGCAGTGGCTTCAC
pRz1-His P35H FOR	CCACCACCGCATCCGGCGTGG
pRz1-His P35H REV	CCACGCCGGATGCGGTGGTGG
pRz1-His P36Q FOR	CCACCACCGCCTCAGGCGTGGATAATG
pRz1-His P36Q REV	CATTATCCACGCCTGAGGCGCTGGTGG
pRz1-His I39V FOR	CCGGCGTGGGTAATGCAGC
pRz1-His I39V REV	GCTGCATTACCCACGCCGG
pRz1-His P44S FOR	CAGCCTCCCTCCGACTGGC
pRz1-His P44S REV	GCCAGTCGGAGGGAGGCTG
pRz1-His W46R FOR	CCCCCCGACCGGCAGACAC
pRz1-His W46R REV	GTGTCTGCCGGTCGGGGGG
pRz1-His L50P FOR	CAGACACCGCCGAACGGGATTATTTC
pRz1-His L50P REV	GAAATAATCCCGTTCGGCGGTGTCTG
pRz1-His R59E FOR	ATTTCACCCTCAGAGGAAGGCGGCCAC
pRz1-His R59 REV	TGGCCGCCTTACTCTGAGGGTGAAATAATCC

The “*” and “**” symbols indicate a dsDNA gblock (Integrated DNA Technologies) synthesized gene (Genscript) designed with a spanin allele flanked by restriction sites compatible with the pRE plasmid.

Table B.2. List of *Rz* and *Rz1* mutants grouped by their sub-structural regions.

The types of residue change, number of isolates obtained by random mutagenesis (positive integers), lysis function, and the relevant structural region of the mutant position are indicated in the table. A type of substitution between residue is indicated by an arrow symbol (→) for transition and by a double greater sign (>>) for transversion. Mutants indicated in bold were tested for their dominance in the presence of the corresponding wt allele. The ability of each allele either to support or block lysis by complementation is indicated by "+" or "-" symbol, respectively. Mutants created by site-directed mutagenesis are indicated by a "0" isolate number. Nonsense mutations are indicated by "X" and their type in parenthesis. "cdn"= codon

Rz mutants					Rz1 mutants				
cdn	change	isolates	lysis	region	cdn	change	isolates	lysis	region
1	**Met → Lys**	**1**	**-**		**1**	**Met → Lys**	**1**	**-**	
4-24	**Artificial TMD**	**0**	**+**		**12**	**Met » Arg**	**3**	**-**	
14	**Cys → Arg**	**2**	**-**	**TMD**	**19**	**Gly » Cys**	**3**	**-**	**Signal sequence**
19	**Leu → Pro**	**2**	**-**		**20**	**Cys » Ser**	**5**	**-**	
20	**Ser → Pro**	**0**	**+**		**20**	**Cys → Arg**	**1**	**-**	
36	**Gln → Pro**	**0**	**+**		**20**	**Cys » Phe**	**2**	**-**	
50	**Ala → Pro**	**0**	**+**		**20**	**Cys → Tyr**	**3**	**-**	**L**
57	**Arg → Cys**	**2**	**-**		**25-30**	**(Gly-Ser)$_3$**	**0**	**+**	
59	**Arg → Cys**	**1**	**-**		32	Pro » Gln	**1**	**-**	
61	Val → Ala	**1**	**-**		33	Pro →Leu	**1**	**-**	
62	Ala → Pro	**0**	**-**		35	Pro » His	**3**	**-**	
64	Leu » His	**2**	**-**	**CC1**	36	Pro » Gln	**3**	**-**	
64	**Leu → Pro**	**3**	**-**		**36**	**Pro → Leu**	**2**	**-**	**Proline-rich-region**
64	Leu » Arg	**1**	**-**		**39**	**Ile → Val**	**1**	**-**	
65	Asp → Gly	**2**	**-**		**42**	**Pro → Ser**	**1**	**-**	
65	**Asp → Asn**	**1**	**-**		44	Pro → Ser	**1**	**-**	
65	**Asp » Val**	**1**	**-**		**45**	**Asp → Lys**	**0**	**+**	

66	**Ala » Pro**	**1**	**-**	
67	**Lys » Asn**	**1**	**-**	
70	**Lys → Glu**	**1**	**-**	
72	Leu » Phe	**1**	**-**	
72	**Leu → Ser**	**1**	**-**	
73	**Ala » Pro**	**1**	**-**	
73	Ala → Val	**1**	**-**	
77	**Ala » Pro**	**1**	**-**	
82	**Leu → Pro**	**1**	**-**	
83	**Arg » Pro**	**1**	**-**	
86	Val » Gly	**1**	**-**	
88	**Ala » Pro**	**2**	**-**	
91	Arg » Pro	**2**	**-**	**L1**
93	Leu → Ser	**3**	**-**	
100 - 115	**(Gly-Ser)$_8$**	**0**	**+**	
107	**Thr→Pro**	**0**	**+**	
125	**Arg→ Pro**	**0**	**+**	
127	Tyr » Asn	**1**	**-**	
134	**Leu → Pro**	**3**	**-**	**CC2**
141	**Leu → Pro**	**1**	**-**	
143	Gly » Arg	**1**	**-**	
147	**Tyr → His**	**1**	**-**	
147	Tyr → Cys	**2**	**-**	
150	**Glu→ Arg**	**0**	**-**	

46	Trp » Cys	**1**	**-**	
46	Trp →Arg	**2**	**-**	
50	**Leu » Arg**	**1**	**-**	**CTD**
50	Leu →Pro	**7**	**-**	
54	Ile » Asn	**1**	**-**	
54	**Ile → X (Ochre)**	**0**	**-**	
55	**Ser → X (Amber)**	**0**	**+**	
56	Pro → X (Opal)	**0**	**+**	
57	Ser → X (Ochre)	**0**	**+**	
58	Glu→Lys	**1**	**-**	
58	**Glu → X (Ochre)**	**0**	**+**	
59	**Arg → Glu**	**0**	**+**	
59	**Arg →X (Ochre)**	**0**	**+**	
60	**Gly →X (Ochre)**	**0**	**+**	

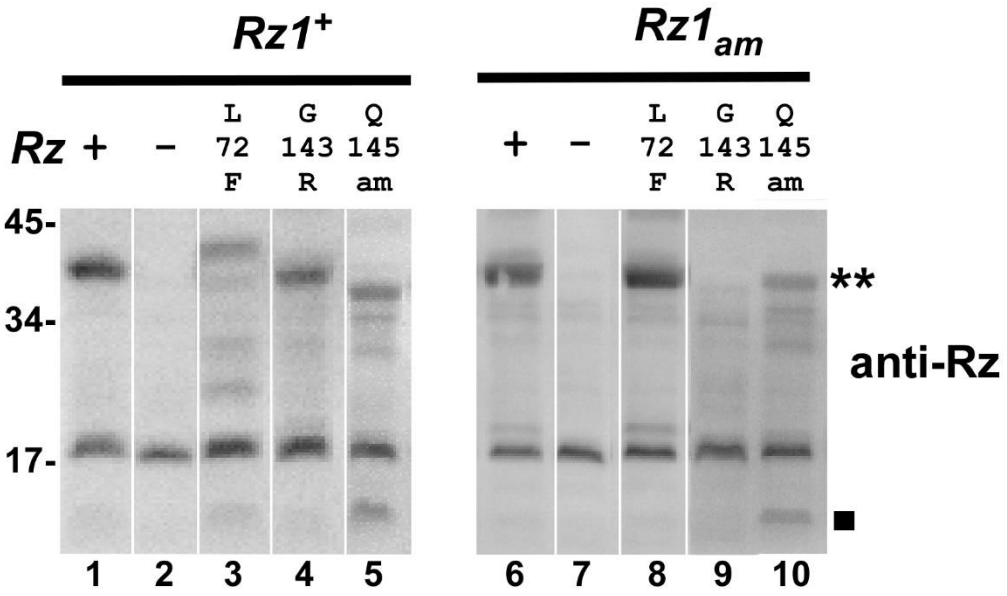


Figure BS.1.

Accumulation defects of Rz mutations in the presence and absence of Rz1.

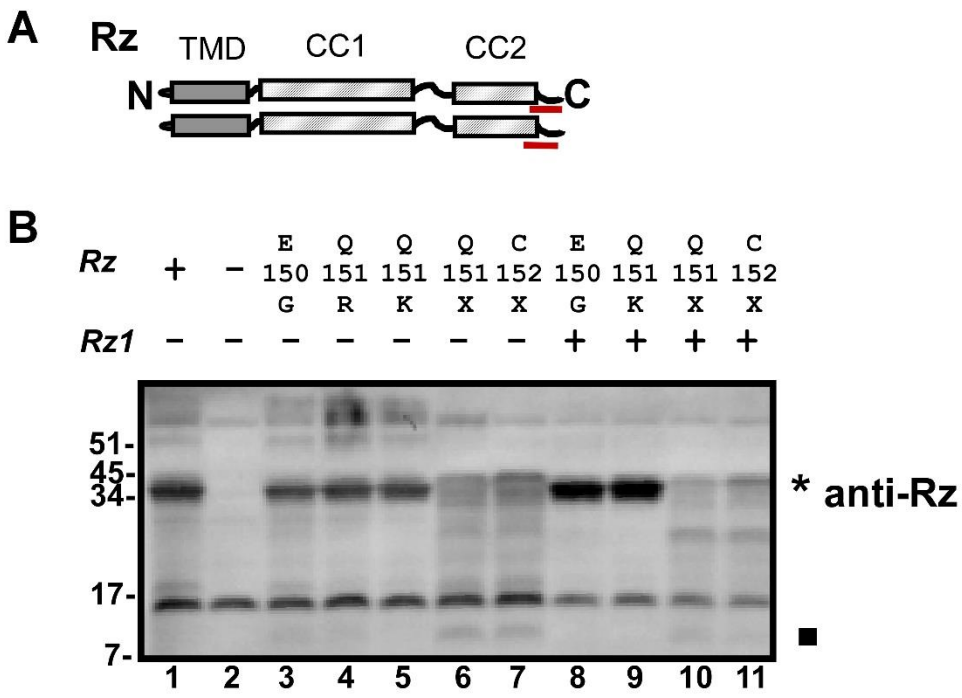


Figure BS.2.

Accumulation of Rz with mutations in the C terminal domain in the presence or absence of Rz1.

APPENDIX C

FIGURES AND TABLES

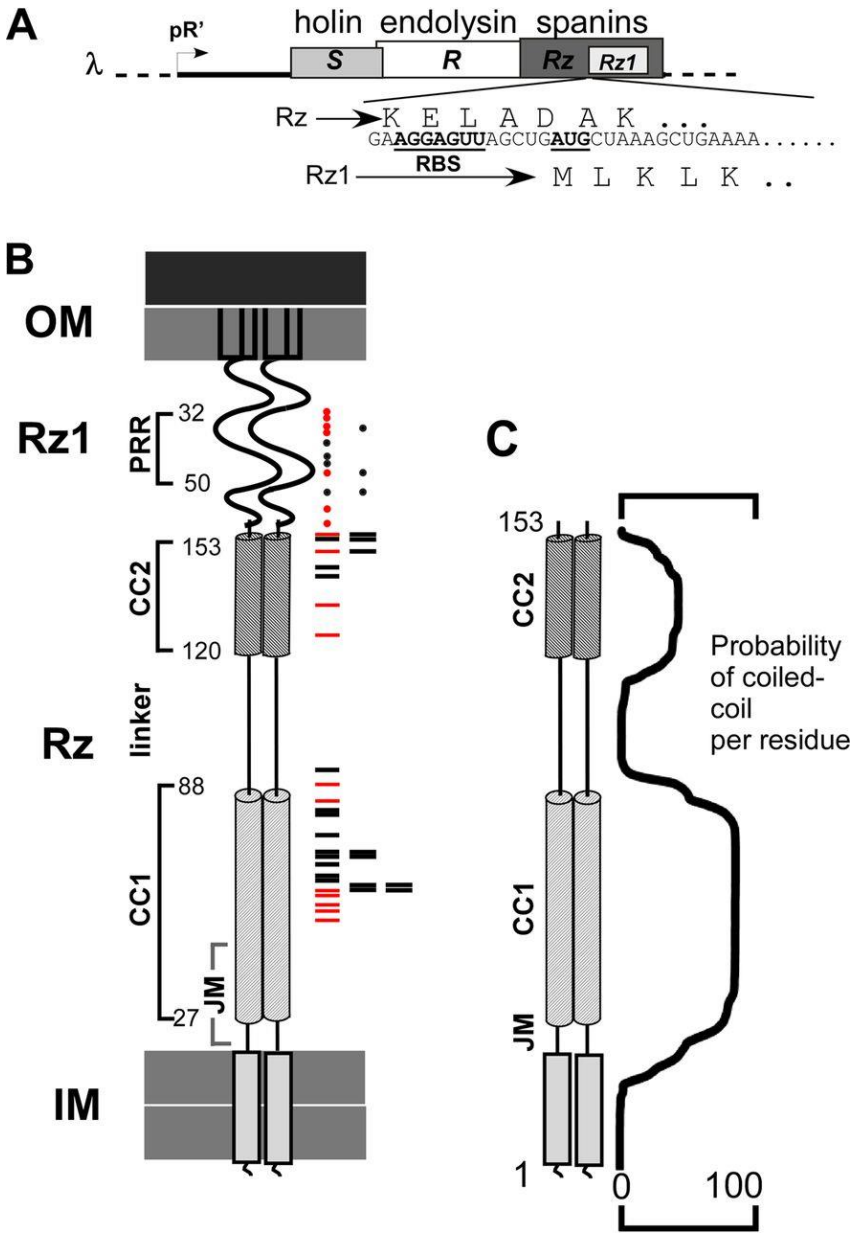


Figure C.1. Features of λ spanins.

