# THE ROLE OF SPANINS IN PHAGE LYSIS

A Dissertation

by

MANOJ RAJAURE

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, | Hays Rye |
| | Paul Straight |
| | Michael Benedik |
| Head of Department, | Gregory D. Reinhart |

May 2015

Major Subject: Biochemistry

## ABSTRACT

The work described in this dissertation addresses how spanins function during phage lysis, focusing on the 2CS system of phage lambda, in which Rz and Rz1 are the prototype i-spanin (inner membrane subunit) and o-spanin (OM lipoprotein subunits**)** of the spanin complex.

Results obtained from genetic and biochemical approaches found that there must be at least one homodimerizing intermolecular disulfide bond near the heterotypic interface. The covalent linkage of spanin homodimers was found to be mediated by the host Dsb system, the first time that this system has been shown to catalyze the formation of intermolecular disulfide bonds. Genetic and biochemical studies with the paradigm phage T4 revealed striking differences in the requirements for position and disulfide bond formation in the PseT.3/PseT.2 2CS system, presumably reflecting the separated architecture of the T4 spanin genes.

Mutational analysis using an extensive collection of Rz and Rz1 mutants highlighted important domains and motifs of Rz and Rz1 that are required for function. A method has been developed to isolate intragenic and intergenic suppressors for these mutants, which should allow future studies to interrogate interacting interface between the different domains of spanin subunits.

In addition, it was found that the spanin complex, in the absence of holin function, can cause rapid lysis of a host mutant that is compromised for peptidoglycan biosynthesis, suggesting that spanin function is regulated by the meshwork of intact

peptidoglycan. Furthermore, we developed a spheroplast method to test the model proposed for spanin function: that the last step of lysis is fusion between the IM and OM. The method is based on the fusogenic properties of the spanin subunits externalized on the outer surface of the cytoplasmic membranes of spheroplasts. The unambiguous result was that the lambda 2CS subunits supported efficient, spheroplast fusion, and that spanin missense mutations that abrogate lysis *in vivo* also blocked the spheroplast fusion activity. These results suggest the last step in lysis of the Gram-negative host is OM disruption by fusion with the IM.

## DEDICATION

To my late grandfather Jageshwor Rajaure. Still lingers in my memory, his words of inspiration and encouragement.

## ACKNOWLEDGEMENTS

I would first like to thank my committee chair, Dr. Ry Young, for giving me the opportunity to work under his guidance. He has been a greatest mentor and a great teacher, I have ever known, to learn from. Without his expert ideas and guidance, I would not succeed in all the works that are presented in this dissertation. Thank you Ry!, for everything and your kind support throughout my graduate career. I would also like to thank my committee members, Dr. Hays Ry, Dr. Paul Straight, and Dr. Michael Benedik, for their guidance, helpful discussion, and constructive comments throughout the course of this research.

Thanks to all of the Young lab members (past and present) and other colleagues in the department who have always been helpful by listening and providing me their honest opinion regarding my work. I would like to especially thank Dr. Joel Berry whose mentorship and outstanding foundational work provided me the guidance for my thesis project. I would like to acknowledge his collaboration in a majority of work presented in the second chapter of this dissertation. I would like to thank my other "Spanin Team Members" and colleagues, Rohit Knongari and Jesse Cahill, for their helpful discussion, comments, and support. I would like to appreciate Dr. Gabriel F. Kuty, Karthik Chamakura, Yi Chen, Adriana Hernandez, and Anthony Sperber for their collegial support, as well as for all the excitement they brought to the lab. I truly want to thank all of the administrative staffs for their support, especially Mrs. Daisy Wilbert for her clerical assistance. She had been phenomenal in helping me to organize meetings,

process documents, and provide assistance with other administrative work. It was such a great experience being in the Department of Biochemistry at Texas A&M University.

Finally, thanks to my parents and my wife for their encouragement, patience, and love.

**TABLE OF CONTENTS**

Page

CHAPTER Page

## LIST OF FIGURES

## LIST OF TABLES

TABLE Page

# CHAPTER I

# INTRODUCTION

Bacteriophages (phages) are viruses that infect bacteria. They are the most abundant and diverse biological entity in the biosphere (1). Current estimates suggest there are ~$10^{31}$ phages in the biosphere and more than $10^{23}$ phage infections occur every second in the ocean alone (2). To initiate an infection, a phage adsorbs onto the surface of a susceptible host and injects its DNA into the host cytoplasm. This is followed by maturation and assembly of progeny virions. At the end of the infection cycle, progeny virions are released, either by extrusion or lysis. Once nascent phages are released from the host, they will infect sensitive hosts in the surrounding environment. It has been suggested that phage infection has a major impact on the ecological balance and dynamics of the microbial population (3). In addition, phages play a key role in bacterial evolution through horizontal gene transfer, including the transfer of virulence genes (4).

The study of bacteriophages has led to the development of modern molecular biology. Many fundamental principles of life science were derived from phage-related work. Such fundamentals include identifying DNA as the basic hereditary element, elucidating processes of DNA replication and transcription, the discovery of the genetic code, the discovery of restriction enzymes, the assembly of viral capsids, and the gene regulation. Phage lambda (λ) is one of those bacteriophages that served as the paradigm for numerous studies, including phage lysis.

λ is a double-stranded DNA (dsDNA) phage of *Escherichia coli*, belonging to the order of *Caudovirales* (or tailed-phages), the largest group in the biosphere. During infection, λ injects its DNA into the host. Injected DNA must access the cytosol, which involves penetration of the host cell envelope. After DNA injection, a cascade of molecular events occurs inside the cytosol leading to phage replication and progeny virion assembly. The infection cycle is terminated by the catastrophic event of cell lysis, which releases newly-assembled virions. Virion release must overcome the three barriers of the cell envelope: the outer membrane (OM), the peptidoglycan (PG), and the inner membrane (IM).

dsDNA phages like λ with a genome size of >10kb employ three types of proteins to lyse Gram-negative hosts. These proteins are holins, endolysins, and spanins, which actively overcome the IM, PG, and OM barriers, respectively. Holins are small hydrophobic membrane proteins that accumulate in the IM and are eventually triggered to form a large hole. Hole formation in the IM permits endolysin to escape into the periplasm and degrade the PG. In the final step, spanins disrupt the OM to complete lysis. In the absence of lysis proteins, virions are trapped inside the host and lysis is delayed indefinitely.

The holin and endolysin of λ are by far the best-characterized lysis proteins and were classically thought to be sufficient for lysis. Until recently, the necessity of spanins was overlooked because they were thought to be only required for lysis when the stability of the OM is enhanced by divalent cations (5, 6). Recently, rigorous characterization of spanins has yielded valuable information, suggesting that spanins

also play a necessary role in lysis (7-10). Based on these studies, a model for spanin function was suggested: spanins cause fusion of the IM and OM in the last step of phage lysis. This mode of function is analogous to the exit strategy of enveloped viruses or the trafficking of secretory vesicles in eukaryotic cells (11, 12). This dissertation aims to describe the spanin proteins of λ to understand the phenomenon of membrane fusion that happens during phage lysis.

In the following section, a brief overview of the cell envelope will be provided, followed by a detailed review on spanins and their role in lysis. Later in this chapter, a brief review of membrane fusion mechanisms in general will be discussed in order to put the membrane fusion model of spanin function in perspective.

**Cell envelope: barriers to phage release**

The cell envelope is a physical barrier to protect the cell from the outside environment. In addition, the cell envelope imposes a critical impediment to phage release. Bacteria have been classified into two major groups based on the composition of the cell envelope: Gram-positive and Gram-negative. The cell envelope of Gram-negative bacteria is comprised of an IM, a PG, and an OM (Fig. 1.1). Gram-positive bacteria lack an OM and possess a thicker PG.

The IM is a hydrophobic phospholipid bilayer that surrounds the cytoplasm. The phospholipid bilayer is mainly composed of phosphatidyl ethanolamine and phosphatidyl glycerol, with small fraction of phosphatidyl serine and cardiolipin (13,

14). It acts as a semi-permeable barrier, preventing leakage of hydrophilic constituents from the cytoplasm.

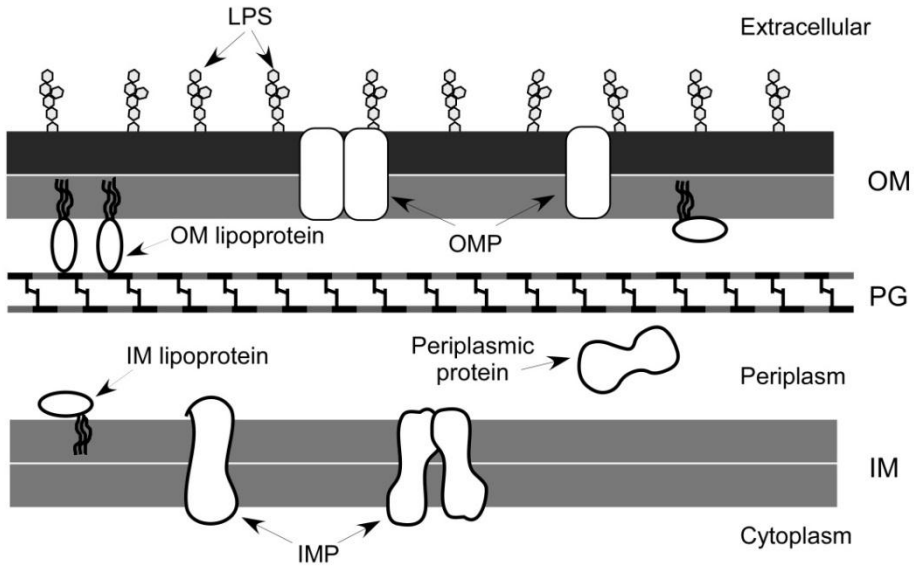



Figure 1.1. The *E. coli* cell envelope. Three major components: outer membrane (OM), peptidoglycan (PG), and inner membrane (IM) are labeled on the right. Other associated components, lipopolysaccharides (LPS), outer membrane proteins (OMP), inner membrane proteins (IMP), inner and outer membrane lipoproteins, and periplasmic proteins are indicated by black arrow. See text for more details.

The IM of the bacterial cell is also the center of many important biological functions such as energy production, lipid biosynthesis, protein secretion, environmental sensing, and assembly of biological precursors (15). These activities are maintained by several

factors, including numerous integral proteins and lipoproteins embedded in the phospholipid bilayer.

Lipoproteins are post-translationally modified proteins (Fig. 1.2). Prior to their modification, these proteins are synthesized in an immature form and translocated to the IM by the Sec machinery. These proteins have a signal sequence at the amino-terminal end characterized by the presence of "lipobox" that has a cysteine residue. The lipobox is a signal peptidase II processing site (L/V-X-G/A/S-C), where X is an amino acid that is not charged or a helix-breaker (16). This signal sequence is first modified by the formation of a thioether diglyceride bond at the cysteine residue catalyzed by the diacylglyceryl transferase, Lgt. The modified signal sequence is cleaved at the amino-terminal end of the modified cysteine residue by a signal peptidase II (SPII). After the removal of the signal sequence, an additional acyl chain is attached to the newly formed amino-terminal end by phospholipid transacylase, Lnt. This mature form is called a lipoprotein (17). IM lipoproteins are anchored to the outer leaflet of the IM by their lipoylated moieties. IM lipoproteins differ from other integral membrane proteins which are composed of transmembrane domains (TMDs), $\alpha$-helical structures that traverse the membrane bilayer.

Another important key feature of the IM is the maintenance of an electrochemical gradient across the membrane bilayer to generate a proton-motive force (pmf) which is crucial for energy production. It is important to mention that once this energy state of the IM is depleted, the cell becomes physiologically inactive and many, if not all, biological functions cease.

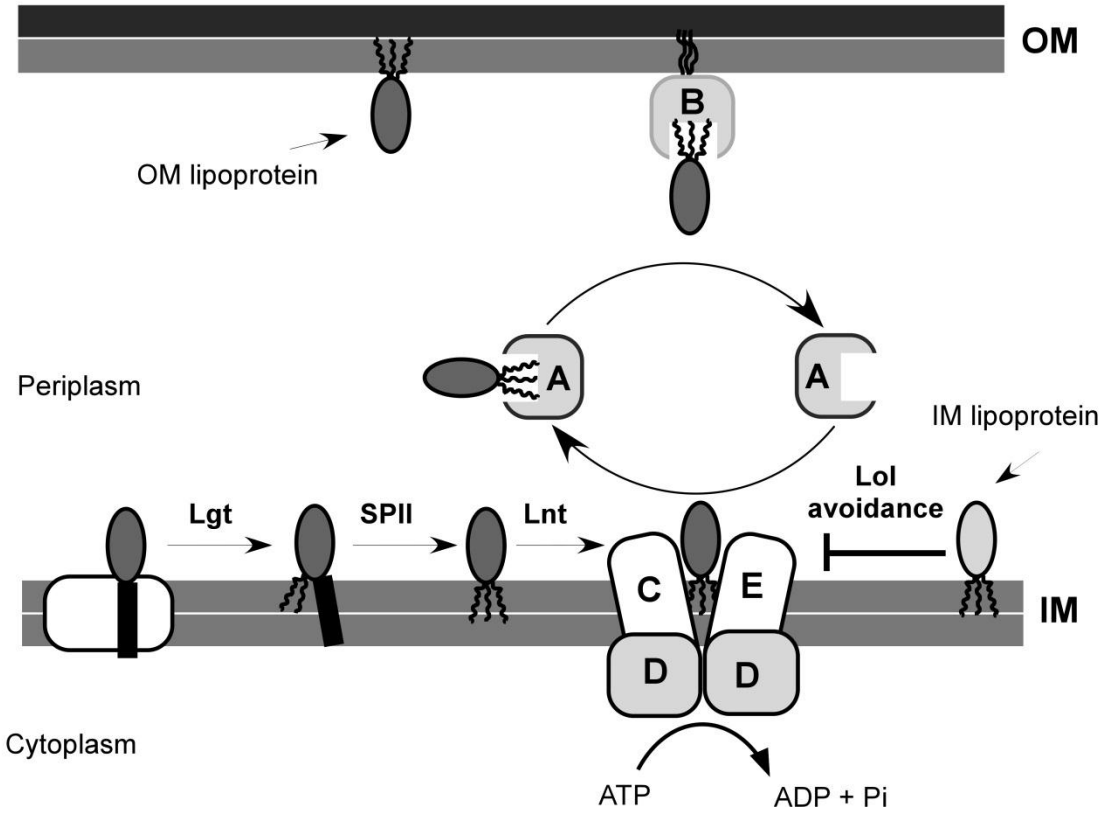



Figure 1.2. The Lol system. Lipoprotein precursors are first translocated across the IM by general secretory pathway. The enzymes Lgt, SPII, and Lnt modify the prolipoprotein into a mature lipoprotein by attaching three acyl chains (see text for details). Mature lipoproteins are then sorted to the IM and OM by the Lol system. The ABC transporter LolCDE complex provides an outer membrane-specific lipoprotein to the periplasmic chaperone LolA, which delivers the lipoprotein to the OM lipoprotein receptor LolB. LolB releases an OM lipoprotein into the inner leaflet of the OM. An Asp residue at the +2 position of mature lipoproteins will lead to Lol avoidance; in this case lipoproteins are not recognized by the LolCDE complex and will be retained in the IM (18, 19).

In contrast to the IM, the OM is an asymmetric bilayer serving as a first line of defense for bacteria. Asymmetry is due to the presence of lipopolysaccharide (LPS) in the outer leaflet of the OM. One important feature of the LPS is that it is known to initiate an endotoxic reaction during bacterial infection (20). Another important feature is that the negatively charged LPS helps to stabilize the OM in the presence of divalent cations such as $Ca^{2+}$ or $Mg^{2+}$ (21). Therefore, it has been suggested that the stability of the OM is dependent, to an extent, on the LPS (22, 23). The treatment of *E. coli* and other Gram-negative bacteria with a cation-chelating agent such as ethylenediaminetetraacetic acid (EDTA) results in sensitivity to hydrophobic antibiotics and lysozyme (21). This is due to the diminished stability of the OM in the absence of divalent cations. The maintenance of the stability of the OM makes Gram-negative bacteria resistant to the detergents, antibiotics, and other agents (24).