(A) The lysis cassette is shown. Lysis genes *S* (holin), *R* (endolysin), and *Rz/Rz1* (the spanins) are under the control of the late promoter pR'. *Rz1* is embedded in the +1 reading frame of *Rz*. The Ribosome Binding Site (RBS) and translated products of *Rz*

and *Rz1* are shown at the start of the embedded *Rz1* reading frame. (B) A cartoon of spanin complex is shown within the cell envelope. The outer membrane (OM), inner membrane (IM), and predicted structural features of Rz and Rz1 are denoted on the left side of the cartoon. The relative positions of single missense mutations that inactivate Rz are represented by red and black rectangles on the right side of the cartoon. Red rectangles represent the relative position of single missense mutations that were used in the suppressor screen. The relative positions of single missense mutations that inactivate Rz1 are represented by red and black circles on the right side of the cartoon. Red circles represent the relative position of single missense mutations that were used in the suppressor screen. (C) A cartoon of Rz is shown with the probability of coiled-coil per residue predicted by MARCOIL to the right of the cartoon.

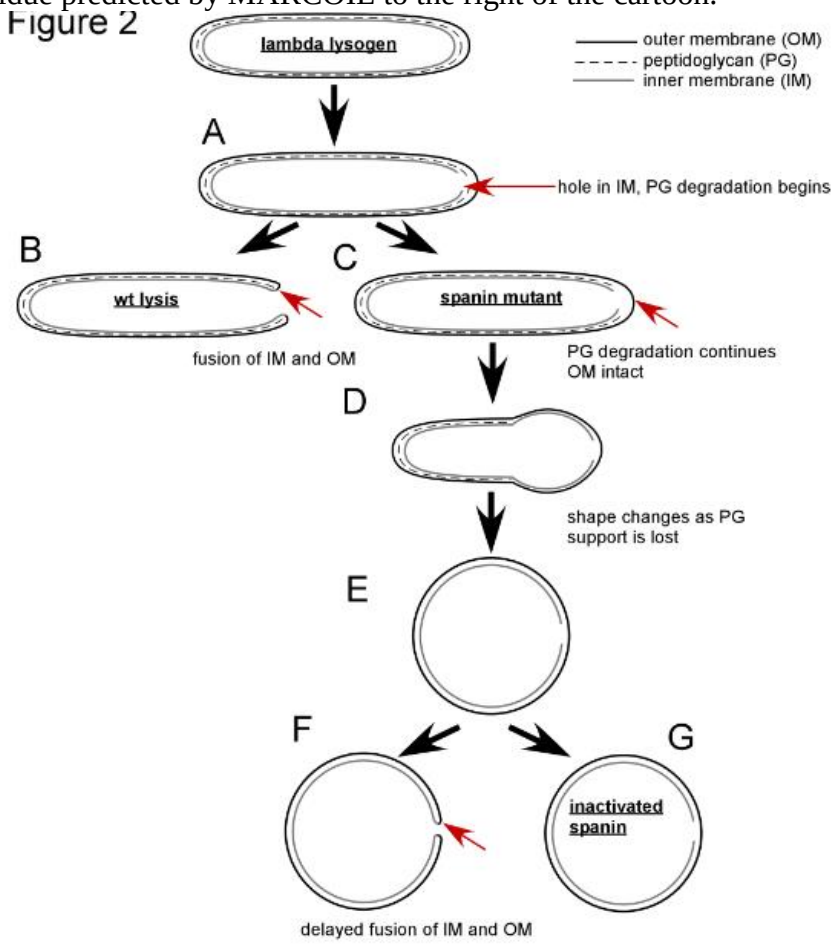


Figure C.2. Model comparing wild type to spanin mutant lysis.

(A) The first step of lysis is holin-mediated hole formation in the IM. After IM permeabilization, the endolysin degrades its substrate, PG (B) In wild type λ, lysis occurs from the poles before loss of rod shape. Lysis is complete after OM disruption. Arrow indicates the putative spanin-mediated fusion of the IM and OM. (C) In the case of spanin mutants, the OM remains intact (marked by the arrow) as the PG is degraded.

(D) As PG degradation continues, cell shape is lost. (E) The cell assumes a spherical shape after complete PG degradation. (F) Lysis mediated by kinetically defective spanins occurs from a spherical cell, arrow indicates site of putative OM-IM fusion. (G) The terminal morphological phenotype of cells carrying non-functional spanins.

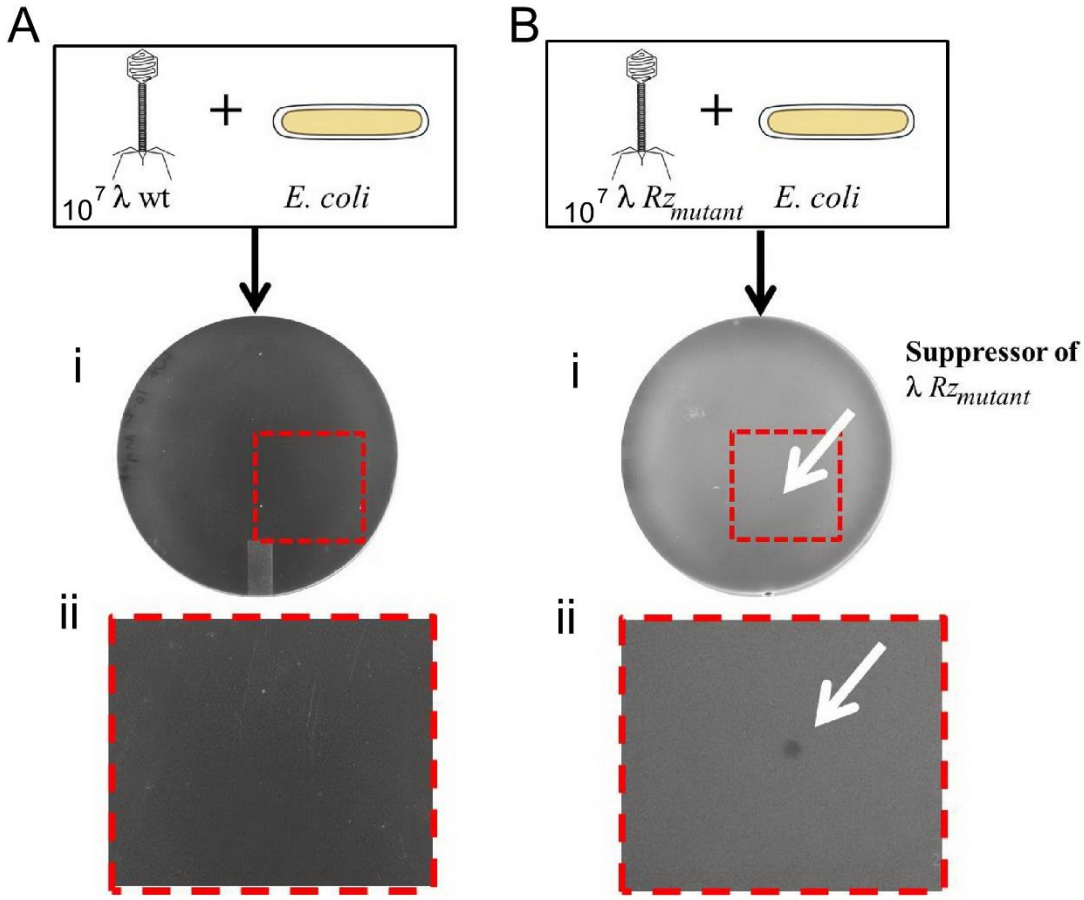


Figure C.3. Suppressors are selected by early plaque formation.

(A) i. Confluent lysis after plating ~10^7 λ910 phage. ii. Magnification of the region of the plate indicated by the dashed red box. (B) i. Plating of ~10^7 λ910 *Rz* mutant phage. Plaque formation due to spontaneous pseudoreversion of an *Rz* mutation is indicated by

the white arrow. ii. Magnification of the region of the plate indicated by the dashed red box.

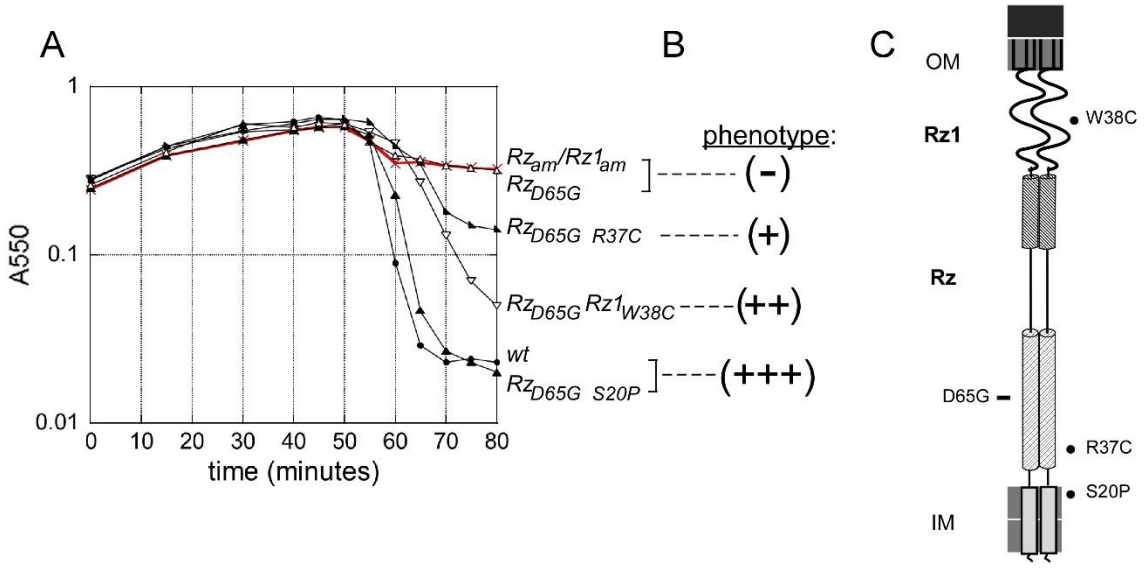


Figure C.4. Subset of lysis profile data used to assess lysis phenotype.

(A)The following lysogens were induced at time = 0 and monitored at A550: MDS12 (λ900$Rz_{am}Rz1_{am}$) (red), MDS12 (λ910Rz_{D65G}), MDS12 (λ910$Rz_{D65G,\ R37C}$), MDS12 (λ910Rz_{D65G}; $Rz1_{W38C}$), MDS12 (λ910), and MDS12 (λ910$Rz_{D65G,\ S20P}$). (B) Categorization of lytic phenotype is denoted respective to the lysis profile traces. (C) Cartoon of the spanin complex within the cell envelope. The relative position of the inactivating Rz_{D65G} mutation is marked with a rectangle to the left of the cartoon. Right of cartoon: the relative positions of suppressors generated from Rz_{D65G}, marked with a filled circle.

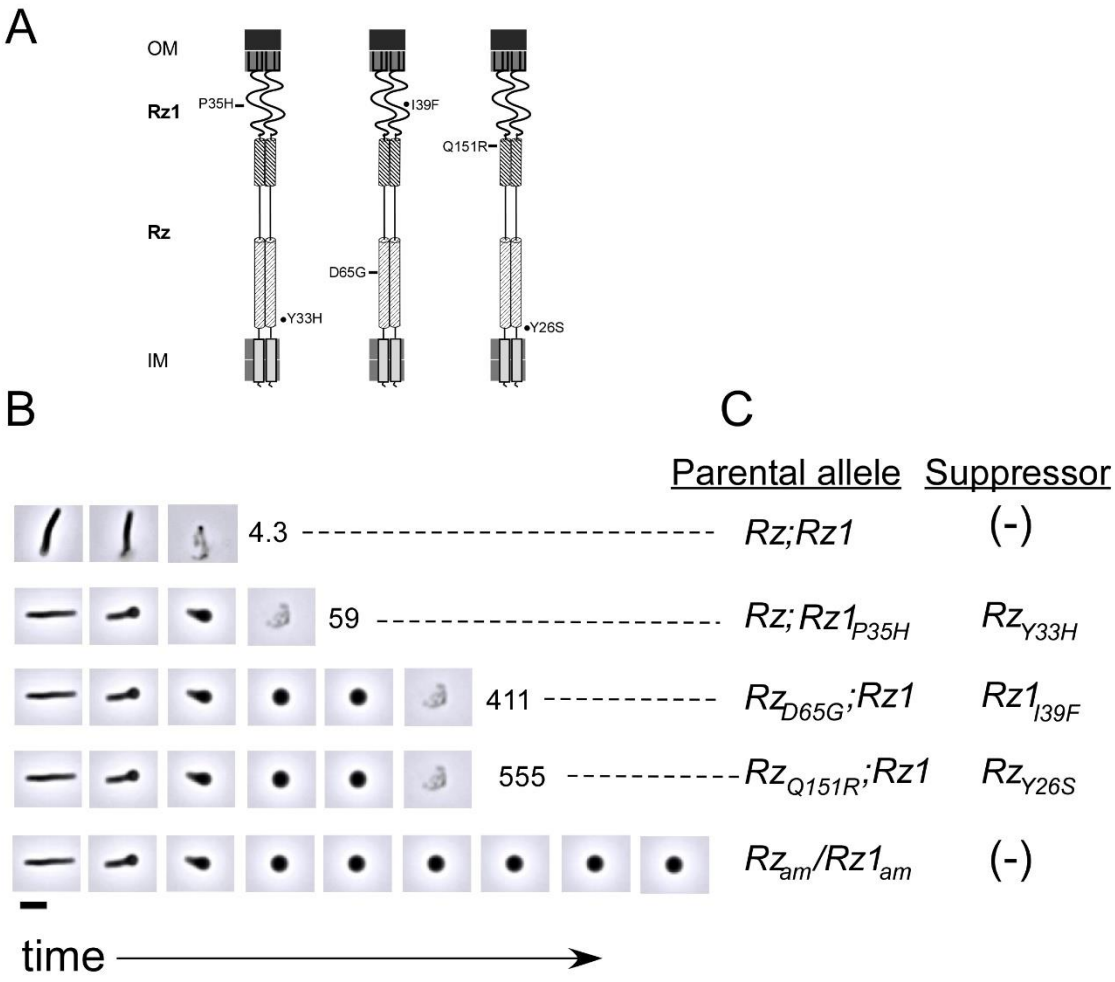


Figure C.5. Suppressors exhibit kinetic and morphological defects in lysis.

(A) Cartoon of the spanin complex within the cell envelope. Relative positions of inactivating mutations used below are denoted to the left of the complex by rectangles. Relative positions of suppressor mutations generated from the parental mutant are denoted to the right of the cartoon by circles.
(B) Lysis was monitored by measuring the time from deformity to lysis. Average time is indicated in seconds and represented by the bar below the micrographs. Red whiskers indicate the standard deviation for the sample. Time lapse micrographs are representative of the lysis morphology observed in each set. (C) The parental allele, suppressor mutation (if present), and type are indicated. Inter=Intergenic suppressor. Intra= Intragenic suppressor. Scale bar = 5 µm.

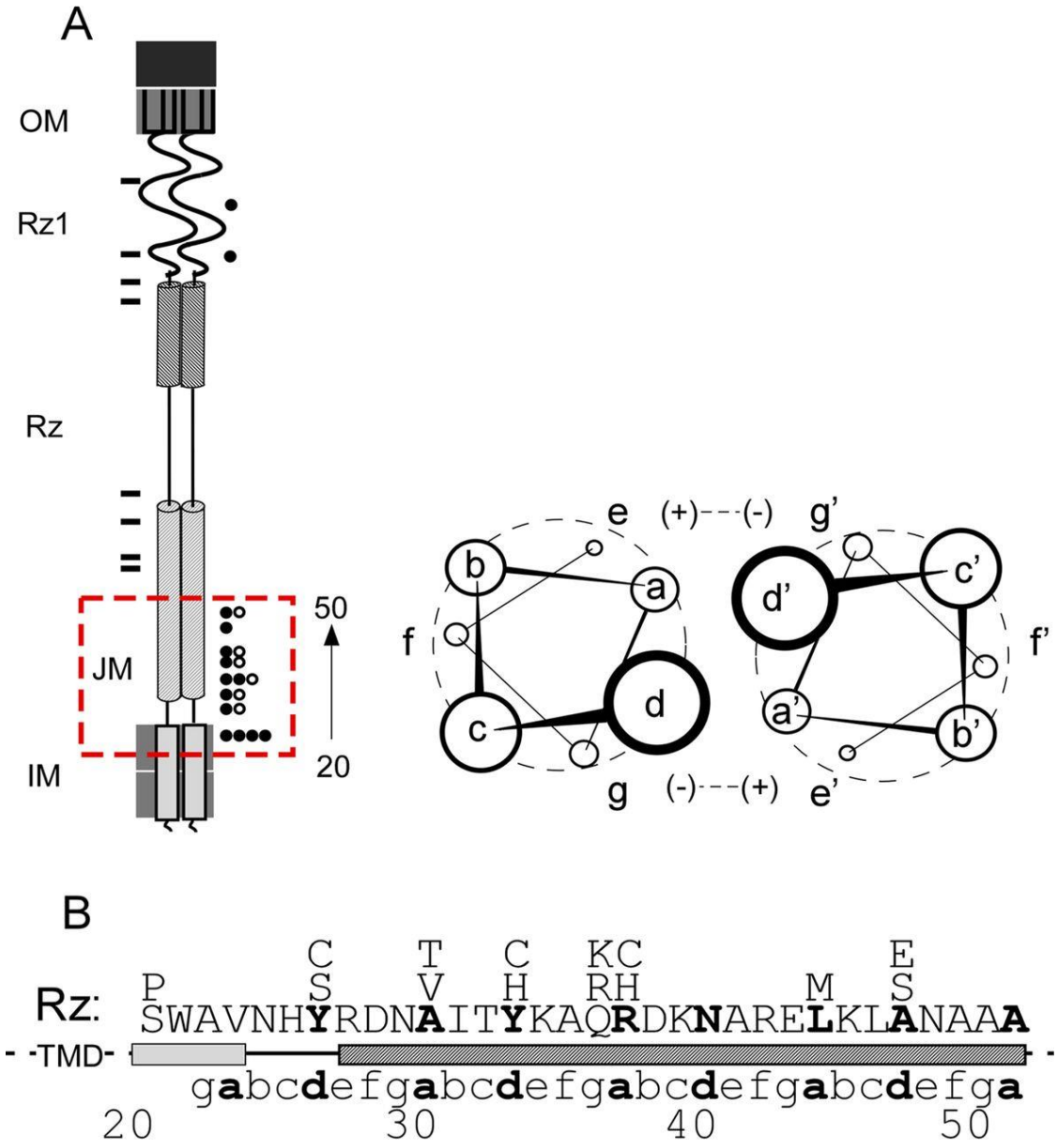


Figure C.6. Suppressors cluster within the core of the juxtamembrane region.

(A) Cartoon of the spanin complex within the cell envelope. Relative positions of inactivating mutations used in the suppressor screen are denoted to the left of the cartoon by rectangles. Relative positions of second-site suppressing changes with wild type lysis phenotype are denoted to the right of the cartoon by circles. Open circles indicate different suppressing alleles that fall within the same position (e.g. A30T and A30V). The dashed red box indicates juxtamembrane (JM) region, approximately residues 20-50. (B) Cartoon of parallel coiled-coil helices oriented so the viewer is looking down

the helical barrels from the N-terminus. Electrostatic attraction is shown between *e* and *g* positions (positive and negative signs), however polar interaction is also supported in these positions. (C) The amino acid sequence of the JM region of Rz. Positions of suppressors that lysed bulk culture as rapidly as wild type (see Table C.3) are denoted above the Rz JM sequence. The heptad repeats within Rz are aligned respective to MARCOIL predictions. Bolded letters indicate positions predicted to be within the hydrophobic core (the *a* and *d* positions).

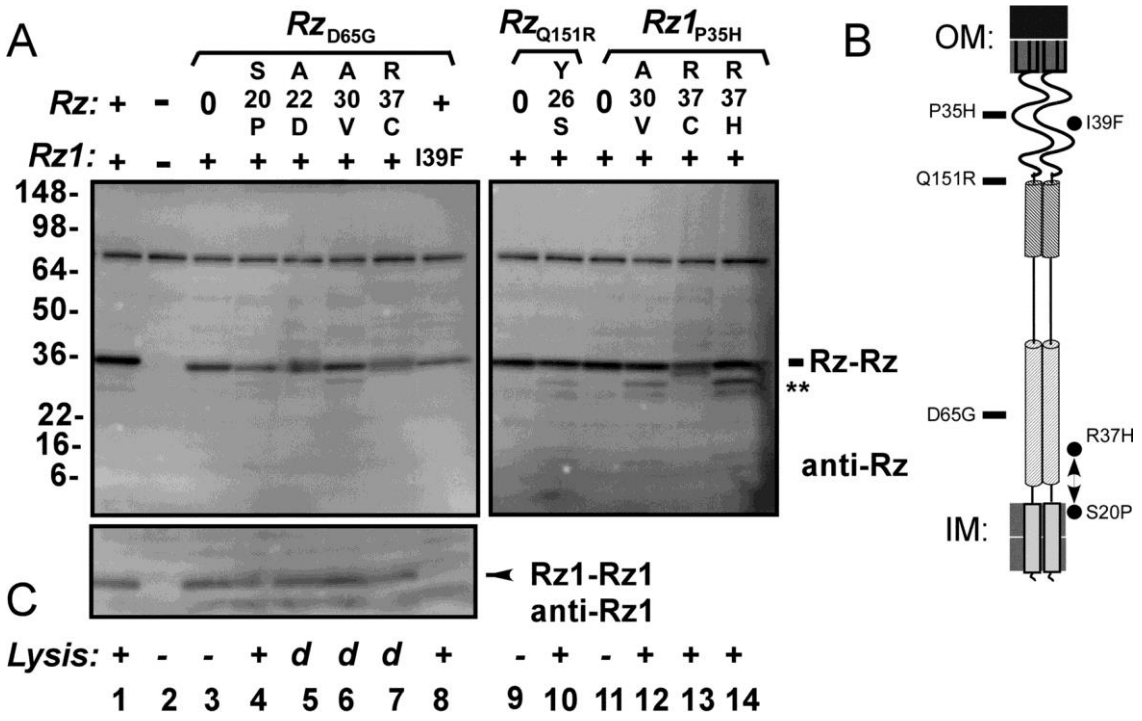


Figure C.7. Suppression reduces Rz stability.

Western blot of Rz and Rz1. MDS12 (λ900) and MDS12 (λ910) lysogens were thermally induced and a one mL sample of cells was collected for TCA precipitate prior to lysis. (A) The presence and allele of *Rz* or *Rz1* is denoted above the blot. (-) indicates λ900$Rz_{am}Rz1_{am}$. Mutant and suppressed spanin alleles are expressed from MDS12 (λ910) lysogens. Parental phages carrying mutants Rz_{D65G}, Rz_{Q151R}, and $Rz1_{P35H}$, are bracketed and suppressing alleles are indicated below each bracket. "0" indicates samples collected from MDS12 (λ910) expressing the non-suppressed (parental) allele. The position of the covalently linked Rz dimer is indicated by (Rz-Rz). Double asterisks indicate apparent proteolytic products. SeeBluePlus2 molecular markers in kDa are indicated to the left. (B) Cartoon of the spanin complex within the cell envelope. Relative positions of inactivating mutations are denoted to the left of the cartoon by rectangles. Relative positions of suppressor mutations are denoted to the right of the cartoon by circles. (C) The symbols "+" indicates wild type lysis rate, "d" indicates a delay in the rate of lysis (see Table C.3), and "-" indicates a non-lytic phenotype.

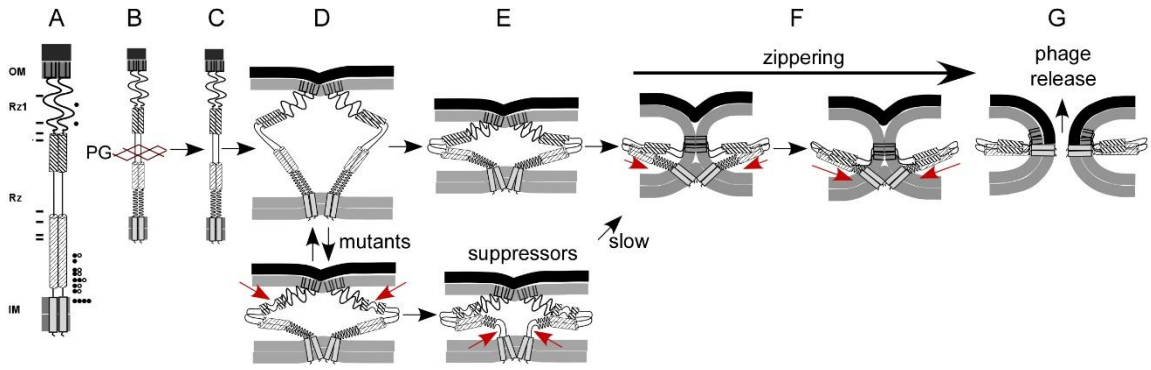


Figure C.8. Model of suppressor function.

(A) Cartoon of the spanin complex within the cell envelope. Relative positions of inactivating mutations used in the suppressor screen are denoted to the left of the cartoon by rectangles. Relative positions of second site suppressing changes with wild-type lysis rate are denoted to the right of the cartoon by circles. Open circles indicate different suppressing alleles that fall within the same position (e.g. A30T and A30V). (B) Prior to lysis the spanin complex accumulates threaded within the lacuna of PG. (C) PG degradation results in spanin activation. (D) Top: The spanin complex undergoes a collapsing conformational change. Bottom: Red arrow indicates possible disruption of spanin structure by an inactivating mutation. Mutations that encode defects in spanin structure prevent further structural rearrangement and spanin function is halted. (E) Top: The complex begins to fold into a hairpin structure. Bottom: Suppression restores function by increasing the flexibility within the juxtamembrane region (arrows). This allows the complex to resume hairpin formation. (F) The juxtamembrane region becomes structured as the complex zippers into a hairpin conformation (red arrows). (G) Phage release occurs upon fusion pore formation.

Table C.1. Phages, strains, plasmids, and primers used in this study.