Unlike the IM, most of the proteins present in the OM are of β-barrel structure. Some of these proteins are porins and efflux pumps that are used for active and/or passive diffusion of nutrients and molecules. Another important characteristic of the OM is the abundance of OM lipoproteins (25). These OM lipoproteins are post-translationally modified in a similar manner to that of the IM lipoproteins described earlier, except that they are transported to the inner leaflet of the OM by the localization of lipoprotein (Lol) system (Fig. 1.2). This is because OM lipoproteins do not possess the negatively charged aspartic residues at the +2 and +3 position of the lipolyated cysteine residue, which would otherwise retain the lipoproteins in the IM (18, 19). After the maturation of the OM lipoprotein in the outer-leaflet of the IM, these proteins are

recognized by a LolCDE complex and delivered to a periplasmic chaperone called LolA using ATP as an energy source. The LolA chaperone ultimately transports the OM lipoprotein to the inner-leaflet of the OM by delivering it to a receptor called LolB. The OM lipoproteins also make a covalent linkage to the PG and it has been suggested that these linkages maintain the integrity of the OM (26). In *E. coli*, the OM is located approximately 25 nm above the IM bilayer (9, 27).The aqueous environment in between is called the periplasm (28). Many soluble proteins are present in this compartment, some of which are enzymes such as nucleases and proteases. Other proteins in the periplasm are chaperones or enzymes that function in cell envelope biogenesis (29).

The third component of the cell envelope is the PG located in the periplasm (Fig. 1.1). The PG, which is composed of the cross-linked glycan strands, is an essential component of the cell envelope. The repeating units of the disaccharide N-acetylglucosamine:N-acetylmuramoyl-(pentapeptide) make up the glycan strands. The glycan strands are cross-linked through the pentapeptide cross bridges by mostly 4-3 and to some extent by 3-3 linkages (30). The PG determines the rod-shaped structure of the cell and confers mechanical strength to resist internal turgor pressure (31). In addition, electron micrographs have indicated that the PG structure in Gram-negative bacteria is thinner compared to Gram-positive bacteria (32).

**The λ lysis paradigm**

Phage-mediated lysis of a Gram-negative host is best studied in the paradigm phage, λ. λ is a siphophage with dsDNA that infects *E. coli*, a Gram-negative bacterium. λ lysis is a three-step process (33). These three steps involve targeting the three barriers of the host cell envelope. The genes responsible for lysis are tightly regulated and are induced only when the phage enters the lytic cycle. In λ, these genes are clustered into a lysis cassette and comprises of four genes: *S*, *R*, *Rz*, and *Rz1*. The λ lysis cassette is part of a single operon which is under the control of the late promoter pR'(Fig. 1.3A). Transcription from this promoter is dependent upon the anti-terminator protein, Q. There are five proteins coded by this lysis cassette. The *S* gene has a dual-start motif, meaning it has two start codons resulting in two translational products, S105 and S107, named after the total number of respective amino acid residues. These proteins have an opposing function: S105 is the holin and S107 is the anti-holin (34, 35). At physiological levels, the production ratio of S105 to S107 molecules *in vivo* is 2:1. This ratio is dictated by the secondary structure of the *S* transcript (36). The *R* gene encodes a soluble endolysin which has transglycosylase activity (37). The two genes at the distal end of lysis cassette have an unusual architecture. The smaller gene, *Rz1*, is entirely embedded in a larger gene, *Rz*. The translational frame of Rz1 is at +1 reading frame of Rz.

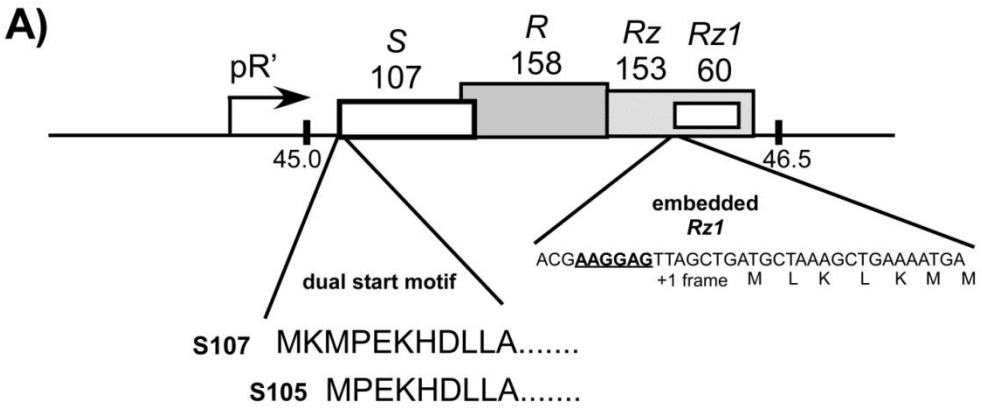



Figure 1.3. The λ lysis cassette, the lysis model, and the lysis morphology
**(A)** The λ lysis cassette. The λ genomic fragment between 45kb and 46.5 kb with lysis genes holin (*S*), endolysin (*R*), spanins (*Rz* and embedded *Rz1*) are drawn relative to the scale. The late promoter pR' is depicted by forward arrow at the beginning of the lysis cassette. Inset: the dual start motif of holin *S105* and *S107* depicted on the left; the embedded open reading frame of *Rz1* is shown on the right. **(B)** The λ lysis model. See the λ lysis paradigm section in the text for more details. **(C)** A series of phase contrast micrographs of representative cells were captured ~50 min after thermal induction of λ prophage. Time 0 indicates the beginning of the lysis event. Upper panel; the presence of spanins (*Rz* and *Rz1*) cause lysis, which is characterized as a local blowout starting from a pole. Lower panel; absence of *Rz* blocks lysis, with local deformation beginning from a pole. Time in milliseconds is indicated at the top of the micrographs and the scale bars represent 5μm. (also see review, (33)).

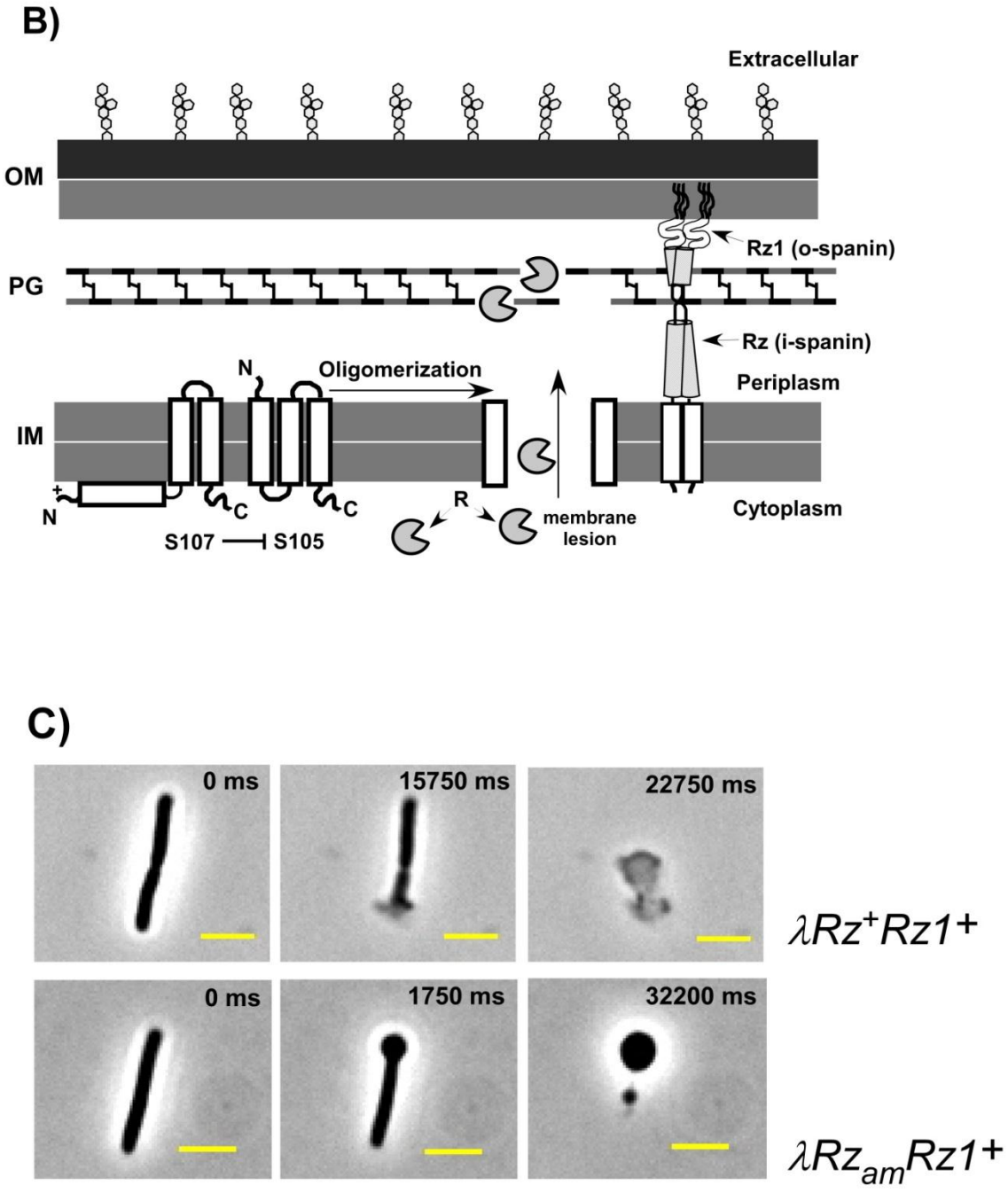



Figure 1.3. Continued.

The lytic cycle begins with the activation of the late promoter pR' resulting in the production and accumulation of the lysis proteins. At a defined time, these lysis proteins act to cause lysis. Among these lysis proteins are the holin (S105) and antiholin (S107), which are the small membrane proteins that target the first barrier of the cell envelope, the IM. Both of these proteins have three transmembrane domains (TMDs). The topology of S105 is N-out-C-in whereas the S107 exist in an N-in C-in topology (Fig. 1.3B) (38). Compared to S105, S107 has two extra residues, Met and Lys, at its N-terminus (inset of Fig. 1.3A). The positive charge at the N-terminus of the S107 confines its TMD1 to the IM-cytoplasm interface by inhibiting its entry into the membrane bilayer (39). The accumulation of S107 in the IM with this alternative topology allows the anti-holin to inhibit S105 function (see Fig. 1.3B) (35). According to the "death raft" model that has been proposed by Wang et al.(40), the holins accumulate in the IM in 2D aggregates, or rafts which are characterized by close helical packing of TMDs that excludes lipid. When this aggregate reaches a critical concentration in the IM, the rafts lose the ability to support the PMF. The notion is that holin molecules within this aggregate somehow undergo a conformational change which makes a gap (40, 41). The formation of this gap is initiated due to holin triggering and a slight drop in membrane potential causes the TMD1 of S107 to assume an N-out topology. This in turn causes S107 to promote hole formation resulting in the formation of large lesions in the IM. Direct visualization of these holes by cryo-electron microscopy has revealed large gaps in the IM ranging from 300 nm to 1000 nm in diameter (42). The major function of the hole is to allow the soluble endolysin to gain access to the periplasm (see Fig. 1.3B).

The λ endolysin is a transglycosylase that cleaves the β-1,4 glycosidic linkage between MurNAc and GlcNAc, the result of which causes degradation of the PG (37). By degrading the PG, the host cell loses its rod shape and is now surrounded by "lesions" in the IM and the intact OM.

The third step in λ lysis involves the disruption of the OM. λ utilizes the translational products of *Rz* and *Rz1*to target the OM for lysis (Fig. 1.3C). Rz, an integral membrane protein, and Rz1, a lipoprotein, accumulate in the IM and OM, respectively (8). These two proteins also interact with each other forming a complex, referred to as spanin complex (see below). The current model for spanin function is that the spanin complexes are trapped within the PG meshwork until the degradation of PG by the endolysin. This allows spanin complexes to become free from the meshwork to disrupt the OM to complete lysis.

**Other ways to break barriers by holin-endolysin-spanin system**

The previous section described the λ lysis paradigm where the holin of λ opens a gap in the IM, allowing endolysin to access the periplasm. In the third step, spanins disrupt the OM to complete lysis. The hole formation event is temporally regulated and endolysin is restricted from PG degradation until the triggering of holin. There are three types of holins and two types of endolysin (Fig. 1.4). All three classes of holins are extensively characterized to the molecular level to understand their function and how they work. Although all holins share the property of permitting timed access of

endolysin to the periplasm, they are diverse in terms of topology and sequence similarity. For example, the holin of T4 (class III) has one TMD, whereas λ (class I) has three (Fig. 1.4). In spite of different topology, both λ and T4 holins make large holes (43). The class II holins have two TMDs (Fig. 1.4) that can make small holes. In lambdoid phage 21, the holin gene, $S^{21}$, has a dual start motif, similar to λ *S*. $S^{21}$ encodes two products: $S^{21}68$ and $S^{21}71$, representing holin and antiholin, respectively. Both products possess two TMDs with N-in, C-in topology. The class II anti-holin has the same functional property as the λ anti-holin but the mechanism of hole formation is different. Genetic and biochemical data indicate that the TMD1 of $S^{21}68$ is dynamic and flips into the periplasm, allowing TMD2 to align to form a hole in the membrane (44, 45). This is drastically different from the way that the class I holin work. In addition, the $S^{21}$ holin makes lesions that are more than 100 fold smaller than canonical holes (46). These smaller size lesions are called "pinholes" (Fig. 1.4B). In this case, the passage of λ R-size endolysins would be impossible. Therefore, these holins always pair with a separate class of endolysin called a SAR (Signal Anchor Release) endolysin (47). SAR endolysins are secreted into the periplasm in a *sec*-dependent manner; however its N-terminal end is anchored to the lipid bilayer by a unique hydrophobic domain which traps the enzyme in an inactive conformation. The typical SAR domain has reduced Leu and Val content in favor of Ala, Gly, and Ser residues. The TMDs of SAR endolysins are sensitive to the collapse of the membrane potential, meaning when pinholins form small holes in the IM, the collapse of the membrane potential causes the release of SAR endolysins into the periplasm. Upon release, the periplasmic domain of SAR endolysins

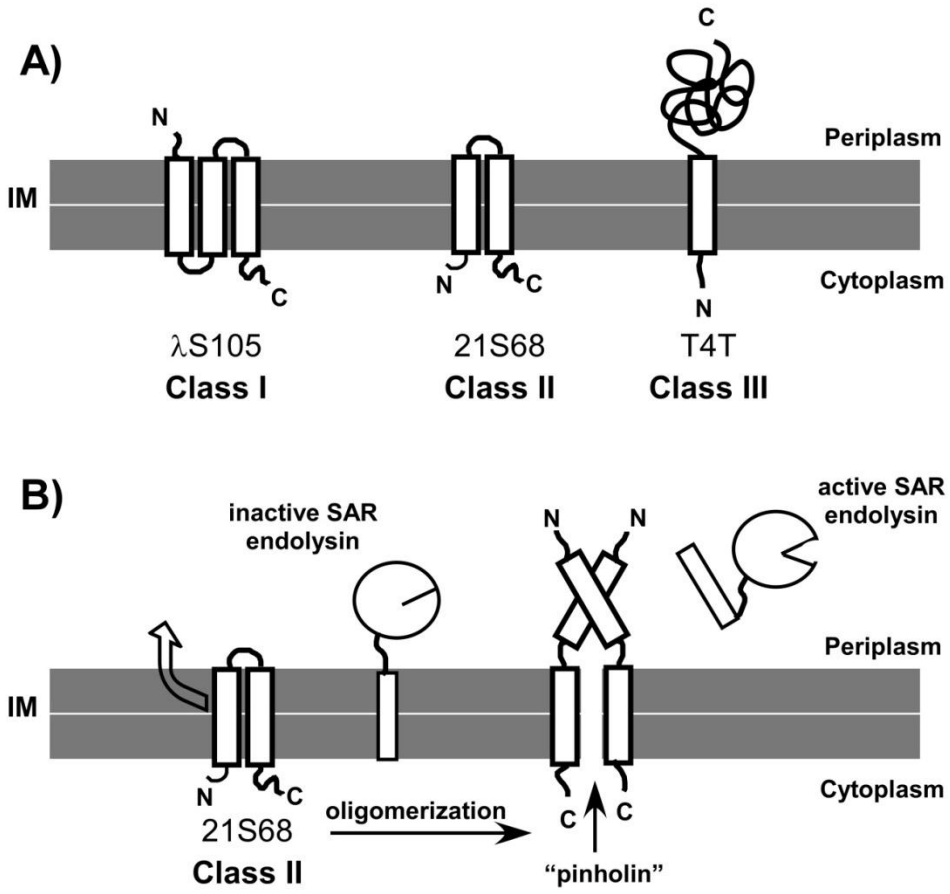



Figure 1.4. The holin classes and topological dynamic of class II holin. **(A)** Shown here are the three classes of holin topology. Prototypes phage members λ, 21, and T4, are indicated at the bottom of each cartoon. The N-terminal and C-terminal ends are indicated by N and C. **(B)** Illustrated here is the topological dynamics of the class II holin of phage 21. The release of TMD1 into the periplasm (block arrow) makes the holin functional, which leads to formation of a small lesion referred to as a "pinholin". Formation of pinholes in the IM (upward arrow) will cause the release of the SAR endolysin into the periplasm in an active form (see text for more detail).

undergoes conformational change, which activates the endolysin, leading to PG degradation (Fig. 1.4B) (48). This alternative lysis strategy is called pinholin-SAR endolysin paradigm.