Bacteriophages	Genotypes and relevant features	Sources
λ900	λ*Δ(stf tfa)::cat* cI_{857} *bor::kan;* carries Cam^R and Kan^R ; wild type *RzRz1*	(98)
λ900Rz_{am} $Rz1_{am}$	λ*Δ (stf tfa)::cat* cI_{857} Rz_{Q100am} $Rz1_{W38am}$ *bor::kan*	(98)
λ*cI857*	λcI_{857} *SRRzRz1*	Laboratory stock
λ910	λcI_{857} *bor::kan;* synthetic alleles of *Rz* and *Rz1* (^{Syn}*Rz;* ^{Syn}*Rz1).* The *Rz1* gene is located downstream of *Rz (*Genbank Accession number = KY609225).	This study
Strains		
MDS12	MG1655 with 12 deletions, totaling 376,180 nt, including all cryptic prophages	(103)
RY17341	MDS12*ΔfhuA;*	(140)
MDS12(λ900)	RY17341 lysogenized with λ900	This study
MDS12(λ900$Rz_{am}Rz1_{am}$)	RY17341 lysogenized with λ900$Rz_{am}Rz1_{am}$	This study
MDS12(λ910*)*	RY17341 lysogenized with λ910	This study
BW25113	*E. coli* Keio collection parental strain	(141)
BW25113(λcI_{857})	BW25113 lysogenized with λcI_{857}	This study
Plasmids		
pER157	pBR322*Δtet (SRRzRz1 bor::kan)*	(98)
pAK1	pER157 with tandem synthetic *Rz* and *Rz1* genes	This study

Primer	Sequence 5' - 3'
pRz S20P FOR	TCTGCCTGCCATGGGCTGTTAATC
pRz S20P REV	GATTAACAGCCCATGGCAGGCAGA
pRz V61A FOR	TCAGCGTGATGCTGCTGCGCTCGATGC
pRz V61A REV	GCATCGAGCGCAGCAGCATCACGCTGA
pRz L64H FOR	TGTTGCTGCGCACGATGCAAAAT
pRz L64H REV	ATTTTGCATCGTGCGCAGCAACA
pRz D65G FOR	GCTGCGCTCGGTGCAAAATACAC
pRz D65G REV	GTGTATTTTGCACCGAGCGCAGC
pRz K70E FOR	GCAAAATACACGGAGGAGTTAGCTG
pRz K70E REV	CAGCTAACTCCTCCGTGTATTTTGC
pRz L72F FOR	ACGAAGGAGTTCGCTGACGCC
pRz L72F REV	GGCGTCAGCGAACTCCTTCGT
pRz V86G FOR	ACGGGACGACGGGGCAGCCGGG
pRz V86G REV	CCCGGCTGCCCCGTCGTCCCGT
pRz L93S FOR	GCGGCGCCGATCACATATTAAGG
pRz L93S REV	CCTTAATATGTGATCGGCGCCGC
pRz Y127N FOR	AGCGAGACAACTTTACACTTAGGGAAAG
pRz Y127N REV	TCCCTAAGTGTAAAGTTGTCTCGCTCGGC
pRz G143R FOR	ACAACTGGAACGAACCCAGAAGTAT
pRz G143R REV	ATACTTCTGGGTTCGTTCCAGTTGT

pRz Y147H FOR	AACCCAGAAGCATATTAATGAGC
pRz Y147H REV	GCTCATTAATATGCTTCTGGGTT
pRz Q151R FOR	ATTAATGAGCGGTGCAGATAG
pRz Q151R REV	CTATCTGCACCGCTCATTAAT
pRz1 P32Q FOR	CGTGAAGCAACCACCGCC
pRz1 P32Q REV	GGCGGTGGTTGCTTCACG
pRz1 P33L FOR	GTGAAGCCACTGCCGCCTCCGGCG
pRz1 P33L REV	CGCCGGAGGCGGCAGTGGCTTCAC
pRz1 P35H FOR	CCACCACCGCATCCGGCGTGG
pRz1 P35H REV	CCACGCCGGATGCGGTGGTGG
pRz1 P36Q FOR	CCACCACCGCCTCAGGCGTGGATAATG
pRz1 P36Q REV	CATTATCCACGCCTGAGGCGCTGGTGG
pRz1 W46R FOR	CCGCCACCTGATCGGCAAACGCCAC
pRz1 W46R REV	GTGGCGTTTGCCGATCAGGTGGCGG
pRz1 I54N FOR	CTAAATGGAATCAATTCGCCATCGGAAAG
pRz1 I54N REV	CTTTCCGATGGCGAATTGATTCCATTTAG
pAK Rz1 E58K	CCACTAAATGGAATCATATCGCCATCGAAAAGGGGATGAAAGCTTGCCC
Supp Seq REV	GAAGCGCGTGTGTATTGCTCAC
Lbd R S121 FWD	CAGCAATATCTGGGCTTCACTGC
AphI kan REV	GCAGTTTCATTTGATGCTCGATGAGTTTTTC
RzRz1 Seq FOR	CAAAATTCAAAGAAGCGGGCGGAAC
pAK Rz R37C	AAGCCAGTGCGACAAAAATCC
pAK Rz R37C	GCATTTTTGTCGCACTGGGC

Table C.2. Suppressors of Rz and Rz1 mutants.

Inactivating mutations used for the suppressor screen are indicated along with corresponding codon changes. Suppressors are grouped according to parental inactivating mutant. Arrow symbols (→) indicate transversions and double greater than signs indicate (>>) transitions.

Suppressors of Rz mutants			
Original Mutation	**Codon change**	**Suppressor**	**Codon change**
V61A	GTT → GCT	A30T	GCC → ACC
		Y33C	TAC → TGC
		R37C	CGC → TGC
		A47S	GCG » TCG
L64H	CTC » CAC	S20P	TCA → CCA
		D38G	GAC → GGC
		H64L	CAC » CTC
D65G	GAT → GGT	S20P	TCA → CCA
		A22D	GCT » GAT
		R27L	CGT » CTT
		A30V	GCC → GTC
			CGC → TGC

		R37C Rz1 W38C Rz1 W38S Rz1 I39F	TGG » TGT TGG » TCG ATC » TTC
L72F	TTA » TTC	F72L	TTC » TTA
V86G	GTG » GGG	R37C A47E	CGC → TGC GCG » GAG
L93S	TTA → TCA	Y26C A30T L44M S93L Rz1 P56T	TAC → TGC GCC → ACC ATG » CTG TCA → TTA CCA » ACA
G143R	GGA » CGA	R143L	CGA » CTA
Y147H	TAT → CAT	S20P Y33C Q36K L44M	TCA → CCA TAC → TGC CAG » AAG ATG » CTG
Q151R	CAG → CGG	S20P Y26S	TCA → CCA TAC » TCC

Suppressors of Rz1 mutants			
Original Mutation	**Codon change**	**Suppressor**	**Codon change**
P33L	CCC → CTC	Rz Q36R	CAG → CGG
P35H	CCG » CAC	Rz A30V Rz Y33H Rz R37C Rz R37H	GCC → GTC TAC → CAC CGC → TGC CGC → CAC
W46R	TGG → AGA	R46I R46G	AGA » ATA AGA → GGA
L50P	CTA → CCA	P50L	CCA → CTA
I54N	ATA » AAC	P56L Rz R37C Rz R37H	CGC → CAC CGC → TGC CGC → CAC

Table C.3. Second site suppressors of Rz and Rz1.

List of inactivating mutant, compensatory second-site mutation, and phenotype of lysis within liquid media are indicated in the table.

Rz Mutation	Suppressor	Phenotype
V61A	A30T	+++
	Y33C	+++
	A47S	+++
L64H	S20P	+++
	D38G	++
D65G	S20P	+++
	A22D	+
	R27L	++
	A30V	+
	R37C	+
	Rz1 W38C	++
	Rz1 W38S	++
	Rz1 I39F	+++
V86G	R37C	++
	A47E	+++
L93S	Y26C	++
	A30T	+
	L44M	+
	Rz1 P56T	++
Y147H	S20P	+++
	Y33C	
	Q36K	

	L44M	+++ +++ +++
Q151R	S20P Y26S	+++ +++

Rz1 Mutation	**Suppressor**	**Phenotype**
P33L	Rz Q36R	+++
P35H	Rz A30V Rz Y33H Rz R37C Rz R37H	+++ +++ +++ +++
I54N	Rz R37C Rz R37H P56L	+ ++ +++

Table C.4. Densitometry of Rz western blots

Lanes and values are arranged corresponding to Fig. C7. The density value of the band corresponding to the expected Rz molecular mass is reported. Each density measurement was normalized to that of the loading control. The highest value of each blot was set to 1. Density readings were proportionally scaled to this value.

Lane[a]	Sample[a]	Density[b]
1	λ910	1
2	λ900 Rz_{am} $Rz1_{am}$	0.027
3	λ910 Rz_{D65G}	0.62
4	λ910 $Rz_{D65G\ S20P}$	0.31
5	λ910 $Rz_{D65G\ A22D}$	0.51
6	λ910 $Rz_{D65G\ A30V}$	0.51
7	λ910 $Rz_{D65G\ R37C}$	0.39
8	λ910 Rz_{D65G} $Rz1_{I39F}$	0.34
9	λ910 Rz_{Q151R}	0.98
10	λ910 $Rz_{Q151R\ Y26S}$	0.82
11	λ910 $Rz1_{P35H}$	1.0
12	λ910 Rz_{A30V} $Rz1_{P35H}$	0.87
13	λ910 Rz_{R37C} $Rz1_{P35H}$	0.85
14	λ910 Rz_{R37H} $Rz1_{P35H}$	0.73

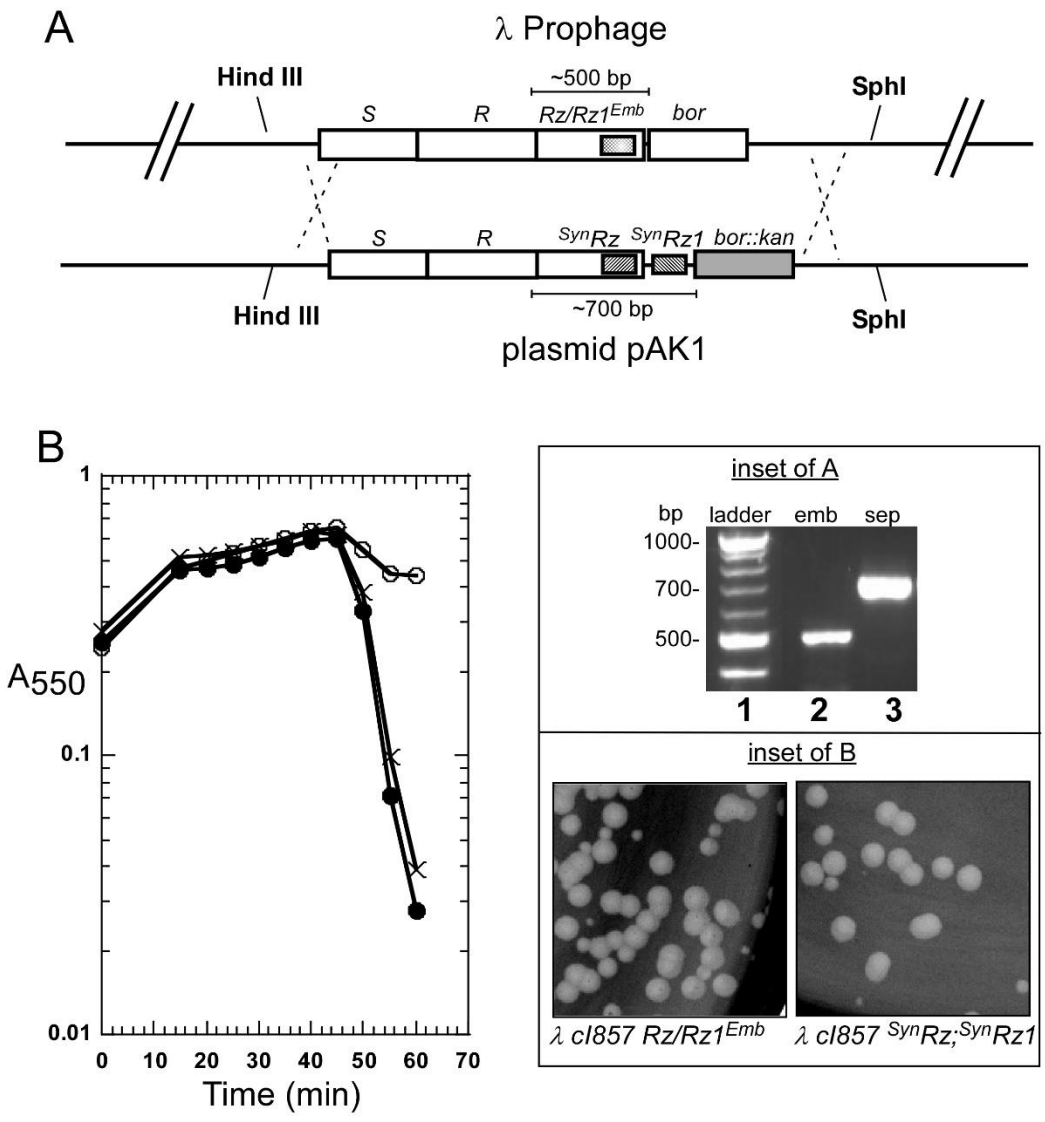


Figure CS.1

Construction of phages carrying separated spanin genes. (A) The plasmid pAK1 was used to recombine the synthetic, separated Rz and Rz1 genes into phage λ. The desired recombinants carry the separated *Rz* and *Rz1* genes along with the *bor*::*kan* marker. A DNA gel (inset) shows the replacement of the wild-type embedded spanins with the synthetic separated spanin genes. An empty well between lanes 1 and 2 was cut from the image and replaced by a vertical white line. (B) Lysis profile demonstrating the lytic

phenotype of phages that carry separated spanins. The following lysogens were induced at time 0 and monitored at A_{550}: ×, MDS12 (λ900); open circles, MDS12 (λ900 Rz_{am} $Rz1_{am}$); filled circles, MDS12 (λ910), which carries the separated spanin genes. (C) Comparison of λ900 and λ910 plaques. (D) Western blots comparing Rz and Rz1 expressed from embedded (wild type; wt) and synthetic (Syn) gene arrangements. −, expression from λ900 Rz_{am} $Rz1_{am}$. Arrows indicate the positions of the Rz and Rz1 bands. LC, loading control (a nonspecific band detected by the Rz antibody).