Canonical holins have no specificity in regard to the type of endolysin they use. However, pinholins can only be paired with SAR endolysins. Regardless of holin-endolysin type, all characterized phages of Gram-negative hosts require spanins to disrupt the OM. Two-component system (2CS) spanins, similar to λ Rz and Rz1 are predominantly found in dsDNA phages of Gram-negative hosts. A second class of spanins, which contains a single-component (1CS) spanin, has been described (7). These spanins, such as Gp11 in phage T1, are referred to as unimolecular spanins or u-spanins (Fig. 1.5). Despite a completely different structural arrangement, Gp11is able to complement the lysis defect of λ spanin null allele (7). Moreover, both types of spanins function independently of holin-endolysin or pinholin-SAR endolysin system. The only common requirement is that spanin function requires the degradation of PG.

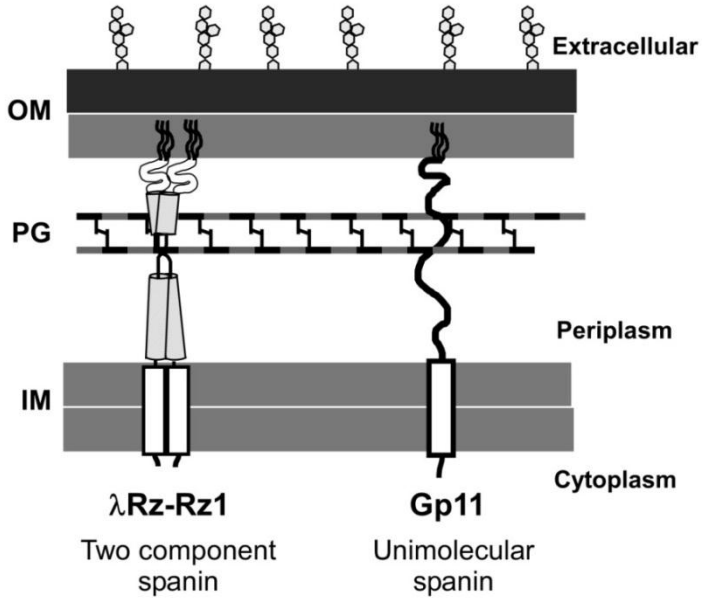



Figure 1.5. The topologies of two classes of spanins. The cartoon on the left shows the 2CS spanins of phage λ. The cartoon on the right illustrates the unimolecular spanin of phage T1 (see text for detail).

**Discovery of λ *Rz* and *Rz1***

The *Rz* gene was first discovered in 1979 by Young et al. (5) during the characterization of mutants obtained from transposon Tn903 insertion in the λ chromosome. Two mutants, λ*dk6* and λ*dk23*, were of special interest. It was evident from the phenotype that both mutations had no effect on the accumulation of infective phage progeny but were defective in lysis. Both mutants were derivatives of phage

λ$S_{am7}$, a phage that infects and lyses a *supF* host. λ*dk6* and λ*dk23* failed to cause lysis of *supF* host. However, the lysis defect depended on the presence of 10 mM $MgCl_2$. Physical mapping of the λ chromosome revealed that the location the transposon insertion was near the essential lysis gene, *R*. The newly-identified gene was named "*Rz*". When the defect was monitored under the microscope, cell morphology changed from rod to spherical shape. This indicated that the host cytoskeleton was compromised by the endolysin activity, although lysis was incomplete. It was suggested that the loss of the rod shape was the intermediate step of the lysis process. In 1983, Taylor et al. confirmed the location of the Tn903 insertion, which was about 40 bp downstream of *R* gene and verified the existence of an open reading frame of *Rz* (49). However, the biochemical properties of Rz and its role in lysis were not yet known. At the time, transglycosylase and endopeptidase activity was detected in the λ lysate. This made it tempting for the Taylor group to assign the latter activity to Rz, since the transglycosylase was attributed to R (50). However, no enzymatic activity of Rz had been demonstrated. The protein sequence of Rz is shown in Fig. 1.6A.

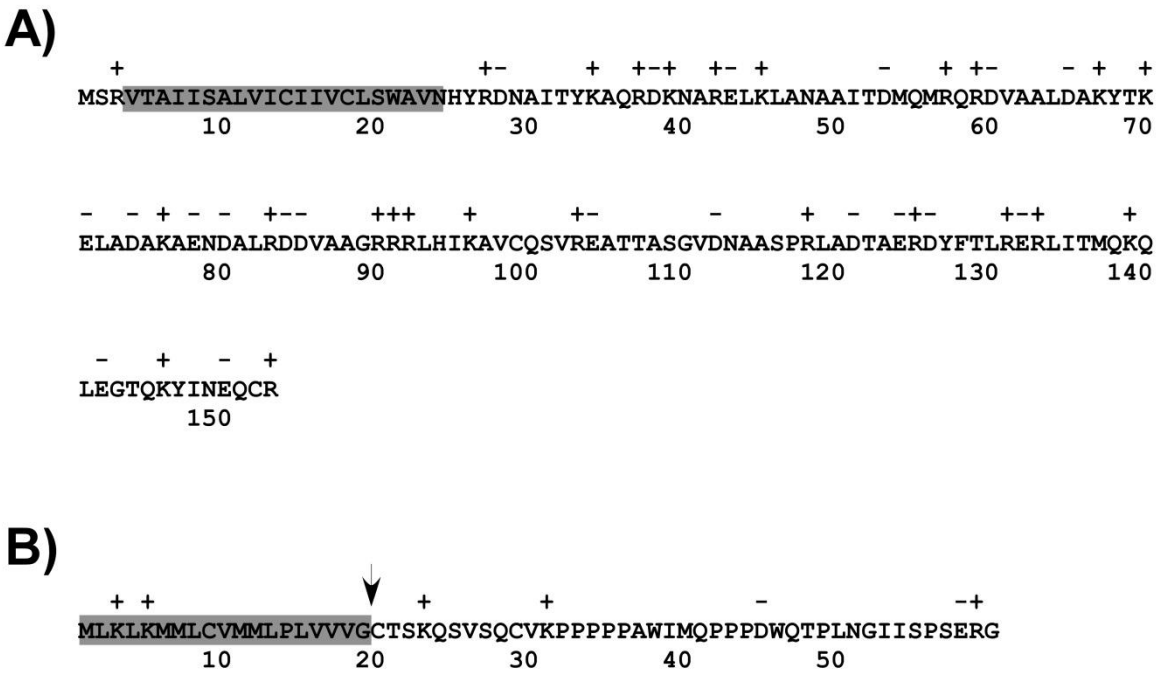



Figure 1.6. The translational product of **(A)** Rz and **(B)** Rz1. Charged residues are depicted by "+" and "-" signs. Residue numbers are indicated at the bottom of each sequence. The respective signal sequence and TMD of Rz and Rz1 are shaded in dark grey near the N-terminal end. The downward-pointing arrow above the Rz1 sequence indicates the cleavage site of Rz1 by signal peptidase II.

The continued effort to purify Rz for biochemical characterization led to the discovery of an unexpected 6.5kDa protein (51). The expression of this protein came as a surprise as the product of an adventitious cloning of the *Rz* gene with a frame-shift mutation. The newly-identified translation product was from a different open reading frame (ORF). The coding sequence of this new ORF was entirely embedded in the *Rz* gene in the +1 reading frame with its own Shine-Dalgarno sequence. This gene was

named "*Rz1*". The function of *Rz1* was not determined at the time but several predictions were made. After analyzing the protein sequence of Rz1 (Fig. 1.6B), a lipoprotein signal sequence was apparent in the N-terminal end. The presence of a lipobox sequence containing the conserved cysteine residue at the N-terminal end of a hydrophobic domain of Rz1 suggested the possibility of a post-translational lipid modification similar to bacterial lipoproteins. Furthermore, the Thr and Ser residues at the +1 and +2 position of the cysteine residue at the 20$^{th}$ position suggested that the Rz1 would be directed to the inner-leaflet of the OM. Based on these criteria, Rz1 was predicted to be an OM lipoprotein. Experiments including membrane separation, palmitate labeling, and use of globomycin as an inhibitor of SpaseII confirmed the prediction that Rz1 was indeed an OM lipoprotein. Despite data indicating the localization and modification, the function of Rz1 remained unknown (52). When a complementation assay failed to rescue the *Rz* defect, it was suggested that *Rz1* might be involved in a pathway independent of lysis. This argument was later nullifed because Zhang et al. (6) demonstrated that the individual plating of the recombinant phages $\lambda Rz_{Q100am}Rz1^{+}$ and $\lambda Rz^{+}Rz1_{W38am}$ on a non-suppressor strain blocked lysis in a divalent cation-dependent manner. It was unprecedented for two genes sharing the same DNA to have an identical phenotype. The shared phenotype suggested that both genes were involved in the same biological pathway. The aforementioned amber mutations could be rescued by co-infection with both amber mutants, indicating that *Rz* and *Rz1* complemented each other *in trans*. In summary, *Rz* and *Rz1* are two genes sharing the same DNA with the ability to express two different protein products involved in the same step of the lytic pathway. This is

unique in biology. The shared DNA architecture raises an interesting question about why evolution paired *Rz* and *Rz1* with such a tight genetic linkage.

**Ubiquity and diversity of *Rz-Rz1* equivalents and the discovery of *gp11***

Following the discovery of λ *Rz* and *Rz1*, functional spanin homologs were identified in many other dsDNA phages of Gram-negative hosts. For example, in 1989, Casjens et al. reported gene *15* of *Salmonella* phage P22 being 59% identical to λ *Rz* and 22% identical to T7 *18.5*, a functional homolog of *Rz* (53). At the time, the P22 *Rz1* equivalent was, however, not identified due to its embedded gene arrangement. In 1994, Ziermann et al. performed a genetic analysis of the P2 lysis cluster and found *lysB* amber mutant that delayed lysis under standard laboratory conditions. At this time the function of *lysB*, a functional homolog of *Rz*, was overlooked for unknown reasons (54). Later in 2004, the overlapping gene *lysC* that extended beyond the stop codon of *lysB* was identified. In addition, the function of *lysB/lysC* pair was identical to λ *Rz/Rz1* pair by complementation (55). This was the first example when *Rz/Rz1*homologs were identified in overlapped architecture.

Recently, Summer et al. (7) performed a global search for spanins in the NCBI phage genome database. Due to the enormous diversity of phages, finding *Rz-Rz1* homologs based on sequence identity alone was difficult. This is especially true with respect to finding a gene within a gene. Therefore a new search strategy was developed.

Search criteria based on gene arrangement and membrane localization signals were employed.

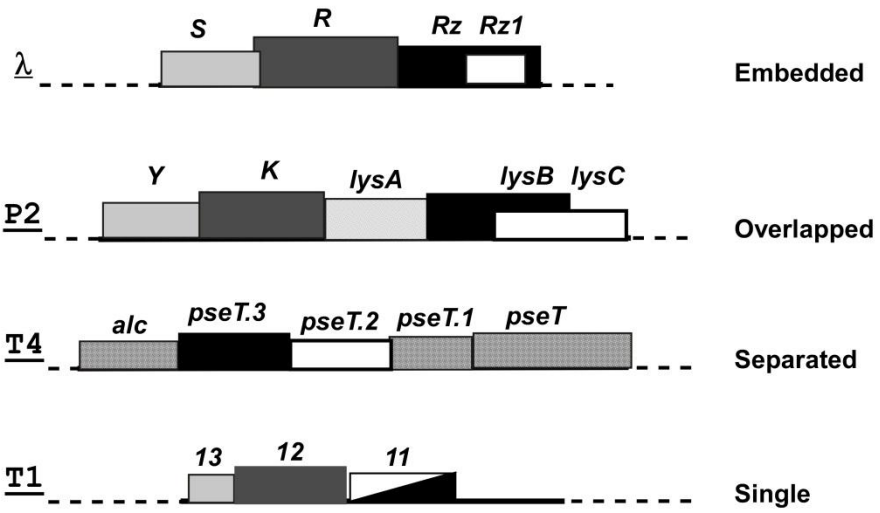



Figure 1.7. The four different classes of spanin gene arrangement. The class of 2CS spanin arrangement is indicated in the right of the figure. From top to bottom: *Rz/Rz1*, *lysB/lysC*, and *pseT.3/pseT.2* are the spanin pairs of phage λ, P2, and T4, respectively. Gene *11* represents the 1CS spanin of phage T1 (7).

After genome analysis of 137 dsDNA phages, it was revealed that nearly all dsDNA phages of Gram-negative hosts possess *Rz/Rz1* equivalent pairs (7). It was clear from this analysis that spanins are ubiquitous in phages of a Gram-negative host. Furthermore, this study had two key findings: **1)** In addition to the embedded and overlapped gene arrangement described above, there was a third class of spanin genetic

architecture. In this arrangement, the genes were either separated by intervening nucleotide sequence or closely associated without any overlap. One such example is *pseT.3/pseT.2* pair of paradigm phage T4 shown in Fig. 1.7. Using classification based on the genetic architecture of spanins, 45 embedded, 51 overlapped, and 18 separated pairs were identified. **2)** A unique class of gene was identified in 7 different phages of T1 family. In this case, the equivalent *RzRz1* pair was represented by a single gene (Fig. 1.7). The predicted translational product of this gene had a signal sequence of OM lipoprotein at its N-terminus (like of Rz1) and a TMD at its C-terminal end that was predicted to localize in the IM (like Rz). This suggested a single gene product would span the periplasm connecting the IM and OM (topology shown in Fig. 1.5). Confirming predictions, *gp11* of T1 complemented the *λRzRz1* pair. This striking feature of a single gene product led to the derivation of the term "spanin," a lysis protein that spans the periplasmic space. This is how single spanin, now referred to as unimolecular spanin or u-spanin, was discovered (7, 56).

This bioinformatic survey led to new questions about spanin function and their evolution. Why are there three different types of gene arrangements in the 2CS? Did these pairs originate as separate genes before evolving into an embedded arrangement? Is the overlapped pair an intermediate state of evolution or the other way around? Recently, two phages were isolated against *X. fastidiosa* named Salvo and Sano. Each phage had a different arrangement of spanin genes, overlapped in Salvo and separate in Sano, although they both belonged to the same family (57). In both pairs, the C-terminal ends were identical yet belonged to different spanin classes based on the gene

arrangement. How is this possible? Based on gene alignment and protein sequence similarity, an answer has been proposed: deletion of an intervening sequence caused two separate genes to overlap each other in a different frame without affecting the functional domains. If this is true then insertion of a large nucleic acid fragment should also be able to separate genes. These possibilities suggest two separate genes can move in or out of each other. It is, however, important to realize that in every case where 2CS spanin genes are separated, i-spanin is located transcriptionally upstream from o-spanin without intervening genes (7). One simplest interpretation is that minimizing the separation of two genes that are involved in the same biological function could prevent their segregation by recombination. It is unclear if there was a linear evolution from separated spanin gene arrangement to embedded arrangement or vice versa. But why do 2CS genes are predominantly found among phages when unimolecular spanins are well equipped to disrupt the OM, demands further research.

In summary, the diversity of spanin genes is high and they are ubiquitously found among dsDNA phages of Gram-negative host. The discovery of Gp11 as a single spanin was a key result of the bioinformatic survey because it justified the necessity of interaction between i-spanin and o-spanin (described latter). It should be noted that spanins have not been identified in phages of cyanobacteria and mycobacteria. This may be due to the unusual OM structure of cyanobacteria and mycobacteria (58-61).