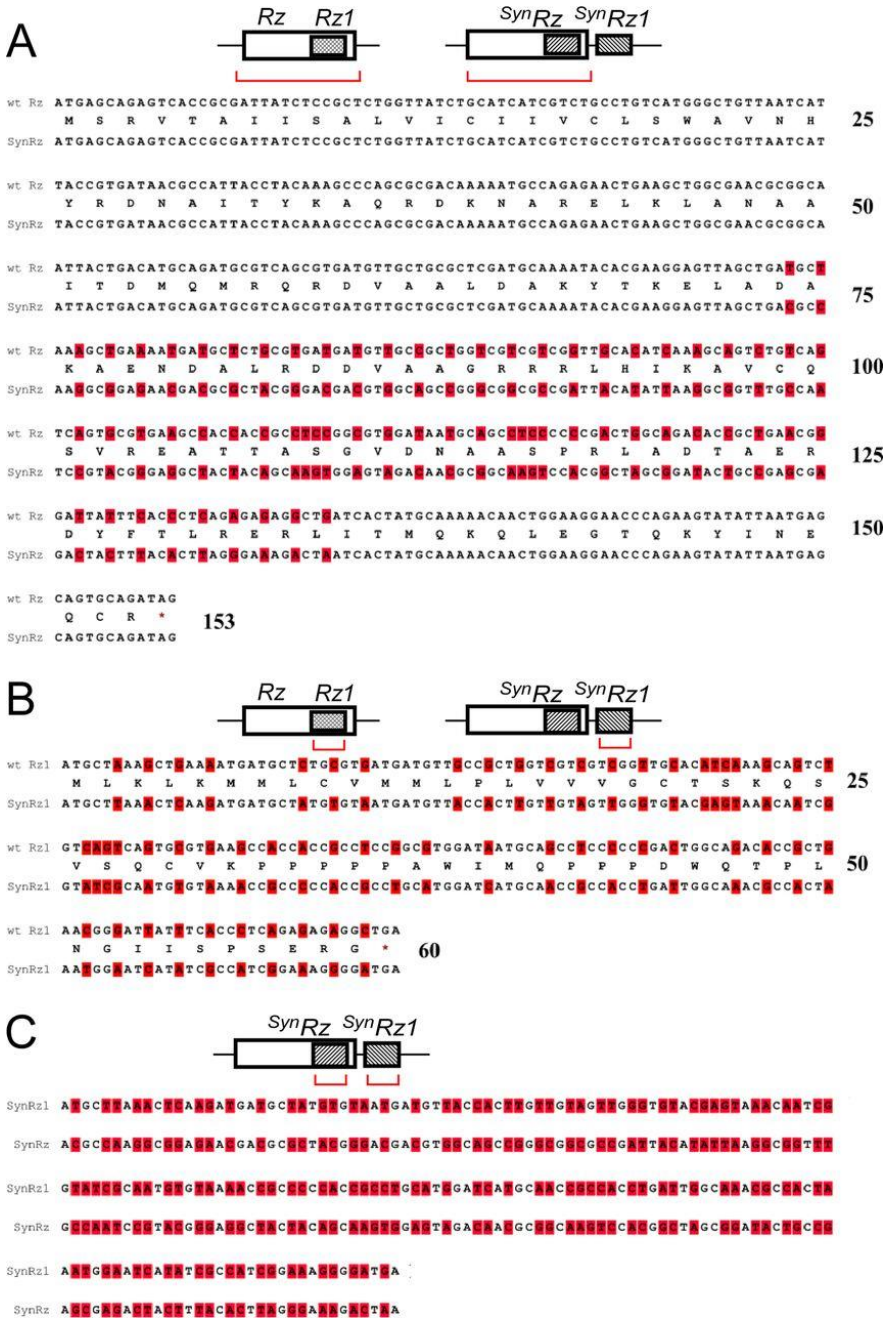


Figure CS.2

Comparison of embedded and synthetic spanin DNA sequences. A cartoon of wild-type and synthetic spanin gene arrangements is shown at the top of each panel, and red

brackets indicate the genes compared in each panel. In each case, the amino acid sequences of synthetic and wild-type genes are the same. (A) Comparison of *Rz* and *SynRz*. The nucleotides that differ between *Rz* and *SynRz* are highlighted. (B) Comparison of *Rz1* and *SynRz1*. The nucleotides that differ between *Rz1* and *SynRz1* are highlighted. (C) Comparison of *SynRz* and *SynRz1*. The region of *SynRz* that was codon optimized to be distinct from *SynRz1* is shown. The nucleotides that differ between *SynRz* and *SynRz1* are highlighted.

APPENDIX D

FIGURES AND TABLES

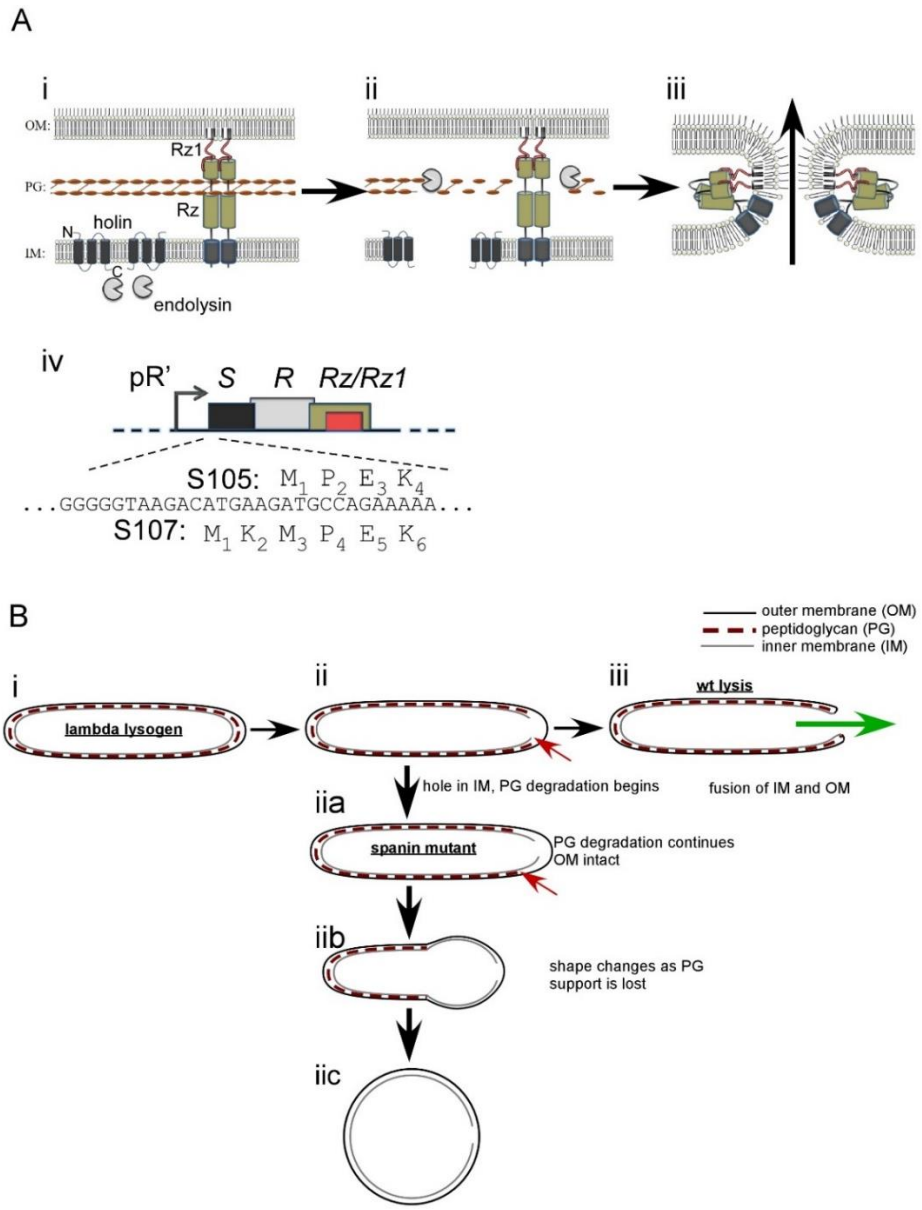


Figure D.1. Overview of the λ lysis system

(A) Lambda lysis genes and model for lysis protein function. (i) The topology and localization of lysis genes before lysis. (ii) The holin forms holes in the inner membrane (IM), which releases endolysin into the periplasm. Endolysin degrades peptidoglycan

(PG). (iii) The spanin complex undergoes a conformational change that fuses the IM and outer membrane (OM), disrupting the OM. Arrow shows direction of phage progeny release. (iv) The λ lysis cassette. The translational starts for the holin and antiholin (*S105* and *S107*) genes are shown above and below the DNA sequence, respectively. (B) Model comparing the shape conversion of spanin mutants to lysis by wild-type λ. (i) Lambda lysogens before lysis. (ii) Hole formation in the IM by the holin and PG degradation is shown (red arrow). (iii) Lysis is complete after OM disruption. The putative fusion of the IM and OM is shown. The green arrow shows expected path of travel for intracellular content, including phage progeny. (iia) In the case of the spanin mutant, PG degradation continues (indicated by red arrow) without OM disruption. (iib) shape loss occurs in zone(s) where PG is absent. (iic) The terminal phenotype of spanin mutants is an un-lysed spherical cell.

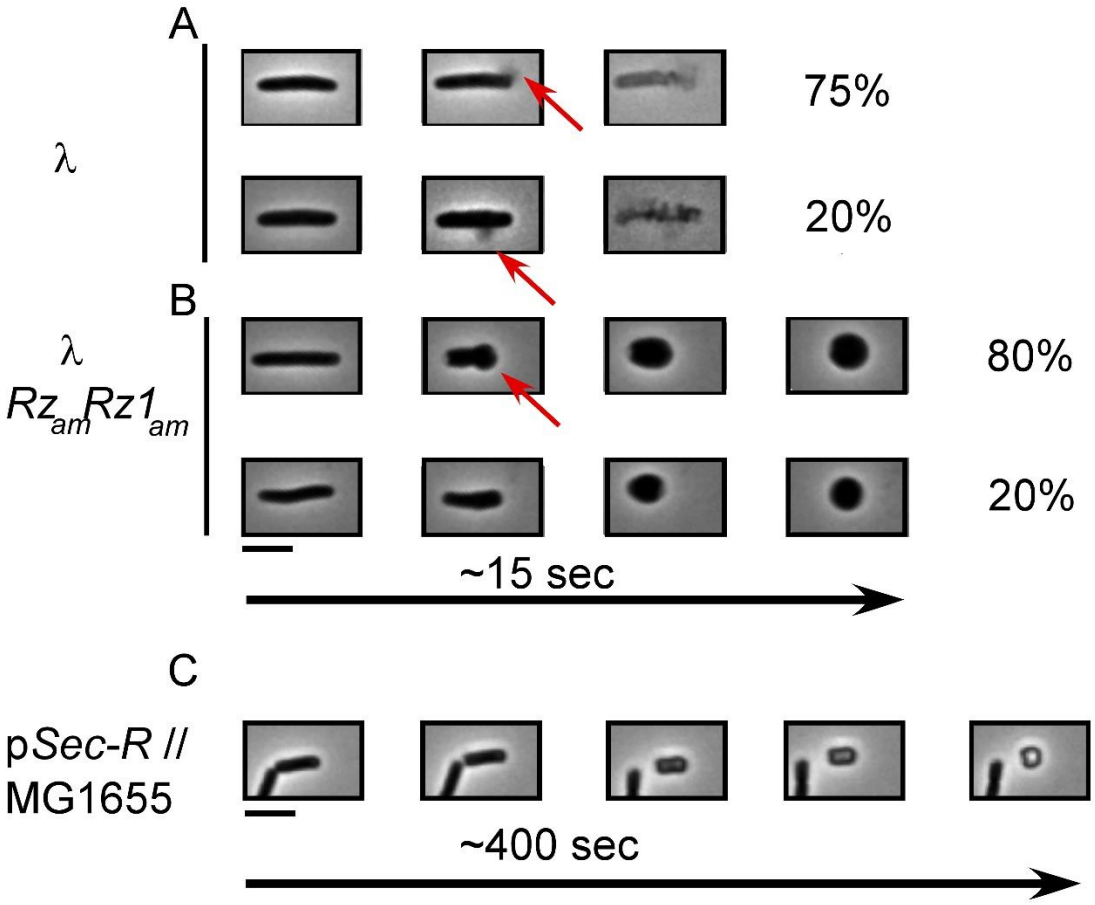


Figure D.2. Lysis morphology and shape conversion of λ lysogens.

MG1655 lysogens and were thermally induced and monitored by phase contrast microscopy 1 min prior to lysis. Average time from the first frame prior to an observable morphological change to lysis or the completion of rod-to-sphere shape conversion is displayed below micrographs. Micrographs are representative of the lysis morphology or shape conversion observed. Percentages are given to the right of micrographs indicate how frequent the displayed type of morphological change occurs. (A) λ900 lysogens showed two primary sites of rupture (indicated with red arrows): in the side wall and in the polar region. The remaining 5% in panel A are instances when the site of rupture was unclear. (B) Similar to the lysogens in panel A, λ900*Rz*$_{am}$*Rz1*$_{am}$ lysogens converted shape either at the poles or by apolar inflation. (C) Cells expressing p*Sec-R* were induced with 0.4% arabinose and monitored as described above. n-values are 134, 40, and 62 for A, B, and C, respectively. Scale bar = 5 μm.

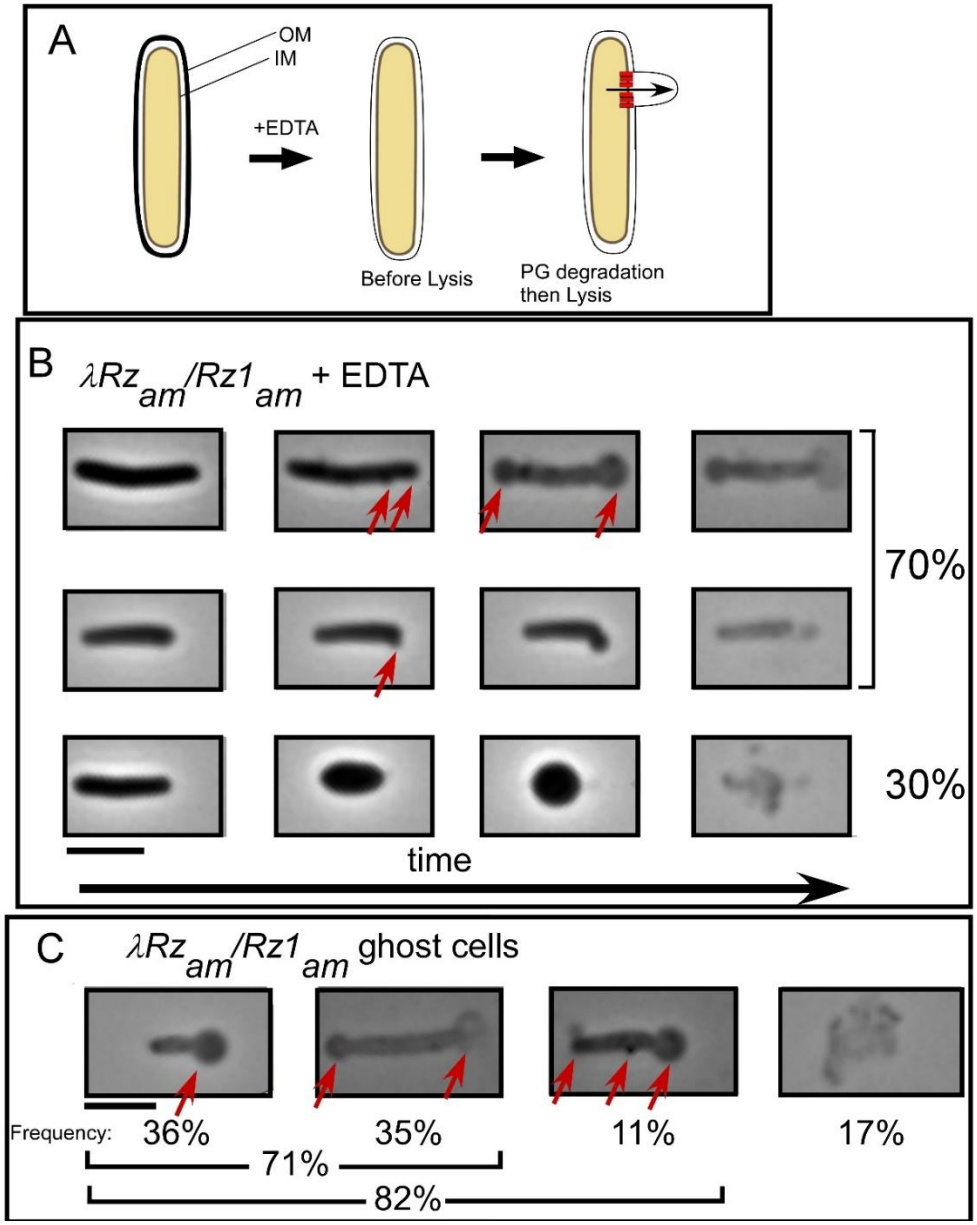


Figure D.3. EDTA complements the lysis defect of spanin mutant cells.

(A) Cartoon of morphological changes expected upon EDTA treatment of spanin mutant lysogens. (B) Time-lapse series of λ900Rz_{am}/$Rz1_{am}$ lysogens treated with EDTA 5 min prior to lysis. Micrographs are representative of the lysis morphology or shape conversion observed. (C) Ghost cells of spanin mutant lysogens treated with EDTA. Red arrows indicate blebs or blowouts. Percentages are given to the right of micrographs indicate how frequent the displayed type of morphological change occurs. The four major representative classes of ghost are shown. n = 122. Scale bar = 5 μm.

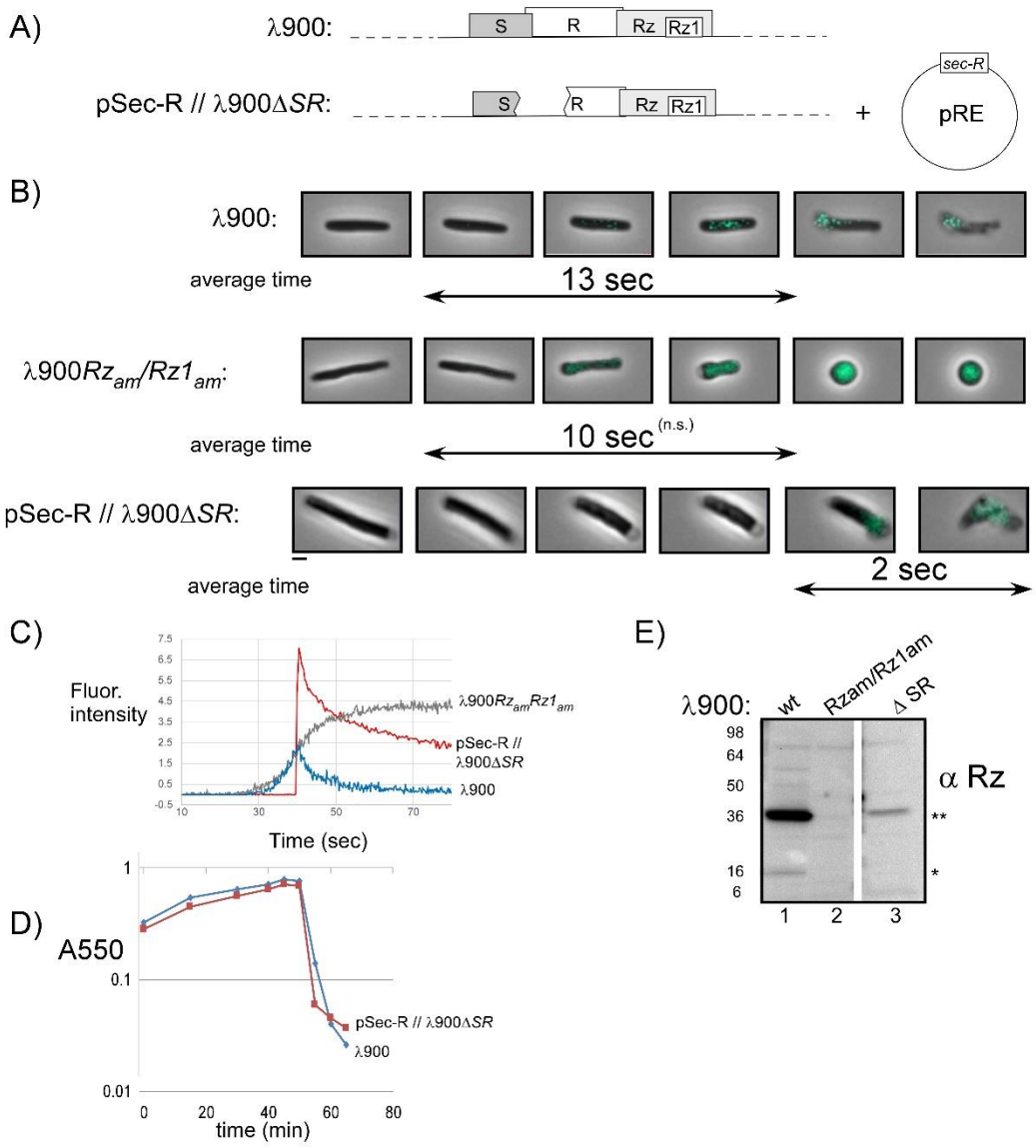


Figure D.4. Thioflavin-T indicates inner membrane permeabilization.

(A) The genotypes used in the panels below. (B) Thioflavin-T (ThT) was added to lysogens at the time of thermal induction. Cells were monitored 1 min prior to lysis with phase-contrast and CFP time-lapse microscopy. Micrographs are a merger of phase and CFP channels and show individual cells progressing through lysis or morphological conversion. Average time from detection ThT signal above background to lysis or shape change is displayed below micrographs. Scale bar = 2 μm. (C) Lysis profile demonstrating the lytic phenotype of λ900Δ*SR* lysogens complemented by *Sec-R*

expression. Lysogens were induced at time 0 and monitored at A550. (D) Western blots comparing Rz expressed from λ900 (wt), λ900$Rz_{am}/Rz1_{am}$, and λ900Δ*SR* lysogens. The position of the Rz monomer and covalently linked Rz dimer are indicated by single and double asterisks, respectively. (E) Maximum fluorescence pixel intensity within cells plotted over time. Signals were normalized. Arbitrary units are shown for the y-axis.

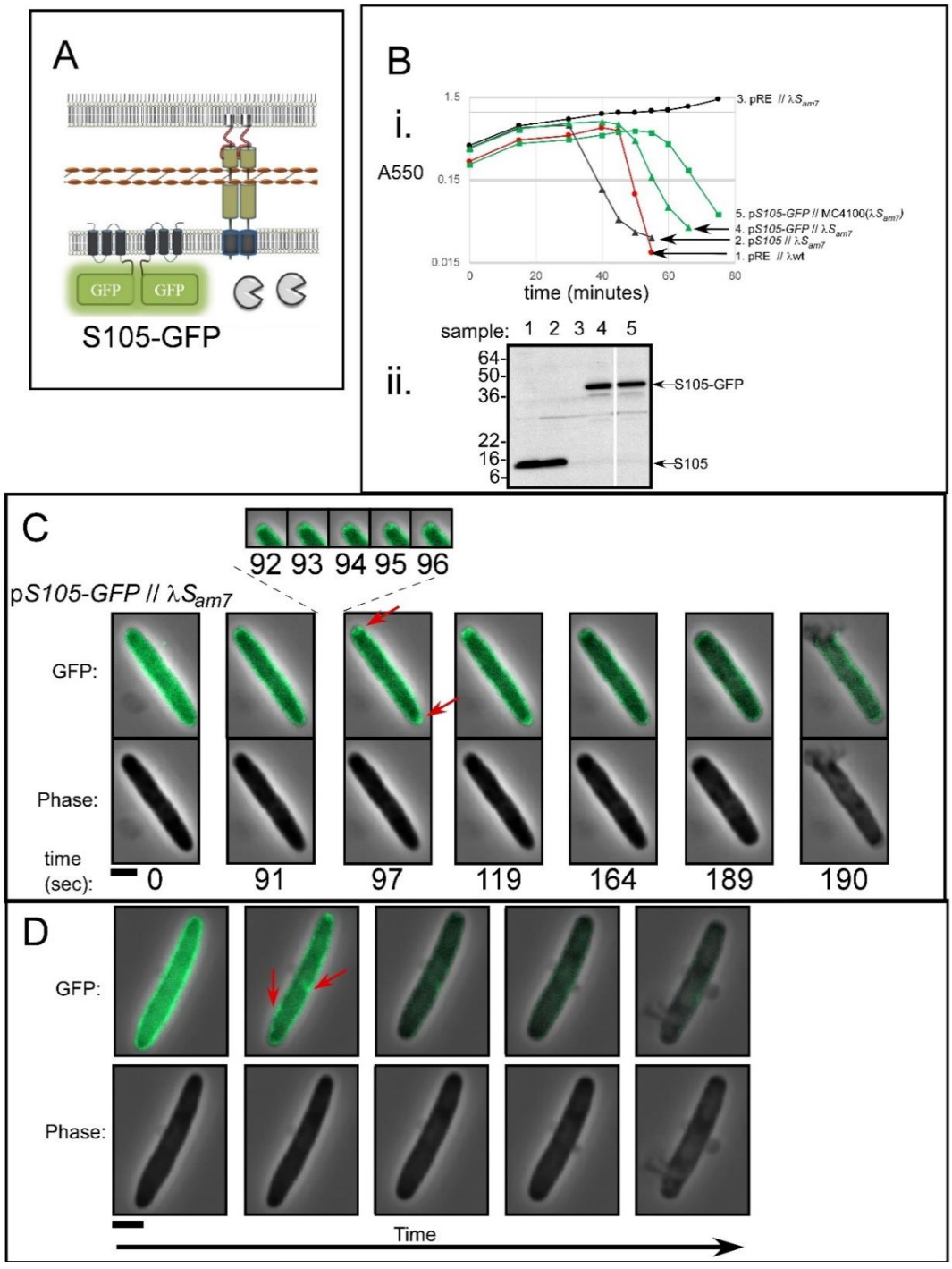


Figure D.5. Lysis of λ900S$_{am7}$ lysogens expressing p*S105-GFP*.