## Molecular properties of λ Rz and Rz1

### *Rz: an integral type II membrane protein*

Rz is a 153-residue protein (Fig.1.6A) with a predicted molecular mass of 17.2kDa. The sequence analysis of Rz indicates that it has a TMD in the N-terminal end. The C-terminal end has a soluble domain consisting of highly charged residues. Moreover, the charge distribution flanking the TMD indicates that Rz is a type II integral membrane protein (N-in, C-out topology). A membrane separation experiment clearly demonstrated that Rz is exclusively localized in the IM (8). Substitution of the TMD of Rz with the heterologous TMD of FtsI does not abrogate the function of Rz. FtsI is a type II integral membrane protein; therefore, the FtsI-Rz construct empirically supports type II topology predictions for Rz (8). What is the role of the TMD of Rz? Another chimera of Rz was constructed with the cleavable signal sequence of PhoA, a secreted periplasmic protein. Expression of PhoA-Rz confers a defect in Rz function, suggesting that the anchoring of Rz to the membrane is essential for its function (8). Moreover, the dominant-negative phenotype of PhoA-Rz indicates that the periplasmic domain of Rz interferes with the function of wild type Rz. Structural prediction suggests the periplasmic domain of Rz consists of two α-helices that are interconnected by a flexible linker. These helices have a signature of heptad repeats, $(abcdefg)_n$, in which positions "a" and "d" are generally hydrophobic residues that are involved in the interaction between α-helices to form a coiled-coil structure (62, 63). Furthermore, the use of

COILS (62) and PairCoil (64) structural prediction programs suggest that these two helices of Rz have strong propensity to form a coiled-coil structure. Despite the diversity in sequence identity, the bioinformatic analysis of other Rz homologs is highly consistent with this predicted structure.

***Rz1: an OM lipoprotein***

Rz1 is prolipoprotein that is translated as a 60-residue long polypeptide (Fig. 1.6B). The sequence analysis of Rz1 indicates that it has a hydrophobic domain at the N-terminal end. This domain matches the signal sequence of a lipoprotein. This 20-residue long signal sequence is cleaved by signal peptidase II, followed by maturation of the protein by addition of lipid moieties in the N-terminal end. Thus, the mature form of Rz1 consists of a 40-residue long lipoprotein. Genetic and biochemical experiments have confirmed that the N-terminusof Rz1 is processed in a similar way to the host lipoproteins (52, 65, 66). Membrane separation studies have demonstrated that Rz1 is exclusively localized in the OM (8). Localization of Rz1 in the OM has been blocked by replacing a negatively charged residue in the +1 and +2 positions of lipoylated Cys. However, directing Rz1 to the IM results in a loss of function indicating that localization of Rz1 in the OM is essential for its function. The periplasmic domain of Rz1is composed of unusually high proline content (10 out of 40 residues), indicative of highly unstructured region. However, the structural role of these polyproline stretches is

currently unknown. High proline content is conserved in Rz1 homologs of λ family, which is an indicator of its importance in o-spanin function.

***The soluble domain of Rz and Rz1 interact to form a complex***

Genetic and biochemical experiments have indicated that Rz and Rz1 interact *in vivo*, forming complex that is essential for function (6, 8, 67). However, there is no available structure of either spanin subunit or spanin complex that has been resolved to the atomic level. This is mainly due to the difficulty of purifying membrane proteins. Nevertheless, structural study of the soluble domain of Rz and Rz1 (denoted sRz and sRz1) was carried out using CD spectroscopy and transmission electron microscopy (TEM) (9). Consistent with structural predictions, CD analysis indicated that sRz exhibits α-helical content whereas as sRz1 is primarily unstructured. Mixing of sRz and sRz1 forms a complex of equimolar stoichiometry accompanied by a significant increase in α-helical content. This is indicative of rearrangement of structure triggered by protein-protein interaction. It has not been resolved which residues are involved in this process. Remarkably, these complexes are rod-like bundles which are~ 24 nm in length and ~4 nm in width. Interestingly, the average length and width of the sRz-Rz1 complex matches the average length of the periplasmic space of *E. coli* and predicted width of the lacunae. These structures also displayed a helical pattern reflecting the property of coiled-coil elements of Rz.

***Rz and Rz1 are covalently linked homodimers***

Another intriguing feature of Rz and Rz1 is that both exist as intermolecular disulfide-linked homodimers (68). There are two cysteines at position 99 and 152 of Rz and a single cysteine at position 29 of Rz1 (Fig.1.5). When the redox state of the individual thiol group was assessed under physiological conditions, all three cysteine residues were found to participate in the disulfide linkage. Given the existence of the u-spanin (Gp11 in T1, which lacks any cysteine residue in the periplasmic domain), it was thought that these cysteine residues must be heterotypically linked between Rz and Rz1. Contrary to this expectation, Rz exists as a homodimer due to intermolecular disulfide linkages between residues 99-99, and residues152-152. Similarly, the Rz1 also exists as a homodimer due to an intermolecular disulfide linkage between residues 29-29. In the earlier experiment, the reduction of each disulfide linkage one at a time by expressing *Rz*$_{C99S}$*/Rz1*, *Rz*$_{C152S}$*/Rz1*, or *Rz/Rz1*$_{C29S}$ allele had no effect in spanin function. Recently we have found that these disulfide linkages do play an important but unknown role in spanin function (more details in Chapter II). The apparent degradation product of single Cys-less allele of *Rz* and *Rz1* in previous experiments have indicated that disulfide linkage might assist in stabilizing the spanin complex (68).

**Development of a model for Rz-Rz1 function**

Described above are the molecular features and phenotype associated with λ spanins and their homologs. It is evident that spanins are essential for lysis. However, the exact mechanism to disrupt the OM is unknown. Several hypotheses have been proposed to address how Rz and Rz1 function at the molecular level.

The enzymatic property to degrade the cross-linking peptide of glycan chains in PG or the peptide of Braun's lipoprotein was once assigned to λ Rz (50). The argument was that the degradation of these peptides by Rz would somehow destabilize the OM causing lysis. This led to the annotation of Rz-like proteins as endopeptidases. However, no such enzymatic activity of purified Rz has been identified. Moreover, the predicted secondary structure and topology of Rz does not align with catalytic domains of known proteases. Therefore, the enzymatic activity related to Rz is no longer valid.

Krupovic et al. proposed a different model while studying Gp36 and Gp37, a functional spanin pair of a phage PRD1 (69). According to this model, spanin complexes are pushed away along with the membrane lipids during the hole formation. Such movement would cause a mechanical stress that is directly transmitted to the OM via the spanin complex subjecting the OM to form a gap. This is an attractive model except that spanin function in this model is dependent on the holin action. Recently, the function of λ spanin was demonstrated to be holin-independent, which invalidates the proposed model (8).

When the membrane composition was studied prior to the establishment of spanin function, it was unclear why a fraction with intermediate density comprising of IM and OM content was enriched during phage infection or prophage induction (70, 71). When a similar observation was made by overexpressing *Rz1*, Bryl et al. (72) interpreted that this lipid mixing was due to Rz1activity and proposed that Rz1has a fusogenic property ascribing to its high proline content (66, 72). The fusogenic property of Rz1 was demonstrated by reconstituting Rz1 into a liposome that catalyzed the release of content from a target liposome. This led authors to conclude that Rz1 alone can disrupt the OM. However, this is contrary to the genetic evidence because it has been established that both *Rz* and *Rz1* are required for lysis. Recently, after an extensive characterization of λ Rz and Rz1, a membrane fusion model has been developed that includes both Rz and Rz1(8-10). This model will become clear once the molecular features described above are superimposed with the brief review on membrane fusion mechanism described in the following section.

**Membrane fusion**

Membrane fusion is a fundamental biological process central to diverse functions. Entry and exit of enveloped virus, mating between two haploid yeast cells, trafficking of vesicles to and from the cell surface, or bacterial cell division by binary fission are all processes which are dependent on the fidelity of membrane fusion (73-76). Despite this diversity, all membrane fusion reactions share the same elementary

processes shown in Fig. 1.8 (77-80). The first step of a general fusion features two membranes which are brought to close proximity (step II). It is followed by membrane merging, referred as the intermediate state or the hemi-fusion state (step III). The last step is the opening of an aqueous fusion pore (step IV). The fusion of two membranes, however, is not a spontaneous process. This is because the lipid bilayer is an inherently stable structure. Despite their stability, the lipid membranes can undergo remodeling processes necessary for membrane fusion, pore formation, or lipid exchanges between different membranes (81). However, to initiate fusion, a localized destabilization of the membrane bilayer is necessary after the membranes are brought into close contact. This is a rate limiting step in the fusion process because destabilization of the membrane bilayer imposes a high energy barrier (82-84). The consensus in the field is that the destabilization of the membrane bilayer is much easier if the membrane is bent (85-87). Membrane contortion is generally produced by fusion proteins that are ectopically localized on the membrane bilayer. These proteins lower the energy barrier by pulling the opposing membranes together. This process is reflected in the hemi-fusion step, where the outer leaflets of the donor and target membrane are highly curved and fused while the inner leaflet is intact (Fig.1.8). According to the stalk-hypothesis, this intermediate structure is a necessary step in membrane fusion (83, 84). In a subsequent step, this stalk radially expands until the trans-monolayer meet to form a new bilayer called hemifusion diaphragm. The rupture of this diaphragm forms the fusion pore completing the fusion process (88). There is an alternate pathway for membrane fusion which is called the direct fusion pathway (88, 89). In this pathway, fusion proteins line

up to form a hemichannel structure on both sides of the opposed membranes. When two hemichannels come together, a fusion pore is established to allow the exchange of lipids and aqueous contents. However, there is less evidence to support this model.

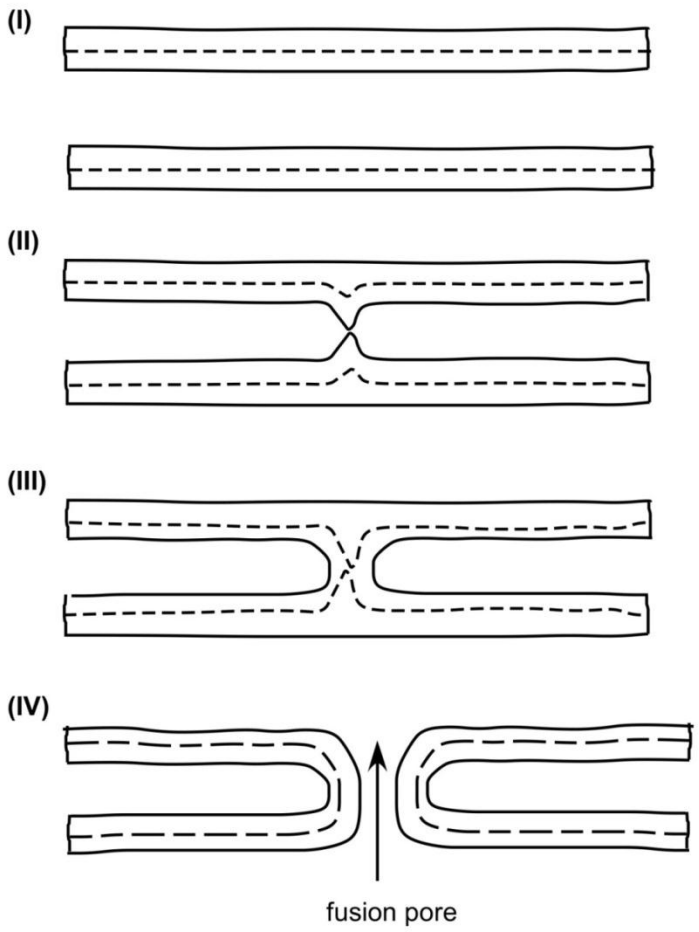



Figure 1.8. General fusion mechanism (I) Two separate membrane bilayers. (II) Initiation of membrane contact between two separate membrane bilayers. (III) Formation of a fusion stalk – the hemifusion state. (IV) Establishment of a fusion pore illustrated by the upward-pointing arrow. The proteins involved in the fusion process are not shown for simplicity.

A detailed characterization of fusion proteins and their interaction with the membrane is required to understand the molecular mechanism of membrane fusion. Fusion proteins of eukaryotic viruses, yeast, plants, and mammals have been studied in great detail. Structural studies of some these fusion proteins indicate that fusion proteins are highly complex and diverse (90-92). Although the fusion proteins of different systems are not identical, extensive amount of work that has been done in the membrane fusion field implies that the membrane fusion pathway is universal. This is because the central features of the most fusion proteins are similar despite the lack of their sequence similarity. The main complexity lies within the regulatory mechanisms of different membrane fusion systems. Here we will briefly describe the central components of the eukaryotic and viral membrane fusion systems in an effort to compare common features.

***SNARE assembly: a central feature of membrane fusion in eukaryotic cells***

In eukaryotic cells, soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs) are known to play a central role in membrane fusion. There are about 36 SNARE genes in mammalian species, 25 in yeast, and 54 in Arabidopsis (93, 94). Remarkably, proteins of all SNARE families share a common structural architecture (12, 95, 96). SNARE proteins are anchored to the membrane either by TMD or a post-translationally modified lipid anchor. All SNAREs feature a domain that has a propensity to form a coiled-coil structure. These conserved features imply conserved function of these proteins. SNAREs are classified based on their association with the

membrane: v-SNAREs for vesicle-associated and t-SNAREs for target membrane-associated (97). v-SNAREs and t-SNAREs interact to form a complex which brings opposing membranes into close proximity. The hallmark of SNARE assembly is the formation of a four-helical bundle mediated by coiled-coil assembly. However, this interaction is highly regulated to prevent fusion initiation when it is not necessary (98). The most well-studied membrane fusion system is the synaptic SNAREs system, which is used for neurotransmitter release. One of the t-SNAREs, called syntaxin 1-A, has N-terminal regulatory domain which interacts with SNARE motifs to inhibit SNARE assembly (99). Similar regulatory mechanisms have also been identified in yeast and plants (100, 101). SNARE triggering is the result of a chemical influx which dislodges the regulatory domain and promotes interaction of the SNAREs. The assembly of SNAREs is accompanied by significant conformational changes that lead in a pathway to membrane fusion, first bringing the opposing membrane bilayers into close proximity (102, 103). After fusion, these *trans*-SNARE complexes are converted to *cis*-complexes, named as such because all complexes end up in the acceptor membrane. SNARE complexes are extremely stable. In eukaryotes SNAREs complexes are disassembled by specialized chaperones using an energy source from ATP hydrolysis (104). Disassembled SNAREs are recycled for the next round of membrane fusion.

***Membrane fusion by mammalian virus***

Based on structure, topology, and size, viral fusion proteins are categorized into four different classes (105). The most well-characterized viral fusion proteins are hemagglutinin (HA) of influenza and gp41 of HIV, both representative of class I. The functional and structural details available from the study of these proteins have largely influenced the field of membrane fusion. This is partly due to the similarity in the fundamental principles between viral fusion proteins and SNARE system described above (77, 106). Mostly, viral fusion proteins, which are glycoproteins, have two membrane-interacting domains: a C-terminal TMD that anchors the protein in the viral membrane and a disparate "fusion peptide" or "fusion loop," composed of hydrophobic patch at the N-terminus that interacts with the target membrane (90). In addition, all these viral fusion proteins identify specific receptors in the target membrane. These fusion peptides are initially hidden until triggered by low pH environment or other cellular receptor ligands (74). This allows the extension of the fusion protein to interact with the target membrane. This extension is considered to be a fusion-active state (107). Remarkably, all class I viral fusion proteins in this state function in a trimeric form (90). The central feature of the trimeric structures is that they are driven by a coiled-coil domain. Another common feature of viral fusion proteins is the conformational change that occurs prior to fusion. The extended fusion peptide holding the target membrane folds back into a hairpin conformation bringing the two membranes into close vicinity. In hemaglutinin of influenza and gp41 of HIV, this hairpin folding mechanism is

achieved by so called “loop to helix” transition (74, 107, 108). It is unclear if the same mechanism is true for all viral fusion proteins. Regardless of the mechanism, all class I viral fusion proteins brings the two membrane bilayers into close proximity for fusion by bringing their C- and N-terminal domains into proximity. The post-fusion conformation of viral proteins is irreversible. Therefore, in viruses, membrane fusion is a “one shot” mechanism. Multiple studies have proposed that the energy necessary for membrane fusion comes from the conformational change of the fusion protein (74, 76, 80, 88, 106).

***Spanins are similar to viral fusion protein***

Given the predicted structural properties and molecular characteristics of spanins, it is likely that the spanins of λ function in a similar manner to other viral fusion proteins. Most importantly, the coiled-coil domains are known to be a key element of all fusion proteins. Such coiled-coil domains are well conserved in λ Rz and its functional homologs. Additionally, the spanin complex is known to make a direct contact with two membrane bilayers (7, 8). Furthermore, *in vitro* experiments with the soluble domains of Rz and Rz1has revealed the accompanied conformational change upon their interaction (9). In general, such conformational changes are thought to be the source of free energy required for fusion reaction. The Rz1 structure contains unusually high proline content in its periplasmic domain. Recently, the feature of similar high proline contents was described for p15 of non-enveloped reovirus (109). These non-enveloped viruses deploy fusogenic protein to the surface of host cell membrane after infection to cause cell-cell

fusion. The polyproline stretch in p15, that is similar to the pentaproline stretch of Rz1, was shown to be critical for fusion. These are different class of viral fusion protein known as fusion-associated small transmembrane (FAST) protein. Altogether, the structural features of spanins are reminiscent of different kinds of viral fusion proteins. . Moreover, spanins act in the last step of lysis for phage release. This suggests that there is no requirement to reverse the spanin complex for another round of fusion which is analogous to viral entry or exit by fusion.

**Questions to be addressed**

As described above, the role of spanin is to disrupt the OM in the final step of phage lysis. It has been proposed that spanins fuse the IM and OM during lysis. The mechanism is, however, unknown. Although most of the work presented in the following three chapters does not provide a mechanism for spanin function it has made a significant progress in understanding some of the key aspect that supports the idea of IM-OM fusion during lysis. Most of the work described in Chapter II was built upon previous finding about spanins being essential for lysis and their existence as a disulfide-linked homodimer *in vivo*. What we have addressed in Chapter II is of two parts: 1) essentiality of the spanins for lambda lysis in different physical conditions and 2) requirement of the disulfide linkage for spanin function. Based on the findings of the second question, in Chapter II we also addressed the molecular pathway for the covalent maturation of spanin subunits. In Chapter III, we took a genetic approach in an effort to

address the steps involved in the spanin function. In Chapter IV, we designed a new technique to address the membrane fusion model. In Chapter V, we have discussed the conclusion and future directions.

# CHAPTER II

# STRUCTURE AND FUNCTION OF THE LAMBDA SPANINS*

## Introduction

Spanin genes are found in nearly every double-stranded DNA phage which infects a Gram-negative host (7). Spanins can be separated into two groups; single-component system (1CS) spanins, and two-component system (2CS) spanins, as found in lambda. To adopt a common terminology, we will refer to *Rz* and *Rz1* equivalents i-spanin and o-spanin, respectively. Many spanin gene pairs share the embedded organization found in lambda *RzRz1* (Fig.2.1). A variant architecture is the overlapped class, whereby the o-spanin gene extends past the end of the i-spanin gene (Chapter I, Fig.1.6). The third gene architecture is called the separated class; the i-spanin and o-spanin genes lack any overlap (Fig.1.6). Alternately, some phages encode a 1CS spanin protein. These proteins are designated as u-spanins (unimolecular-spanins) (Fig.1.6). Despite the ubiquity, diversity, and intriguing molecular features of the spanins, the lysis phenotype associated with the *Rz* and *Rz1* defect was characterized as conditional.

---

* Part of this chapter is reprinted with permission from "The spanin complex is essential for lambda lysis" by Berry J., Rajaure M., Pang T., and Young R.,2012. *J. Bacteriol.* 194:5667-5674, 2012 by American Society for Microbiology.
Part of this chapter is reprinted with permission from "Spanin function requires subunit homodimerization through intermolecular disulfide bonds" by Berry J., Rajaure M., and Young R., 2013. *Mol. Microbiol.* 88(1):35-47, 2013 by John Wiley and Sons.

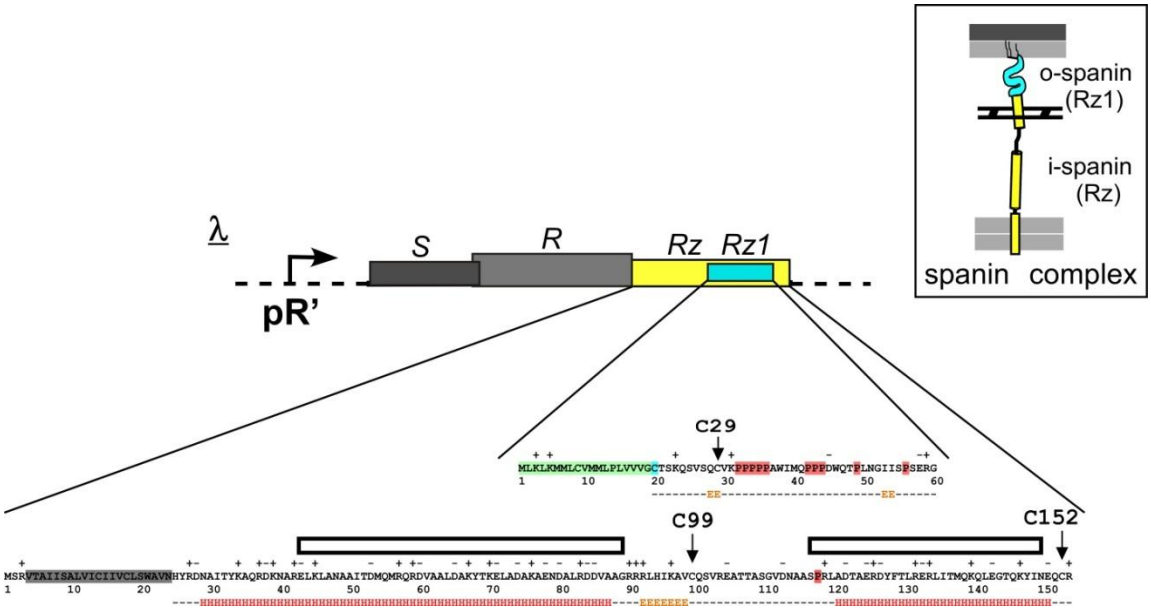



Figure 2.1. Features of λ i-spanin and o-spanin. Shown here is the λ lysis cassette. The last two genes of lysis cassette consist of *Rz*, and *Rz1*. The *Rz1* gene (teal) is embedded in the +1 reading frame of *Rz.* Inset shows the primary structures of Rz and Rz1, aligned in the actual register of the embedded *Rz1* gene (i.e., the start codon of *Rz1* begins at the second base of codon Asp73 of *Rz*). *Rz* (yellow) encodes the 153 aa inner membrane spanin subunit (i-spanin). *Rz1* encodes the 40 aa outer membrane spanin subunit (o-spanin), which is processed from its 60 aa primary translation product. The Rz (Type-II signal anchor) and Rz1 (lipoprotein) signal sequences are highlighted in gray and light green, respectively. Alpha-helical (red capital H) and extended (orange capital E) predicted by JPRED (110)(http://www.compbio.dundee.ac.uk/www-jpred/) are shown below each primary sequence. Coiled-coil regions in Rz predicted by PairCoil2 (111) (http://groups.csail.mit.edu/cb/paircoil2/) are shown by open rectangles above the sequence. Proline residues in the mature periplasmic domains and the modified N-terminal Cys of mature Rz1 are highlighted in red and teal, respectively. Arrows indicate unmodified periplasmic cysteine residues of Rz (C99 and C152) and Rz1 (C29). Boxed inset shows the spanin complex that spans the periplasm, through the meshwork of the peptidoglycan (black hatching).

This is because under standard laboratory conditions the induced lambda lysogen does not require *Rz* and *Rz1* for lysis unless the medium is supplemented with millimolar concentration of divalent cations (5, 6). It has been shown that the divalent cations enhance the stability of the OM, presumably by bridging the negatively-charged phosphate groups of the lipopolysachharide in the outer leaflet (21). When the lysis morphology of $\lambda Rz_{am}Rz1_{am}$ induced cells were examined with phase-contrast microscopy, spherical cells were observed that were indefinitely stable even in the absence of divalent cations (68). This observation inevitably changed our perspective on the role of spanins in lysis. The results were, however, difficult to reproduce under standard laboratory conditions, which include vigorous shaking of the growing culture for aeration. Therefore, reassessment of the spanin null phenotype in a bulk culture under optimized laboratory conditions was required.

Another important discovery was that Rz and Rz1 form a covalently-linked homodimers *in vivo* (68). Rz has two cysteine residues: Cys99 in the predicted hinge region and Cys152 in the penultimate position. Rz1 has only one cysteine residue, Cys29, near the middle of the periplasmic domain (Fig.2.1). Individual assessment of cysteine residues in both Rz and Rz1indicated that these disulfide linkages are homotypic: i.e., Cys99-99 and Cys152-152 linkages in the Rz dimer and Cys29-29 linkage in the Rz1dimer (68). Another feature of the periplasmic domain of Rz is that it has an extraordinarily high content of charged residues arranged in two predicted alpha-helical domains with a strong propensity for coiled-coil formation. This feature allows the purified soluble domain of thiol free Rz ($sRz_{C99,152S}$) to form a non-covalent

homodimer (9). The periplasmic domain of Rz1, on the other hand, is predicted to be mostly unstructured with high proline content (10 out of 40 residues) (Fig.2.1). In addition, the purified soluble domain of thiol free Rz1 ($sRz1_{C29S}$) was unable to form a dimer *in vitro* (9). It is unclear how Rz1 monomers are brought in proximity to form a homodimer *in vivo*. Despite Rz and Rz1 forming homodimers, the $Rz_{C99S}$, $Rz_{C152S}$, $Rz_{C99,152S}$, and $Rz1_{C29S}$ substitution alleles retain lytic function (9). This demonstrates that the aforementioned cysteines are not required for spanin function. Nevertheless, these Cys residues are conserved in the 19 other members of the lambda Rz family, as well as in the closest relatives of the lambda Rz1 family (7). The role of disulfide linkages in 2CS spanin function remains unclear.

Here we report the results of three separate experiments: (1) spanin function in liquid culture was re-examined in the absence of divalent cations; (2) the molecular pathway for the homodimerization of spanin subunits was interrogated by a biochemical approach; and (3) the prevalence of cysteine residues among other spanin homologs was assessed by a bioinformatic approach. We discuss the results in terms of a model for the subcellular localization and covalent maturation of the spanins.

## Materials and methods

### *Bacterial growth, induction, TCA precipitation, and time-lapse phase-contrast microscopy*

Bacterial cultures were grown and plated in standard LB media or on LB agar respectively supplemented with $MgCl_2$ (10 mM), ampicillin (Amp, 100 μg $ml^{-1}$), chloramphenicol (Cam, 10 μg $ml^{-1}$), and kanamycin (Kan, 40 μg $ml^{-1}$) when appropriate. Culture growth was monitored by recording A550 as a function of time. Fresh overnight cultures were diluted by transferring 100 μl of overnight culture to 25 ml of LB in a 250 mL growth flask and grown with aeration at 30ºC for lysogenic cultures and 37ºC for non-lysogens. Lysogens were thermally induced at A550 of 0.3 by aeration at 42ºC for 15 min followed by continued growth at 37ºC. At 50 minutes post-induction of the indicated lysogen and/or plasmid, concentrated Tricholoracetic acid (TCA) was combined with a 1 ml aliquot of culture to a final concentration of 10% (v/v). Further details of TCA based protein precipitation can be found elsewhere (10, 45). Video phase-contrast microscopy was conducted according to J. Berry et al. (10).

### *Strain and plasmid construction*

Bacterial strains and phages used in this study and their relevant genotypes are listed in Table 2.1. The MC1000 derivatives RI89, RI90, and RI179 were obtained from

the Coli Genetic Stock Center. Strain RI90 carries a *dsbA* insertional inactivation, *dsbA::kan*, and strain RI179 carries the knockout Δ*dsbC::cam* allele. Upon receipt of each strain, the presence of *dsbA::kan* in RI90 and Δ*dsbC::cam* in RI179 was confirmed by PCR as described in the following section. To generate strain MR135 that carries both the inactivated *dsbA* and deleted *dsbC* alleles, the *dsbA::kan* locus from RI90 was transferred into RI179 by P1 transduction (112). Briefly, candidate recombinants were selected by plating on LB-kanamycin followed by incubation at 37 ºC overnight. Kanamycin resistant colonies were screened individually for the *dsbA::kan* locus by PCR as described below. Strain RI89, RI90, RI179, and MR135 were lysogenized with phage λ900, λ900 $Rz_{am}Rz1^{+}$,or λ900$Rz^{+}Rz1_{am}$ as previously described (6) with minor modification to the selection procedure regarding RI90 and MR135. Phage λ*900* contains an insertional inactivation of the non-essential genes *stf* (*stf tfa::cam*) and *bor* (*bor::kan*). Accordingly, the RI90 selection was carried out by plating on LB-chloramphenicol. As MR135 contains both kanamycin and chloramphenicol resistance markers, a selection based on resistance to these antibiotics could not be employed. Instead, MR135 was infected with λ900 and plated as described previously (6) with the exception that overnight incubation took place at 30 ºC to ensure the temperature-sensitive *cI857* allele remained active. Candidate lysogens were then isolated by streaking cells located within the turbid center of a λ900 plaque to a fresh LB plate followed by incubation at 30 ºC overnight. MR135λ900 was isolated by cross-streaking candidate colonies against λ20, λ101, and λ108 as described previously (113). As an isogenic control, RI89 was lysogenized with λ$cI857Rz_{am}Rz1_{am}$ as described above. For

all lysogenic strains, the presence of a single copy of the λ chromosome was confirmed by screening candidate colonies as described previously (114).

Plasmids used in this study and their relevant features are described in Table2.1. The pRE plasmid is a pBR322 derivative that contains the λ late promoter pR′ located directly upstream of a multiple cloning site. The plasmids $pRz_{C99S}Rz1_{C29S}$, $pRz_{C152S}Rz1_{C29S}$, and $pRz_{C99S,C152S}$ $Rz1_{C29S}$, were generated by subcloning $Rz1_{C29S}$ fragment from $pRz1_{C29S}$ into the corresponding restriction sites of $pRz_{C99S}$, $pRz_{C152S}$ and $pRz_{C99S,C152S}$. Sub-cloning of the *Rz1* gene from pRz1 into the pRE plasmid has been previously described (8).

Table 2.1. Bacteriophages, strains, plasmids, and primers

| **Bacteriophages** | **Genotypes and relevant features** | **Sources** |
|---|---|---|
| λ900 | λ*Δ(stf tfa)::cat* $cI_{857}$ *bor::kan;* carries $Cam^R$ and $Kan^R$*;* $Rz^+Rz1^+$ | Lab stock |
| λ900$Rz_{am}$ $Rz1^+$ | | Lab stock |
| λ900$Rz^+Rz1_{am}$ | | Lab stock |
| λ900$Rz_{am}$ $Rz1_{am}$ | | Lab stock |
| λ20 | λ$b_{221}cI$ | Lab stock |
| λ101 | λ*vir* | Lab stock |
| λ108 | λ*h80* $b_{221}cI$ | Lab stock |
| **Strains** | | |
| MC1000 | *araD139 Δ(argABC-leu)7679 galU galK Δ(lac)X74 rpsL thi* | (115) |
| RI89 | MC1000 *phoR* D*ara174 leu+* | (115) |
| RI90 | RI89 *dsbA::kan1* | (115) |
| RI179 | RI89 *ΔdsbC::cam* | (115) |
| MR135 | RI179 *dsbA::kan1ΔdsbC::cam* | This study |
| RI89(λ900) | | This study |

Table 2.1. Continued

| **Strains** | **Genotypes and relevant features** | **Sources** |
|---|---|---|
| RI89(λ900$Rz_{am}$ $Rz1_{am}$) | | This study |
| RI90(λ900) | | This study |
| RI90 (λ900$Rz_{am}$ $Rz1^{+}$) | | This study |
| RI179 (λ900) | | This study |
| RI179 (λ900 $Rz_{am}$ $Rz1^{+}$) | | This study |
| RI179 (λ900 $Rz^{+}Rz1_{am}$) | | This study |
| MR135(λ900) | | This study |
| 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 |
| **Plasmids** | | |
| pRE | Plasmid with the λ later promoter pR' that is transcriptionally activated by λQ | (45) |
| pRz | pRE carrying Rz with inactivated Rz1 | (8) |
| $pRz_{C99S}$ | pRz carrying C99S mutation in Rz | (8) |
| pRzC152S | pRz carrying C152S mutation in Rz | (8) |
| pRzC99,152S | pRz carrying C99S and C152S mutation in Rz | (8) |
| pRz1 | Rz1 cloned in pRE | (8) |
| $pRz1_{C29S}$ | pRz1 carrying C29S mutation in Rz1 | (8) |
| pRz/Rz1 | Rz cloned upstream of Rz1 in pRE | (8) |
| $pRz_{C99S}Rz1_{C29S}$ | $Rz1_{C29S}$ cloned downstream of $Rz_{C99S}$ | This study |
| $pRz_{C152S}Rz1_{C29S}$ | $Rz1_{C29S}$ cloned downstream of $Rz_{C152S}$ | This study |
| $pRz_{C99,152S}Rz1_{C29S}$ | $Rz1_{C29S}$ cloned downstream of $Rz_{C99,152S}$ | This study |
| **Primers** | **Sequences** | |
| DsbA_for | 5'- TAATCGGAGAGAGTAG- 3' | |
| DsbA_rev | 5' - GGGCTTTATGTAATTTAC - 3' | |
| DsbC_for | 5' - CGGCGACGAAGTTGTATCTGTTGT TTCACGCG - 3' | |
| DsbC_rev | 5' - GCGGGCGTGATGTCTGAAAAGAACGGGAAG - 3' | |

### *Standard DNA manipulations, PCR, site directed DNA mutagenesis and DNA sequencing*

Plasmids and oligonucleotide (primer) sequences are described in Table 2.1. Isolation of plasmid DNA, DNA amplification by PCR, DNA transformation, and DNA sequencing were performed as previously described (116). Oligonucleotides (primers) were obtained from Integrated DNA Technologies (Coralville, Ia.) and were used without further purification. Restriction and DNA-modifying enzymes were purchased from New England Biolabs (Ipswich, Ma.); all reactions using these enzymes were performed according to the manufacturer's instructions. Site-directed mutagenesis was performed using the QuikChange kit from Stratagene (La Jolla, Calif.) as described previously (38). For the PCR based confirmation *dsbA* and *dsbC* mutant loci, genomic DNA served as a template and the *dsbA* (dsbA_for and dsbA_rev) and *dsbC* (dsbC_for and dsbC_rev) flanking oligonucleotides served as primers. PCR products were resolved by agarose gel electrophoresis and bands of the expected mobility were excised and then sequenced as described above using the DsbA_for and DsbC_for oligonucleotides as primers respectively.