(A) Cartoon showing the localization of lysis proteins produced by λ900S$_{am7}$ plasmid-borne p*S105-GFP*. (B) (i) Lysis profile showing complementation of *S$_{am7}$* by plasmid-

expressed *S105-GFP*. λ900S$_{am7}$ lysogens expressing pRE plasmids as indicated. All cells are MG1655 unless indicated otherwise. (ii) Western blot comparing the *S105* expressed from the chromosome (lane 1), to *S105-GFP* expressed from the pRE plasmid (lanes 2-5). (C, and D) Representative time lapse images of isogenic λ900S$_{am7}$ lysogens induced for *S105-GFP*. Cells were imaged 1 min before lysis. Time is shown below panels. Red arrows indicate rafts. Scale bar = 2 μm

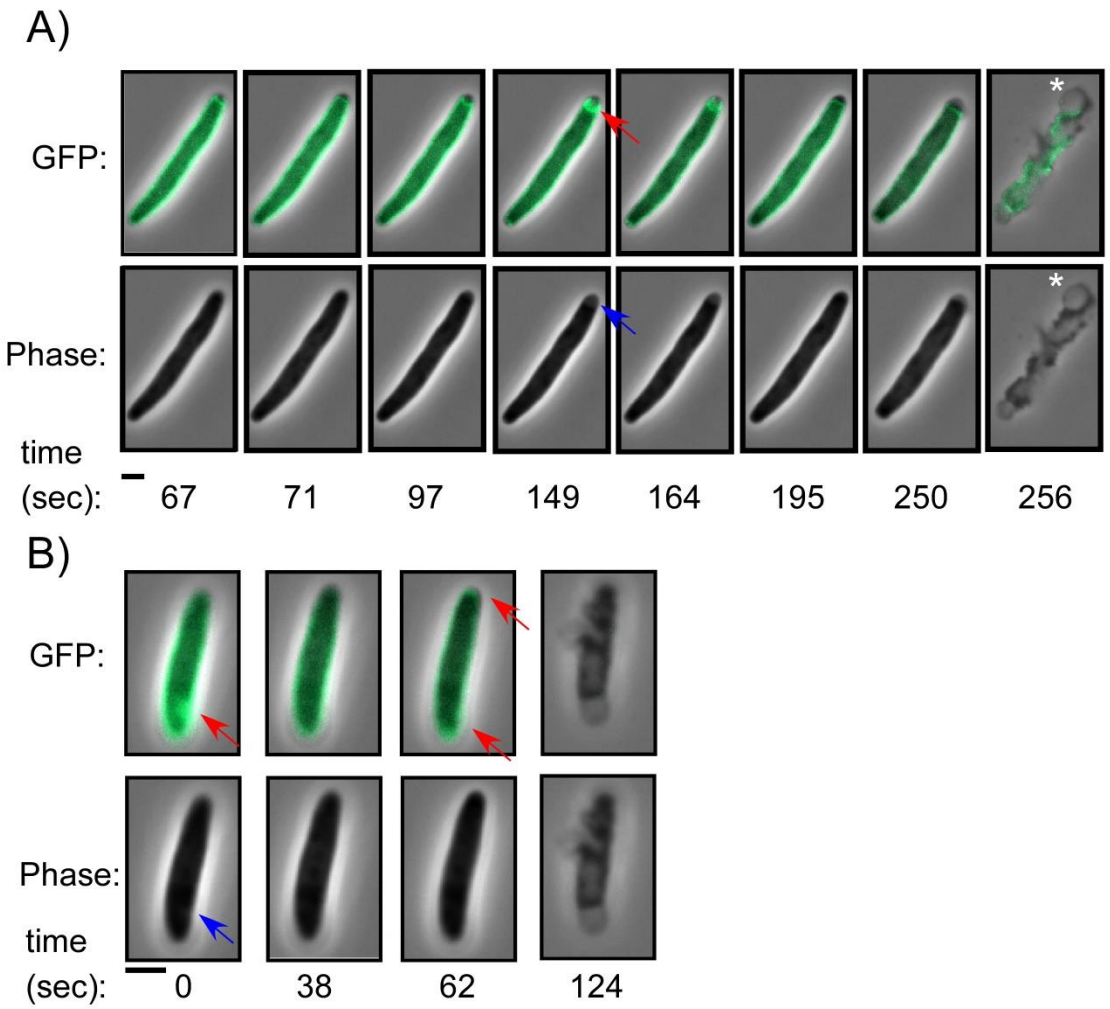


Figure D.6. The dynamics and features of S105-GFP rafts.

λ900S$_{am7}$ lysogens induced for lysis and expressing *S105-GFP*, isogenic to Fig. D.5C and D. Two examples of intermittent holin raft formation (denoted "flickering" in the text) can be seen. Arrows indicate rafts (red) and phase light spots (blue). Vesicles generated after lysis are denoted by asterisks. Scale bar = 2 μm.

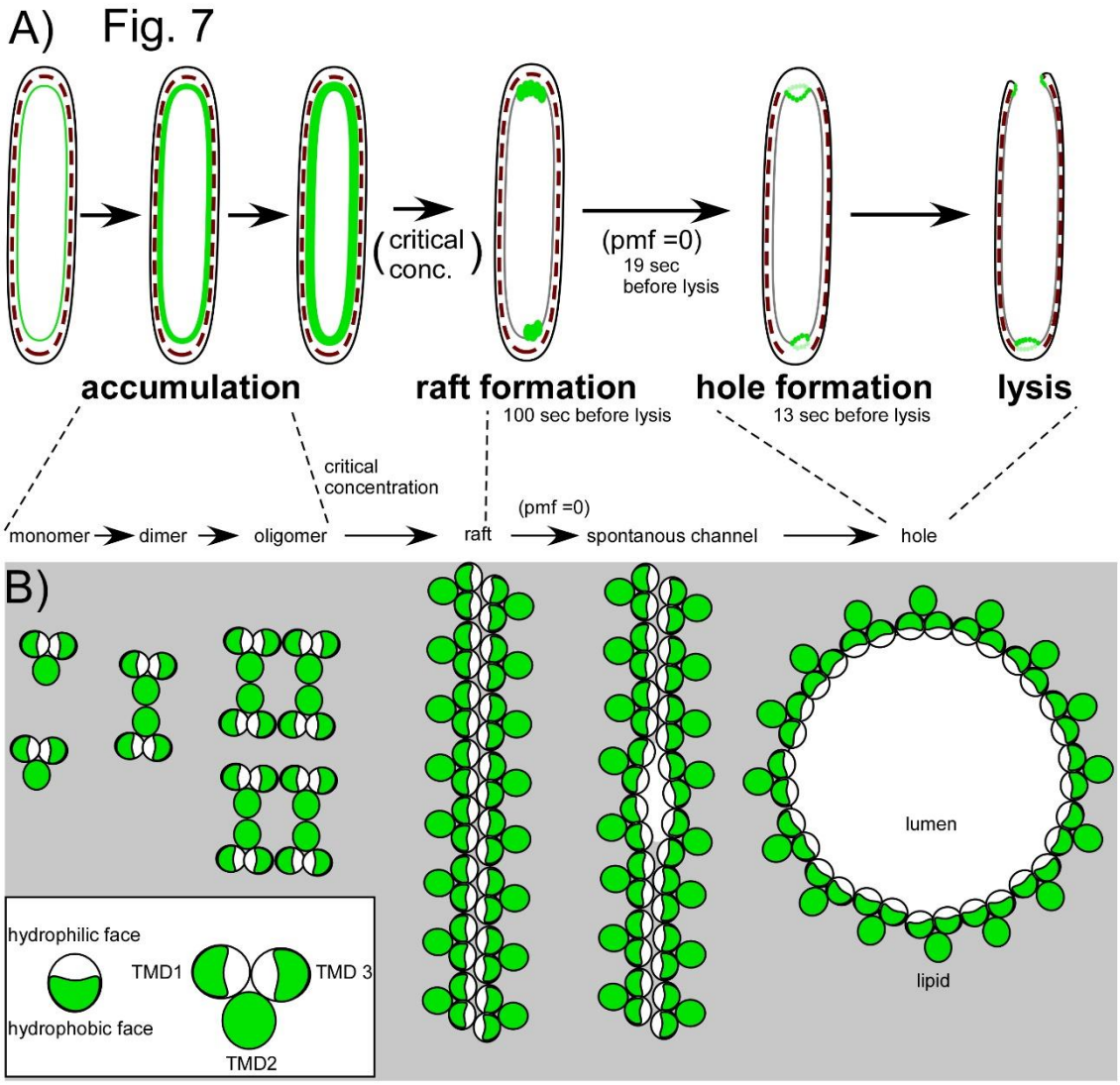


Figure D.7. Model of holin function incorporated into observed morphological changes prior to lysis.

(A) The holin is shown in green accumulating in the IM and nucleates into rafts upon reaching a critical concentration. Rafts are most often associated with the poles. Rafts trigger, forming holes in the IM which result in localized blowout proximal to the site of raft formation. (B) As the holin accumulates in the IM, the hydrophilic faces of TMD1 and 3 are in contact, sequestered from the lipid. In the raft form the hydrophilic faces are associated by intermolecular contacts. Rafts persist until the energy production of the cell is blocked. When pmf is lost, the raft rearranges into a hole-forming structure, which allows endolysin access to the periplasm. Lysis begins.

Table D.1. Phages, strains, plasmids, and synthetic genes used in this study.

Bacteriophages	**Genotypes and relevant features**	**Sources**
λ900	λ*Δ(stf tfa)::cat cI*$_{857}$ *bor::kan;* carries Cam^R and Kan^R	(125)
λ900*Rz*$_{am}$	λ*Δ (stf tfa)::cat cI857 RzQ100am bor::kan*	Laboratory stock
λ900*Rz*$_{am}$ */Rz1*$_{am}$	λ*Δ (stf tfa)::cat cI857 RzQ100am Rz1W38am bor::kan*	Laboratory stock
λ900Δ*SR*	λ*Δ (stf tfa)::cat* Δ*SR* deletion nt 45136–45815 *bor::kan*	Laboratory stock
λ*R*$_{am}$	λ*Δ (stf tfa)::cat cI857* mutations in *R*$_{Q26am}$ and *R*$_{W73am.}$	Laboratory stock
λ900*S*$_{am7}$	λ*Δ (stf tfa)::cat cI857 S*$_{am7}$ *bor::kan, a*mber in position 56 of *S*	Laboratory stock
Strains		
RY17303	MC4100Δ*tonA*	Laboratory stock
RY16390	MG1655Δ*tonA lacI*^{q} Δ*lacY*	Laboratory stock
MG1655 (λ900)	Lysogens carrying indicated λ prophage	This study
MG1655(λ900*Rz*$_{am}$/ *Rz1*$_{am}$)		This study
MG1655 (λ900Δ*SR)*		This study
MG1655 (λ*R*$_{am}$*)*		This study
MG1655 (λ900*S*$_{am7}$*)*		This study
MC4100 (λ900*S*$_{am7}$*)*		(37)

Plasmids		
pRE	Medium copy plasmid carrying Q-dependent λ pR'	(139)
pBAD24	*bla araC* P_{ara}	(142)
pSec-R	The *sec-R* gene is a fusion of the *phoA* signal sequence (residues 1-25) inserted in-frame at the 5' end of the entire *R* gene. The RBS of *R* was used. Under P_{ara} control and cloned into KpnI and HindIII sites of pBAD24.	This study
pS105-GFP	pRE carrying GFP_{A206K} inserted in frame with *S105*, separated by a 30 aa linker.	(37)
Synthetic gene fragments *Sec-R*	TTAGGTACCGCCGGAGTAGAAGATGAAACAAAGCACTATTGCA CTGGCACTCTTACCGTTACTGTTTACCCCTGTGACAAAAGCCC GCACACCAGAAATGGTAGAAATCAATAATCAACGTAAGGCGTT CCTCGATATGCTGGCGTGGTCGGAGGGAACTGATAACGGACGT CAGAAAACCAGAAATCATGGTTATGACGTCATTGTAGGCGGAG AGCTATTTACTGATTACTCCGATCACCCTCGCAAACTTGTCAC GCTAAACCCAAAACTCAAATCAACAGGCGCCGGACGCTACCAG CTTCTTTCCCGTTGGTGGGATGCCTACCGCAAGCAGCTTGGCC TGAAAGACTTCTCTCCGAAAAGTCAGGACGCTGTGGCATTGCA GCAGATTAAGGAGCGTGGCGCTTTACCTATGATTGATCGTGGT GATATCCGTCAGGCAATCGACCGTTGCAGCAATATCTGGGCTT CACTGCCGGGCGCTGGTTATGGTCAGTTCGAGCATAAGGCTGA CAGCCTGATTGCAAAATTCAAAGAAGCGGGCGGAACGGTCAGA GAGATTGATGTATGAGGATCCAAGGGCCTAAGCTTGGCTGTTT TGGCG	This study

Table D.2. Lysis of spanin mutant cells treated with EDTA prior to lysis.

The number of blebs detected per cell, subcellular localization of blebs and lysis morphology are described per cell. P = polar; pp = parapolar; m = midcell. "-" indicates blebs were not detected.