### *SDS-PAGE and Western blotting*

SDS-PAGE and Western blotting were performed as described previously (38, 117). SeeBlue Plus2 (Invitrogen) prestained standard was included as a molecular mass standard. Acetone washed TCA pellets were resuspended in 1X SDS-PAGE buffer (50

mM Tris-HCl, 2%SDS, 5% glycerol, pH 6.8) with or without a β-mercaptoethanol (100mM BME) as indicated and incubated at 100°C for 5 minutes. Protein sample loading volumes were normalized according to $A_{550}$ units at the time of TCA addition and each blot was repeated at least 3 times in order to demonstrate reproducibility of the observed Rz or Rz1specific band (monomer, oligomer, and degradation products). Samples destined for western blotting with anti-Rz serum were resolved by SDS-PAGE on 10% resolving/4% stacking Tris-tricine polyacrylamide gels. Samples destined for western blotting with anti-Rz1 serum were resolved by SDS-PAGE on 16.5% resolving/4% stacking Tris-tricine polyacrylamide gels. Proteins were transferred to PVDF membrane (Pall Life Sciences) using a Hoefer TE unit at 0.1 mA overnight at 4º C. Antibodies (Sigma Genosys) were generated in rabbits against the synthetic peptides CELADAKAENDALRDD, corresponding to the Rz residues 71-85 (modified with an N-terminal Cys for conjugation) and SQCVKPPPPPAWIMQ, corresponding to Rz1 residues 27-41 and used at a dilution of 1:1000. Similarly, in T4 the antibodies were generated against synthetic peptides CERENEKLRKDAKKA, corresponding to the PseT.3 residues 74-87 and CWLNDVKRYVHDQKT, corresponding to the PseT.2 residues 71-84 and used with similar dilution of 1:1000. The secondary antibody, goat-anti-rabbit-HRP (Thermo Scientific), was used at a dilution of 1:5000. Chemiluminescence was detected using a Bio-Rad XR Gel Doc system.

### *Manual alignment of the spanins*

In order to determine the conservation of periplasmic Cys residues across different 2CS spanins, 31 representative members of 24 different families were aligned using CLC workbench 7.5 (CLC Bio, Aarhus, Denmark) with or without maximum penalty for gaps. With the exception of closely related phages such as lambda, 933W, and P22, the sequence did not align according to the position of the Cys residues. Next, the sequences of periplasmic domains were manually aligned without any gaps, and the number of Cys residues in each i-spanin and o-spanin pair was manually counted. In order to illustrate the relative position of the Cys residues, a graphical representation was generated and drawn to a linear scale where each amino acid residue occupied the same space (see Fig. 2.9).

## Results

### *Shearing forces attendant to aeration complement the RzRz1 lysis defect in vivo*

Spanin function in batch aerobic culture is dispensable for lysis unless the medium is supplemented with 10 mM divalent cations; under these conditions, the terminal phenotype is spherical cells (6, 53). Recently, our laboratory used phase contrast microscopy to re-examine the process of phage lysis as well as the phenotype of spanin mutants in the absence of divalent cations. In the controlled conditions under the

microscope slide and free from shaking, spherical cells were stable for extended period of time, even without stabilization of the OM with divalent cations (10). These results led us to hypothesize that the shearing forces attendant to aerobic growth of a culture in a shaker flask was capable of complementing the *RzRz1* lysis defect. To test this hypothesis, liquid cultures of λ and λ$Rz_{am}Rz1_{am}$ lysogens were induced as previously described. However, in this case the induced lysogens were removed from agitation 10 minutes prior to the normal onset of lysis. The results clearly showed that lysis was defective in the $Rz_{am}Rz1_{am}$ culture, but not the wild type (Fig.2.2). As expected, the terminal phenotype of these lysis-defective spanin mutants was the accumulation of spherical cells which was observed under the microscope. Several physiological parameters were shown to affect the degree of the lysis defect, which presumably reflects the stability of the spherical cells. Delaying induction until later in exponential phase (Fig.2.2B) or lowering the temperature of the culture to 30°C 10 minutes prior to lysis (Fig.2.2C) resulted in a retardation of lysis. Conversely, increasing the shearing forces by using a baffled flask suppressed the lysis defect even in the presence of $Mg^{++}$ ions (10). Delaying induction to mid-logarithmic phase also produced a partial $Mg^{++}$-independent lysis-defective phenotype for the $Rz_{am}Rz1_{am}$ culture (Fig.2.2B, open black squares).

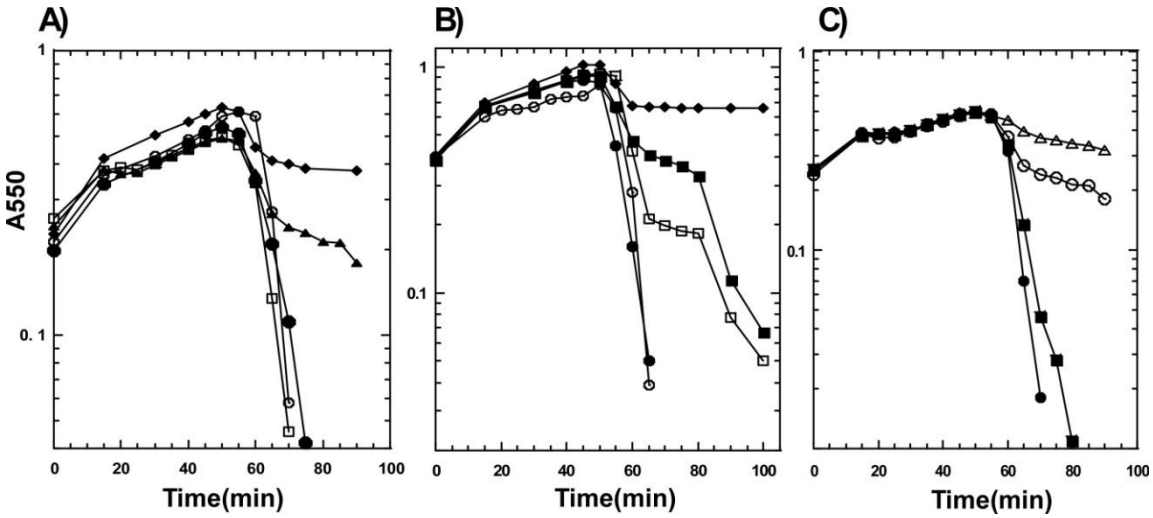



Figure 2.2. Saltatory λ*Rzam/Rz1am* lysis is an artifact. The effect of various physiological parameters on cell lysis was assessed following induction of λ prophages. (A) Shearing forces attendant on flask shaking complement the λ$Rz_{am}$*/Rz1am* lysis defect. The following prophage inductions were conducted with constant shaking: λ (open circles), λ$Rz_{am}$*/Rz1am* in LB plus10mM $Mg^{2+}$ (closed diamonds), and λ$Rz_{am}$*/Rz1am* (open squares). For the following prophage inductions, culture shaking was ceased at 45min after induction: λ (closed circles) and λ$Rz_{am}$*/Rz1am* (closed triangles). (B) Late-log phage induction leads to a $Mg^{2+}$ independent lysis defect. For the following prophage inductions, culture shaking was ceased at 45 min after induction: λ (closed circles) and λ$Rz_{am}$*/Rz1am* (closed squares). The following cultures were subjected to constant shaking: λ (open circles), λ$Rz_{am}$*/Rz1am* in LB plus 10mM Mg2+ (closed diamonds), and λ$Rz_{am}$*/Rz1am* (open squares). (C) Influence of temperature in lysis defect. For the following cultures the temperature was kept at 37˚C when culture shaking was ceased: λ (closed circles) and λ$Rz_{am}$*/Rz1am* (open circle). For following culture the temperature was dropped to 30˚C when culture shaking was ceased: λ (closed squares) and λ$Rz_{am}$*/Rz1am* (open traingles).

### *Synthetic lethality of $Rz_{C152S}$ $Rz1_{C29S}$*

The previous finding that Rz and Rz1 both formed covalent homodimers *in vivo* was striking in light of the molecular properties of the purified sRz and sRz1 (9, 56). Both of these proteins had been converted to thiol-free forms, by C99S and C152S changes in sRz and the C29S change in sRz1, in order to minimize the chance of oxidative damage during purification (9). Previously, gel filtration chromatography showed that purified sRz and sRz1 existed as a homodimers and monomers, respectively. This indicates that the periplasmic domain of Rz can homodimerize in solution in the absence of disulfide bond linkages between the monomers, which is consistent with its strong coiled-coil prediction (Fig.2.1). However, the periplasmic domain of Rz1 apparently requires the disulfide linkage to homodimerize in solution (9). Since $Rz1_{C29S}$ was a functional allele when tested in an $Rz^{+}$ context, we wondered if the presence of the covalently dimerized Rz was compensating for the Rz1 dimerization defect. To address this question, complementation tests were performed using all combinations of the *Rz* and *Rz1* Cys to Ser substitutions Fig.2.3A. The results were unambiguous, showing that either $Rz_{C152S}$ or $Rz_{C99,152S}$ but not $Rz_{C99S}$, is synthetically lethal with $Rz_{C29S}$, in terms of spanin function We conclude that for function, the interaction between the C-termini of Rz and Rz1 in the spanin complex requires a covalent homodimerizing linkage near the heterotypic junction; i.e., either between the distal segments of the two Rz molecules or between the two Rz1 periplasmic domains.

The corollary is that the disulfide bond at N-proximal Cys residue is irrelevant to function.

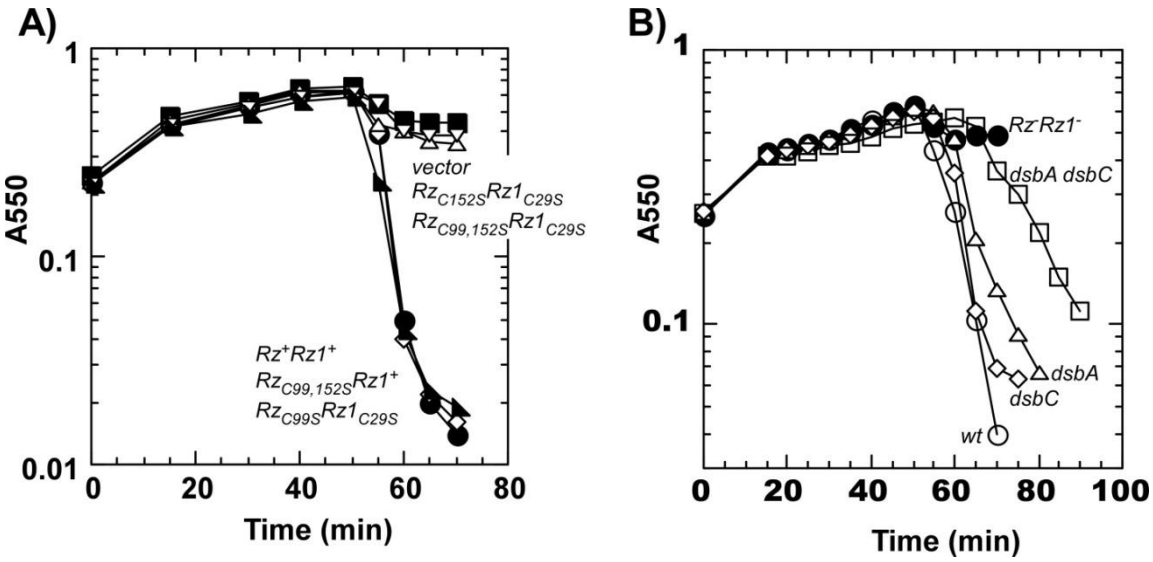



Figure 2.3. Lysis requires either Cys29 of Rz1 or Cys152 of Rz and can be achieved without DsbA or DsbC. The MC4100 lysogenic strains grown in LB supplemented with 10 mM $MgCl_2$ were induced at time=0 and growth monitored as A550. (A) MC4100 (λ900$Rz_{am}Rz1_{am}$) carrying the following plasmids : pRE, square (vector only control); pRzRz1, circle; p$Rz_{C99S}Rz1_{C29S}$, diamond; p$Rz_{C152S}Rz1_{C29S}$, upright triangle; p$Rz_{C99,C152S}Rz1_{C29S}$, downward triangle; $Rz_{C99,C152S}$Rz1, right triangle. (B) MC100 λ lysogens : RI89(λ900), open circle; RI89(λ900$Rz_{Q100am}$ $Rz1_{W38am}$), filled circle; RI90(λ900), triangle; RI179(λ900), diamond; and MR135(λ900), square.

***The role of the dsb system in the covalent homodimerization of the spanin subunits***

In bacteria, the formation of intramolecular disulfide bonds between consecutive Cys residues in secreted proteins is catalyzed by DsbA. When non-consecutive linkages are required for proper folding, the non-native disulfide bonds are removed by DsbC (118). We wondered if the *dsb* system was required for the formation of the intermolecular homodimerizing disulfide links in the spanin subunits. To address this question, λ lysogens deficient in DsbA, DsbC or both DsbA and DsbC production were induced and monitored for lysis (Fig.2.3B). In the *dsbA* and *dsbC* inductions, no significant effect was seen on the lysis kinetics. There was a delay of ~5 min in the onset of lysis in the *dsbA dsbC* double mutant and also a slightly reduced rate of lysis, as judged by the slope of the decline in turbidity. However, the double mutant also had a slower growth rate, as noted by others (119). Thus the slight delay in lysis timing observed in the *dsbAdsbC* background is difficult to interpret. To address how the allelic state of the *dsb* genes affected the morphological pathway at the cellular level, we employed the same video phase-contrast microscopy previously used to establish the essential role of the spanins in λ lysis (Fig.2.4, and Supporting Information available in Berry et al.(56)). As observed previously, the λ lysis in the $dsb^+$ background was rapid, with no deformation of the native rod morphology prior to the sudden localized disruption of the envelope and loss of phase contrast (Fig.2.4A). By contrast, the $\lambda Rz_{am}Rz1_{am}$ induction was non-lytic and resulted in the transition of cells from rod to spherical shape (Fig.2.4E). However, in all three *dsb* backgrounds (*dsbA*, *dsbC*, and

*dsbA dsbC*), there was a defect in the kinetics and morphological pathway for lysis at the single cell level. Compared to the wild type host, cell disruption occurred 10-15 sec later in *dsbA* and ~900 sec later in *dsbC* (Fig.2.4B and C respectively). Interestingly, in the *dsbA dsbC* background, the kinetic defect in lysis kinetics was indistinguishable to the slight defect observed for *dsbA* (Fig.2.4D).

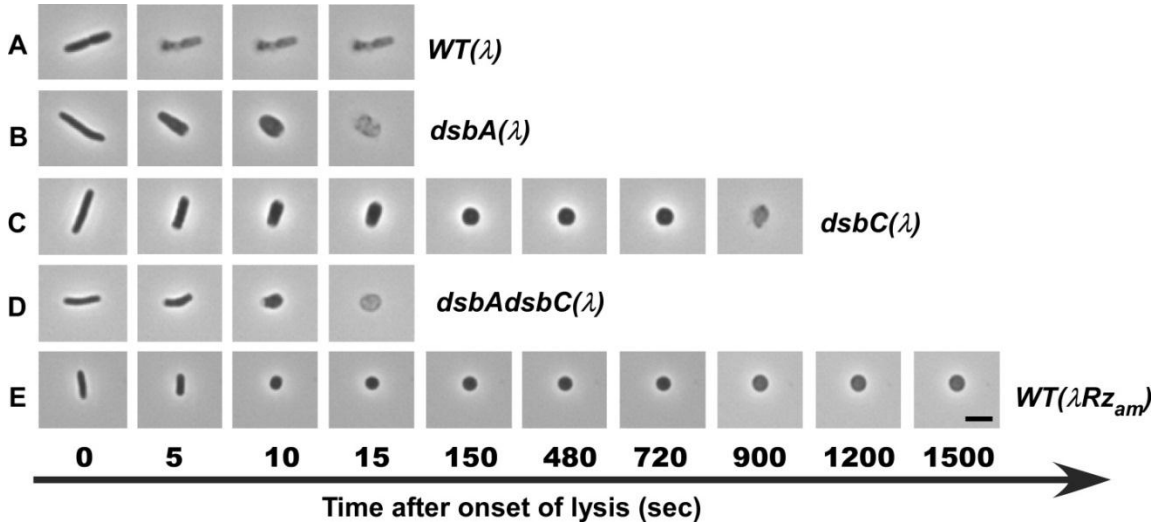



Figure 2.4. *dsb* mutant hosts exhibit a kinetic defect in lysis characterized by a transient spherical cell phenotype. λ900 lysogens of MC1000 marked with the indicated *dsb* alleles were thermally induced and single cells subsequently monitored by time-lapse phase contrast microscopy. Time 0 marks the first frame available prior to the onset of a visible deformation of cell shape or lysis. Each cell was monitored until lysis occurred or a total of 20 minutes had lapsed. The scale bar is 5μm. The lysogen in panel E carried the prophage λ900$Rz_{am}$.

To address this apparent epistatic effect of the *dsbA* mutation in spanin function, Western blot analysis of the Rz and Rz1 proteins was performed under reducing and non-reducing conditions in the various *dsb* backgrounds (Fig.2.5). The amount of Rz covalent dimer was somewhat diminished in the *dsbA* background (Fig.2.5A compare lanes 1 and 3). Since DsbC acts either as an isomerase, with no net increase in disulfide bonds, or as a reductase (120), this result indicates that the intermolecular Rz disulfide bond can form spontaneously. Moreover, the amount of Rz covalent dimer was even more drastically reduced in both *dsbC* and the *dsbA dsbC* backgrounds (compare lane 1 with lanes 4 and 5). This is consistent with the severely retarded lysis kinetics in Fig.2.4, panels C and E. It should be noted that the long delay in the *dsbC* backgrounds is not observed in the bulk lysis experiments (Fig.2.3B), suggesting that the severely reduced amount of functional spanin is still sufficient to suppress the stabilizing effect of the 10 mM $MgCl_2$ supplement that stabilizes the spherical cells against the shearing forces in $Rz_{am}Rz1_{am}$ conditions. In any case, these results again show that covalent Rz dimers can form spontaneously as the covalent dimer is evident even in the *dsbA dsbC* background (Fig.2.5A, lane 5). In the *dsbC* background, an Rz monomer species accumulated that had a reproducibly increased mobility compared to both the reduced form and to the $Rz_{C99,152S}$ mutant protein (Fig.2.5A; compare lanes 4, 6, 7). Presumably this is Rz with the Cys99 and Cys152 linked in an intramolecular disulfide bond, which would severely constrain the ability of Rz to span the periplasm or support the formation of the coiled-coil bundles observed in vitro. Support for this interpretation can be gleaned from analysis of the Rz species of the Cys-substitution alleles in the *dsbC* host, where the

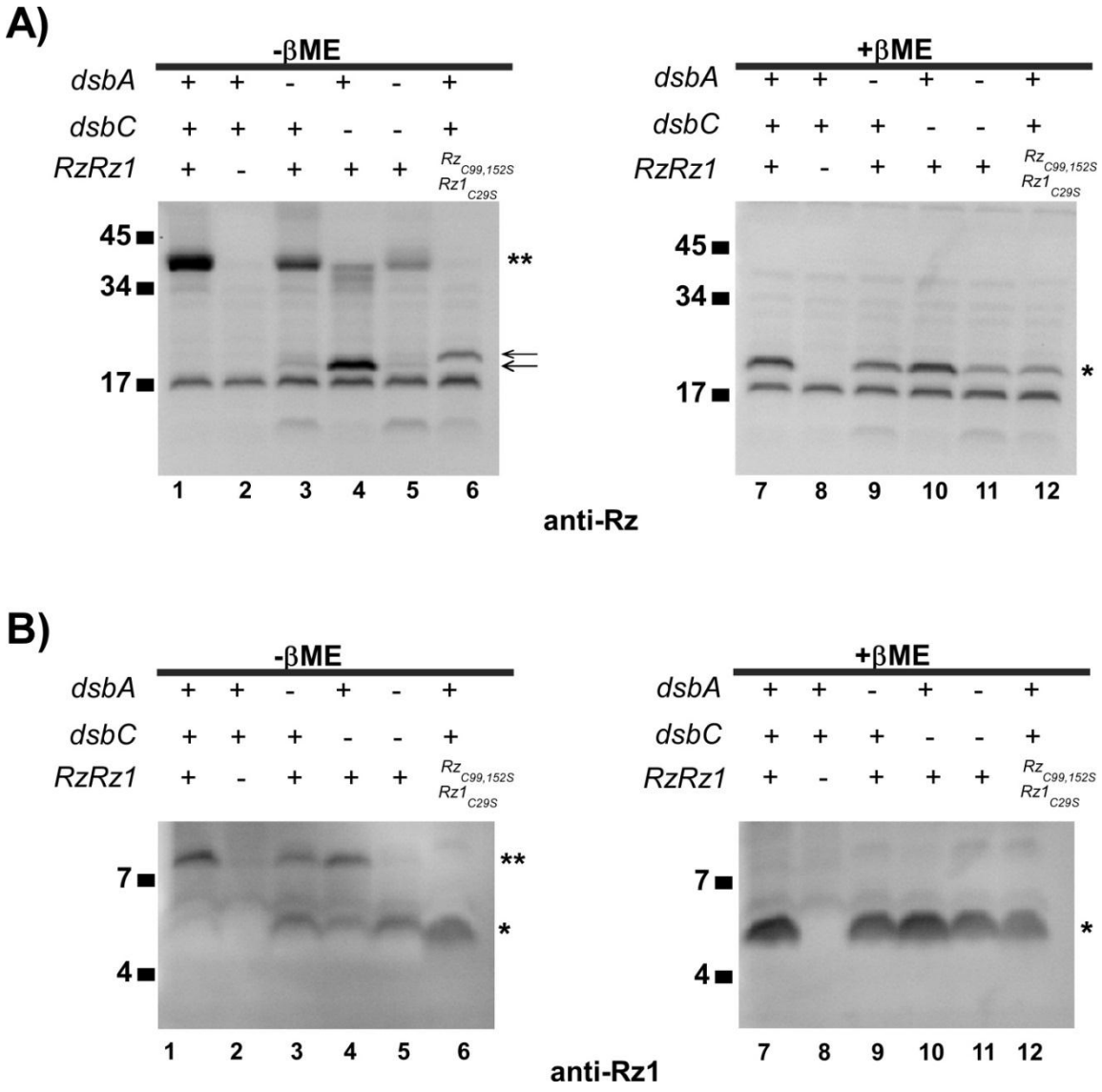



Figure 2.5. Dsb system mediates intermolecular disulfide bond formation. Samples from MC1000 lysogens were prepared by TCA precipitation at 50 minutes post induction and subjected to SDS-PAGE and western blotting using anti-Rz or anti-Rz1 antibody, in the presence or absence of β-mercaptoethanol (β-ME) as indicated. The allelic state of *dsbA*, *dsbC*, *Rz*, and *Rz1* is indicated. The monomer and dimer species are indicated by single and double asterisk sign, respectively. In lanes 6 and 12 of both panels, samples were from cysteine-less alleles of *Rz* and *Rz1*. Arrows indicate monomer species of Rz with (lower arrow) and without (upper arrow) an intramolecular disulfide bond.

intramolecular bond cannot form and covalent dimers are formed instead (Fig.2.5B). Thus both Rz periplasmic Cys residues are essential to formation of the monomeric Rz with increased mobility. We propose that the accumulation of this looped, oligomerization-incompetent species underlies the more drastic morphological transition phenotype of the *dsbC* host (see Discussion). The *dsb* backgrounds also affected the covalent dimerization of Rz1, with the dimer species accumulating to reduced levels in both *dsbA* and *dsbC* backgrounds (Fig.2.5C). In the double mutant, the covalent Rz1 dimer is undetectable. In view of the predicted lack of secondary structure for Rz1, we considered it likely that the DsbA-independent formation of the Rz1-Rz1 disulfide linkage would be templated by complex formation with Rz. To address this possibility, we assayed the formation of Rz1 covalent dimers in the presence and absence of Rz in both *dsbA* and *dsbC* backgrounds (Fig.2.6A). The results showed that, in the absence of DsbA, Rz1-Rz1 intermolecular disulfide bond formation required Rz (lanes 1, 2). It seemed likely that the converse would not be true for Rz; i.e., that the coiled-coil structure of Rz would provide the steric environment for DsbA-independent intermolecular disulfide bond formation. To address this idea, we assayed Rz-Rz disulfide bond formation in the presence and absence of Rz1 in the *dsb* backgrounds (Fig.2.6B). Surprisingly, the results showed that DsbA-independent formation of the covalent Rz dimer was nearly as dependent on the presence of Rz1 as the Rz1-Rz1 linkage was dependent on Rz (lanes 1, 2). These results will be considered below in the context of an overall model for covalent homo-dimerization of the spanin subunits (see Discussion).

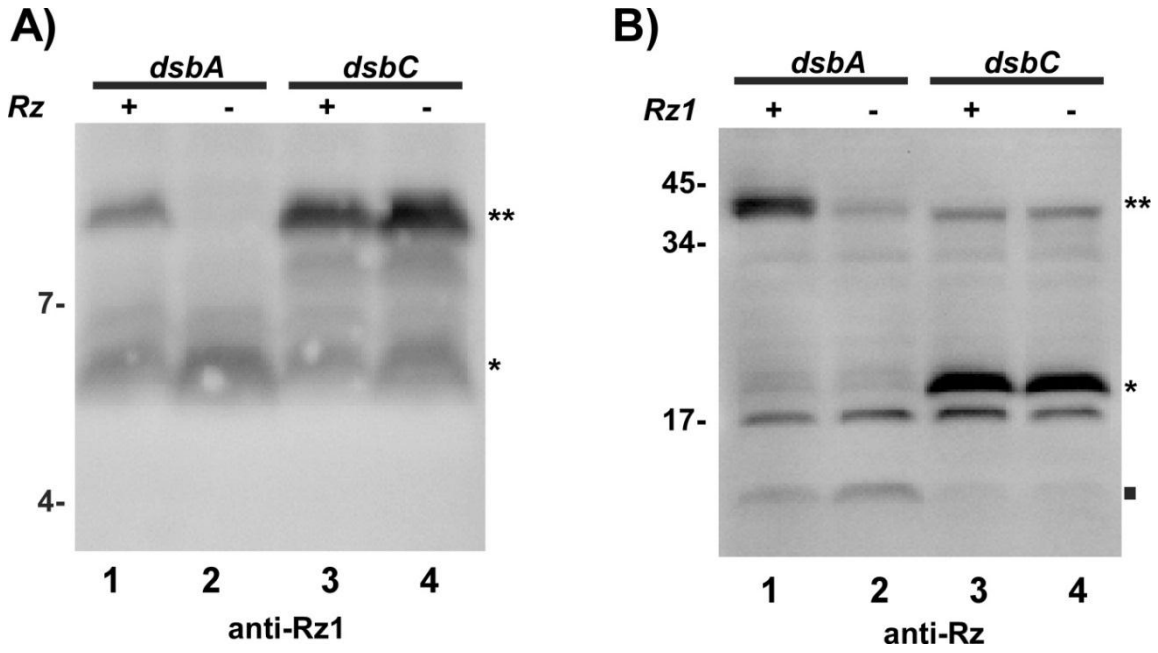



Figure 2.6. Spontaneous dimerization of either spanin subunit is supported by its cognate partner. Samples from induced MC1000 lysogens were prepared and analyzed as in Figure 2.4. In panel A, Rz1 monomer and dimer species in *dsbA* and *dsbC* hosts are indicated by asterisk and double asterisk, respectively. In panel B, Rz monomer and dimer species are similarly labeled. The black square indicates a degradation product previously observed to accumulate in the absence of Rz1(8).

***Kinetic defects associated with Cys substitutions in Rz and Rz1***

The defects in the kinetics and morphological pathway associated with the *dsb* mutant hosts suggested to us that similar defects might be detectable at the cellular level in the various Cys substitution mutants. Video microscopy confirmed this notion (Fig.2.7). As expected, the terminal phenotype of the lysis-defective *RzRz1* alleles

lacking both the distal Cys of Rz and the sole Cys of Rz1 was stable spherical cells (Fig.2.7B and D). However, even the three lytic allelic combinations (*Rz*$_{C99S,C152S}$*Rz1*, *Rz*$_{C99S}$*Rz1* and *Rz*$_{C99S}$*Rz1*$_{C29S}$) also showed evidence of an altered lysis pathway, in that a partially spherical intermediate lasting 10-20 sec was formed (panels A,C,E), instead of the abrupt lysis of the rod-shaped cell that characterizes the wild type situation (Fig.2.4A, also Berry et al., (10)).

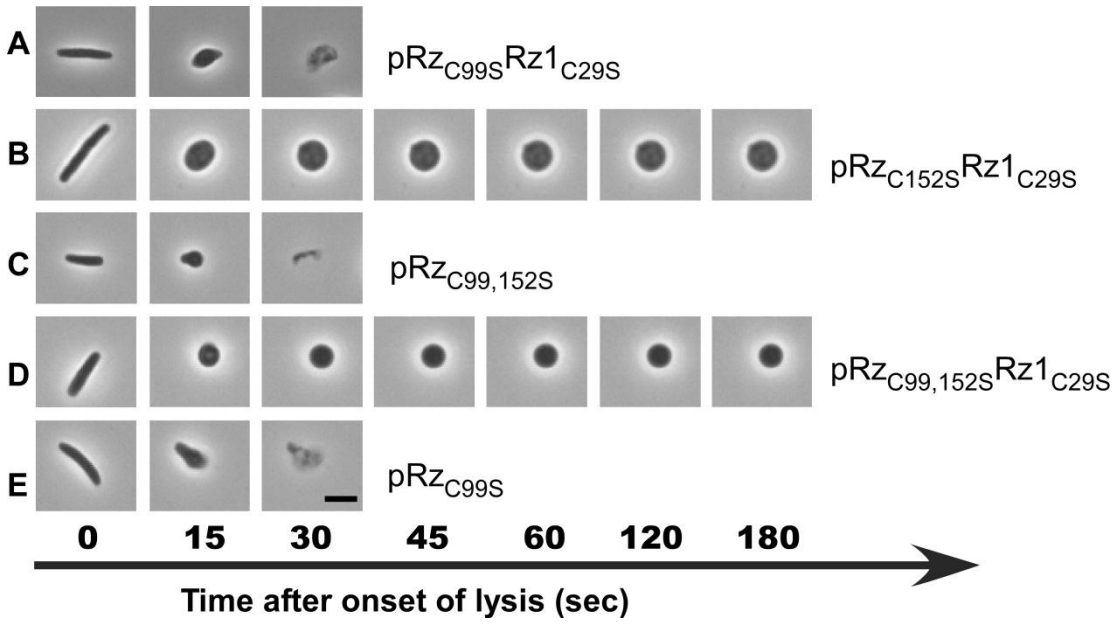



Figure 2.7. Lysis phenotypes of cysteine-substitution alleles of the spanin genes. MC4100(λ900*Rz*$_{am}$ *Rz1*$_{am}$ ) lysogens carrying the indicated *Rz* and *Rz1* alleles on a trans-activating plasmid were thermally induced and subsequently monitored by time-lapse phase contrast microscopy as described in Fig. 2.3. Each cell was monitored until lysis occurred or was transformed into a spherical form (see Supplementary Video in Berry et al. (56)). The scale bar is 5μm.