Cell	Number of blebs	Subcellular location	Lysis morphology
1	1	P	Rod
2	5	P, P, P, pp, m	Sphere
3	1	P	Rod
4	2	P, P, m	Rod
5	1	P	Rod
6	1	P	Rod
7	3	P, P, m	Rod
8	3	P, P, P	Rod
9	2	P, P	Rod
10	3	P, P, P	Rod
11	2	P, P	Rod
12	3	P,P, pp	Rod
13	4	P, P, pp, pp,	Rod
14	3	P, P, pp	Rod
15	2	P, P,	Rod
16	-	-	Sphere
17	3	P, P, m	Rod
18	2	P, P,	Rod

19	3	P, P, m	Rod
20	-	-	Sphere
21	2	P, P,	Rod
22	2	P, P,	Rod
23	1	P	Rod
24	-	-	Sphere
25	-	-	Sphere
26	3	P, pp, m	Rod
27	3	P, P, P	Rod
28	2	P, P,	Rod
29	-	-	Sphere
30	-	-	Sphere
31	3	P, P, pp	Rod
32	2	P, pp	Rod
33	-	-	Sphere
34	4	P, P, pp, m	Rod
35	-	-	Sphere
36	1	P	Rod
37	-	-	Sphere
38	3	P, P, m	Rod
39	-	-	Sphere

40	3	P, P, pp	Rod
41	3	P, P, pp	Rod
42	2	P, P	Rod
43	-	-	Sphere
44	-	-	Sphere

Table D.3. Lysis of cells expressing *S105-GFP*.

The table is ordered by the monitoring time before lysis. Time is reported in seconds. Breach site and raft locations are indicated by P=polar, m= midcell, pp =parapolar. NA. = uncertain because the parameter could not be detected. The "time raft to lysis" measures the seconds between raft formation and lysis. "time raft lost" is a measurement of the interval between when raft detection was lost and lysis.

cell	time to lyse	lysis site	# of rafts	raft location	time raft appears	time raft gone	time raft to lysis	Time raft lost	phase light raft	raft flicker	rafts predict lysis site
1	1	P	NA	NA	NA	NA	NA	NA	NA	NA	NA
2	1	m	NA	NA	NA	NA	NA	NA	NA	NA	NA
3	1	m	NA	NA	NA	NA	NA	NA	NA	NA	NA
4	41	m	NA	NA	NA	NA	NA	NA	NA	NA	NA
5	57	P	1	P	0	19	NA	38	-	-	+
6	80	P	2	P, P	0	21	NA	59	-	-	+
7	85	pp	4	P, P, pp, m	0	71	NA	14	-	-	+
8	86	m	1	P	0	43	NA	43	-	-	-
9	121	m	2	P, P	57	100	64	21	+	-	-
10	126	P	2	P, P	17	59	109	67	+	+	+
11	137	P	2	P, pp	0	NA	NA	NA	+	-	+
12	141	P	3	P, P, m	64	122	77	19	+	-	+
13	150	m	2	P, P	20.5	139	129.5	11	+	-	-
14	156	P	2	P, P	64	145	92	11	+	-	+
15	161	m	2	P, m	52	145	109	16	+	-	+
16	162	P	3	P, pp, m	0	NA	NA	NA	+	-	+
17	166	P	2	P, pp	69	94	97	72	+	-	+
18	170	m	2	P, P	72	132	98	38	-	-	-
19	170	P	1	P	0	157	NA	13	+	-	+
20	179	pp	3	P, P, pp	77	137	102	42	-	+	-
21	179	P	5	P, P, P, pp, pp	104	165	75	14	+	-	+
22	185	P	3	P, P, pp	93	167	92	18	-	-	-
23	186	P	2	P, P	68	118	118	68	+	-	+

24	191	P	2	P, P	95	158	96	33	+	+	+
25	195	pp	2	P, P	130	165	65	30	+	-	-
26	196	P	2	P, P	116	166	80	30	+	-	+
27	203	P	2	P, pp	82	150	121	53	-	-	+
28	205	m	2	P, P	96	136	109	69	+	-	-
29	211	m	1	P	133	190	78	21	+	-	-
30	214	P	2	P, P	66	177	148	37	+	+	+
31	215	m	5	P, P, P, pp, m	119	NA	96	NA	-	-	+
32	219	pp	2	pp, pp	140	208	79	11	+	-	-
33	221	pp	4	P, P, pp, pp,	134	175	87	46	+	-	+
34	222	pp	4	P, pp, pp, m	152	208	70	14	-	-	+
35	223	P	1	P	124	189	99	34	+	-	+
36	224	P	2	pp, m	153	206	71	18	+	-	-
37	225	pp	4	P, P, pp, pp	135	175	90	50	+	-	+
38	227	m	4	P, P, m, m	143	NA	84	NA	+	-	+
39	252	P	2	P, P	108	164	144	88	+	-	+
40	252	pp	2	P, P	166	225	86	27	+	+	-
41	259	P	2	P, P	167	NA	92	NA	+	-	+
42	265	pp	1	P	186	229	79	36	+	-	-
43	280	P	2	P, P	178	263	102	17	+	-	+
44	282	P	1	P	193	230	89	52	-	-	+
45	292	PP	4	P, P, pp, m	215	252	77	40	+	-	-
46	303	P	3	P, P, m	222	271	81	32	-	-	+
47	374	P	2	P, P	289	329	85	45	-	-	+
48	392	P	1	P	246	325	146	67	+	-	-
49	472	P	1	P	401	444	71	28	+	+	+
50	478	PP	3	P, P, m	422	455	56	23	+	-	-

Average	2.3	Average	93.4	35.7	33/46	6/46	30/46
Standard Deviation	1.1	Standard Deviation	22.2	20.0			65%
	n=46		n=39	n=41			

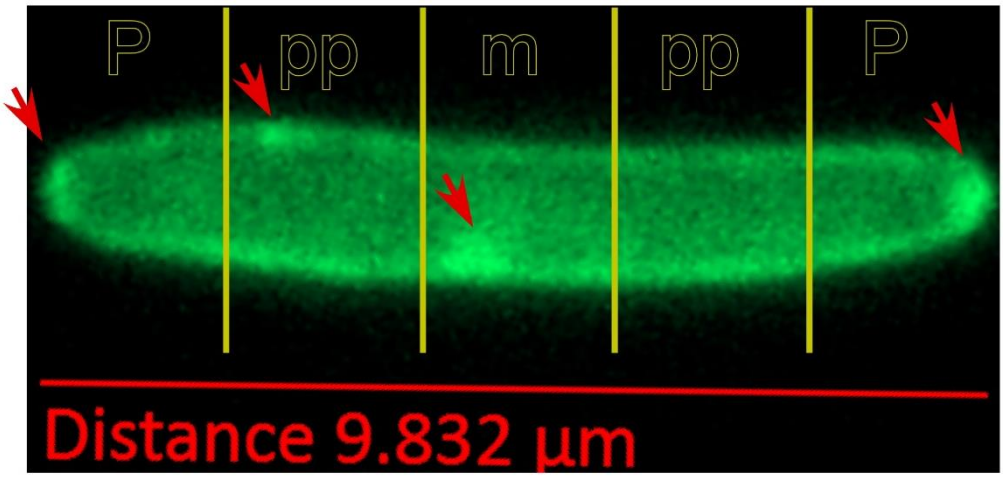


Figure DS.1.

Assigning rafts to subcellular compartments. P=polar, pp=parapolar, m=midcell.

APPENDIX E

FIGURES AND TABLES

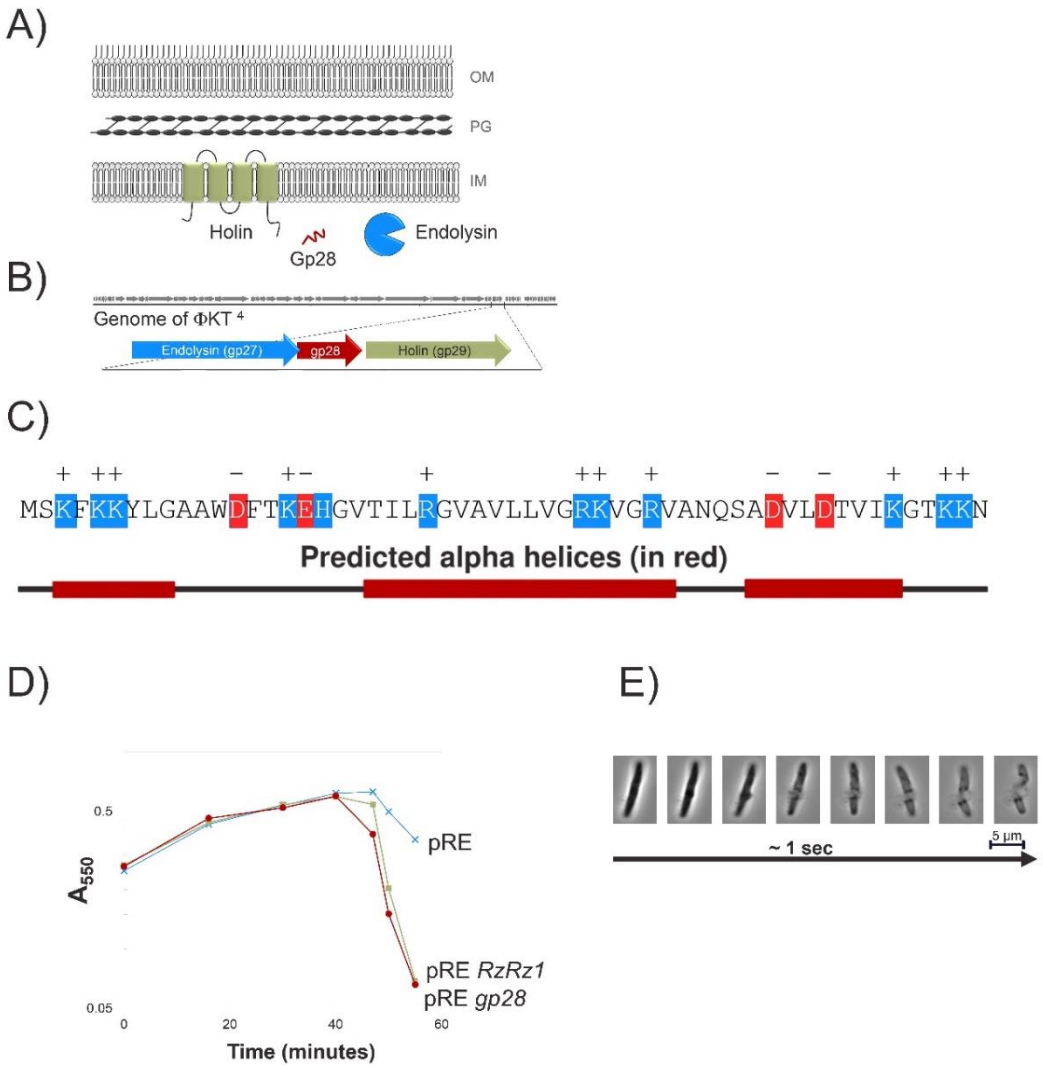


Figure E.1. Gp28 is a novel OM disrupting protein.

(A) Predicted localization and topology of the phiKT lysis gene products. (B) PhiKT lysis cassette is shown. (C) Amino acid sequence of Gp28 is shown with charges above residues. Blue= basic residues. Red = acidic residues. Predicted alpha helical structure is aligned below the aa sequence. (D) Lysis profile of $\lambda Rz_{am}Rz1_{am}$ lysogens expressing pRE vector only (blue), pRE RzRz1 (green) or pRE gp28 (red). Lysogens were induced

at t=0 and A550 was monitored over time. (E) The lysis cassette in B was plasmid-expressed and cells were monitored by phase-contrast at the time of lysis. Representative micrographs of a single cell lysing.

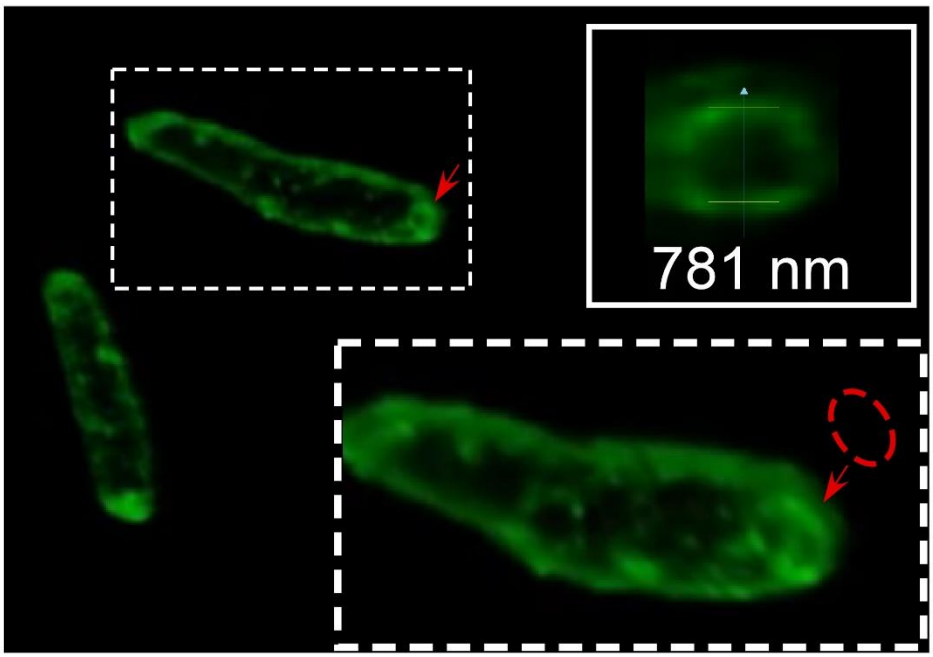


Figure E.2. 3D image of the S105-GFP ring

After triggering, z-stacks were taken and deconvolved using a confocal microscope. Two cells are shown. The boxed cell in the top center is enlarged and the right pole is tilted toward the viewer. The ring structure is marked with a red arrow and dashed red ring. Top right inset: The ring viewed from the pole of the cell inward. Measurements of the ring are given.