***Kinetic defect associated with the DTT treatment during lysis***

Additional evidence for the requirement of intermolecular disulfide bonds was obtained using a reducing agent on induced cells. In this experiment, induced λ lysogens were treated with 15mM of DTT 5 min prior to lysis and observed under a phase contrast microscope. The lysis phenotype of the DTT-treated cells resembled that of cells which had defective spanins (Fig. 2.8), suggesting that the essential disulfide bonds for spanin function were being reduced. However, the round cells were not stable enough compared to the amber or Cys to Ser substitution mutants (compare Fig.2.4 and 2.7). The simplest interpretation is that disulfide linkages in some fraction of spanin complexes were not reduced to completely block lysis.

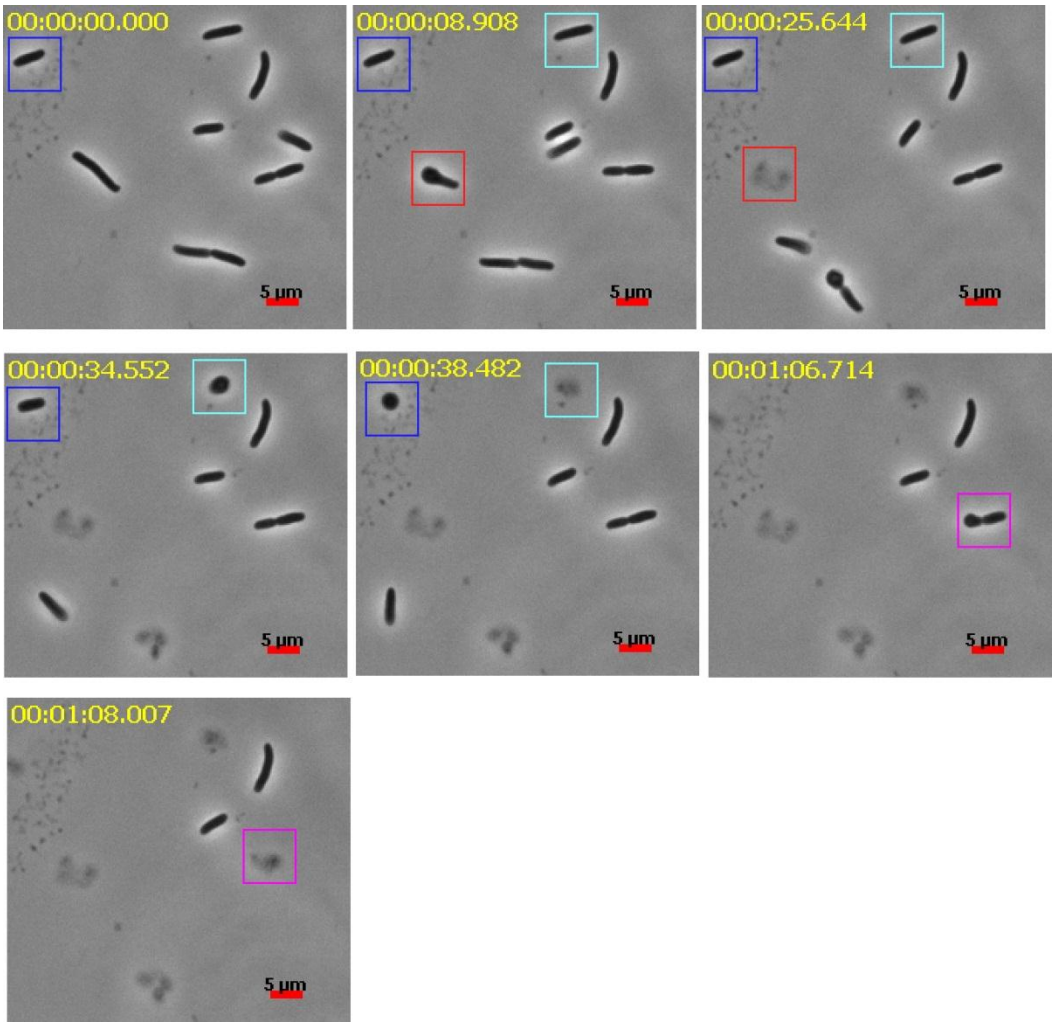



Figure 2.8 The wild type λ spanins are sensitive to the DTT-treatment. MC4100 (λ900) lysogens with wild type *Rz* and *Rz1*alleles were thermally induced and 15mM DTT was added at 45 min post induction. Beginning at 55min, the induced culture was monitored by time-lapse phase contrast microscopy as described earlier. Compared to wild type lysis phenotype in Fig.2.3A, most of the cells within a frame transformed into a spherical form either completely or partially before lysis (follow marked colored boxes). The scale bar is 5µm.

***Cysteine requirement is conserved among 2CS spanins***

The initial bioinformatics survey of spanin homologs was reported in 2007 (7). Since then, the number of sequenced phage genomes has greatly increased. A comprehensive study of spanins started through collaboration with Dr. Joel Berry and Rohit Kongari of our lab. Using membrane localization signals and gene arrangement as the primary search criteria for determining new spanins (7), we have been able to identify 420 spanin homologs among 526 phages of Gram-negative hosts. Thirty seven of the five hundred and twenty six phages possess u- spanins, while the rest contain 2CS spanins. There were 106 phages in which spanins could not be identified; these phages mostly had cyanobacterial hosts. We then focused our analysis on the number of cysteines in the periplasmic domains of identified spanins. Using this criterion, we were able to determine that cysteines are widely conserved in 2CS spanin; all but 12 spanin pairs contained cysteines in the periplasmic domain. Table 2.2 shows the prevalence of cysteine residues in the spanins of 371 phages, grouped on the basis of spanin gene architecture. Out of 371 spanin pairs, ~20% had at least one cysteine in the i-spanin, 29% had at least one cysteine in o-spanin, and remaining 51% had at least one cysteine in both i-spanin and o-spanin (Table2.2, column 2, 3, & 4). Moreover, when the cysteine residue was present only in the i-spanin, the i-spanin and o-spanin was typically embedded gene architecture. In contrast, when the cysteine residue was present only in the o-spanin, it was equally distributed between overlapped and separated architecture. Interestingly, when the cysteines were present in both components, the equal distribution

Table 2.2. The prevalence of the cysteine residues in 2CS spanins.
Column 2 shows the number of phages that have periplasmic cysteine residue in i-spanin only. Column 3 shows the number of phages that have periplasmic cysteine residue in o-spanin only. Column 4 shows the number of phages that have periplasmic cysteine residue in both spanin subunits. The total phages with embedded, overlapped, and separated spanins gene arrangement are indicated in the last column.

| Gene arrangement | Cys in i-spanin | Cys in o-spanin | Cys in both | Total |
|---|---|---|---|---|
| **Embedded** | **60** | **1** | **83** | **144** |
| **Overlapped** | **4** | **53** | **89** | **146** |
| **Separated** | **10** | **52** | **19** | **81** |
| **Total** | **74** | **106** | **191** | **371** |

shifted from separated and overlapped to overlapped and embedded. Overall, when we examined these spanin pairs more closely, we found that in all cases cysteines were located towards the C-terminus if they were exclusively present in the i-spanin or o-spanin alone. Among 317 total spanin pairs, we were able to find only one case where embedded spanin architecture had cysteine in the o-spanin alone; usually it is present in the i-spanin or both (see column 2 in Table 2.2). Conversely, we were able to find few cases where overlapped and separated architecture had cysteine in the i-spanin alone; usually in these architectures it is present in o-spanin or both. This suggests that overlapped and separated are more closely evolutionarily related. In 2007, the spanins of 127 phages were organized into 37 families, among which 28 of the phage families had a

single member (see Table 2 of Summer et al. (7)). We randomly chose 31 phages with 2CS spanin from 24 different families to construct a representative diagram shown in Fig.2.9. In spite of the diversity among families that are not detectably related by sequence homology, the representative diagram shown in the Fig.2.9 suggests that the cysteine residues are primarily located near the C-terminal end of both i-spanins and o-spanins. This seems to correlate with the data represented in Table 2.2.

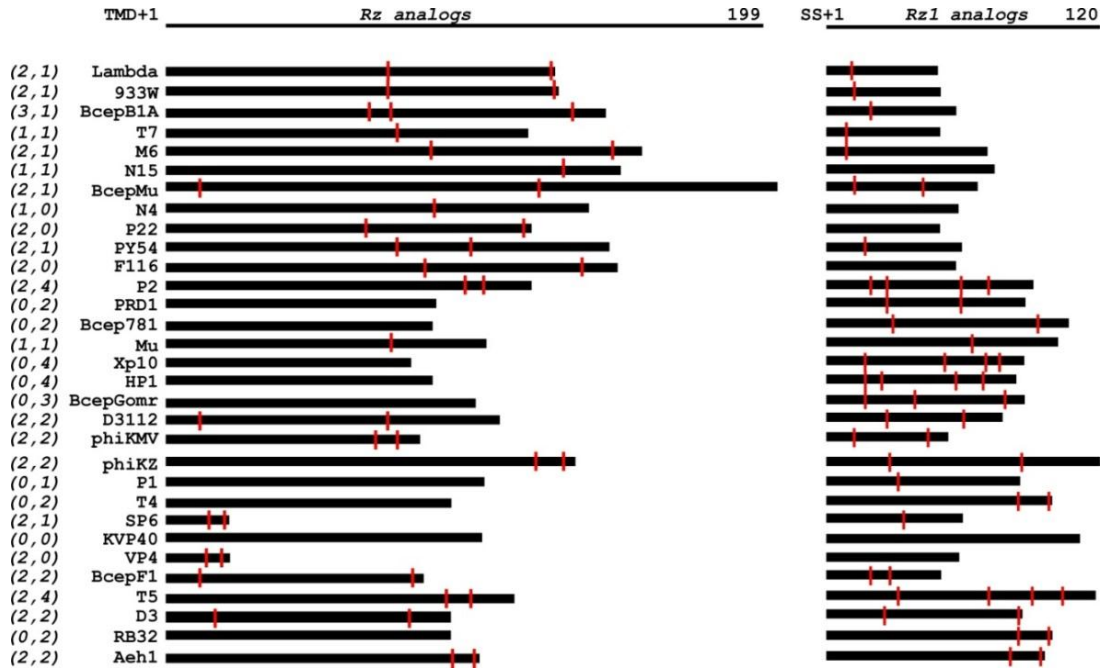



Figure 2.9. 2CS spanins require at least one cysteine in the periplasmic domain. The representative diagram illustrates the number of cysteine residues (parenthesis in the left) in the Rz and Rz1 equivalents and their relative position in the periplasmic domain (red mark). The length of the periplasmic domains of i-spanin and o-spanin equivalents is drawn to scale and indicated (top) by the last amino acid residue of the largest domain. *KVP40 is an exception that lacks a periplasmic cysteine residue*.

In contrast, none of the 37 1CS spanins had a cysteine residue in their periplasmic domain. We suggest that a minimum of one cysteine is the universal requirement for all the 2CS, exemplified by λ. To test our hypothesis that a 2CS spanin requires at least one cysteine residue in the i-spanin or o-spanin, we selected spanins of T4 phage. The spanins in T4 and λ do not have sequence similarity, although their predicted structures are more or less conserved (compare Fig.2.10A and Fig.2.1A). Furthermore, in T4, the i-spanin (PseT.3) lacks a cysteine residue whereas the o-spanin (PseT.2) has two cysteines in its periplasmic domain, at positions 87 and 98 (Fig. 2.10A). A complementation assay indicated both *pseT.2* and *pseT.3* are required for lysis in a $\lambda SRRz_{am}Rz1_{am}$ background (Fig. 2.10B). To assess the role of cysteines at positions 87 and 98 of *pseT.2*, single and double serine substitution alleles were generated, all of which were lysis-defective (Fig. 2.10B). Thus, in the case of T4, both cysteines in the o-spanin were essential for lysis. Contrary to the λ spanins, the i-spanin of T4 did not form a covalent-linked dimer *in vivo* due to the lack of cysteine residues in its periplasmic domain (Fig.2.10C). Also, most of the o-spanin of T4 existed as monomers despite the presence of two cysteine residues in the distal portion of the protein (Fig.2.10C). This indicates that the majority of the T4 o-spanin might exist as a small loop conformation due to an intramolecular disulfide linkage between two distal cysteines. This may be the functional form of the spanin, since the two distal cysteine residues are present in the o-spanin rather than i-spanin. This is unlike spanins in λ where an intramolecular disulfide in the i-spanin would create a large hairpin preventing from spanning the periplasm. What caught our attention was the presence of a small fraction of covalently linked dimer of o-spanin that is detectable

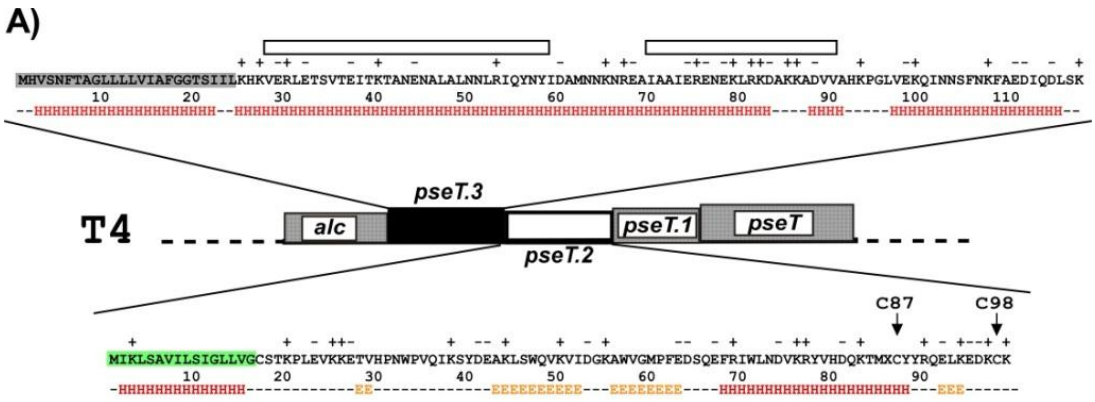
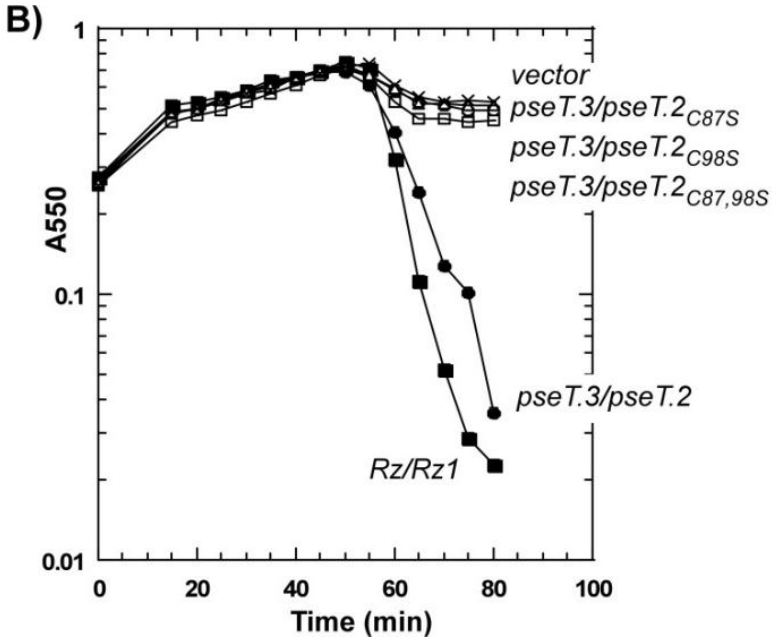
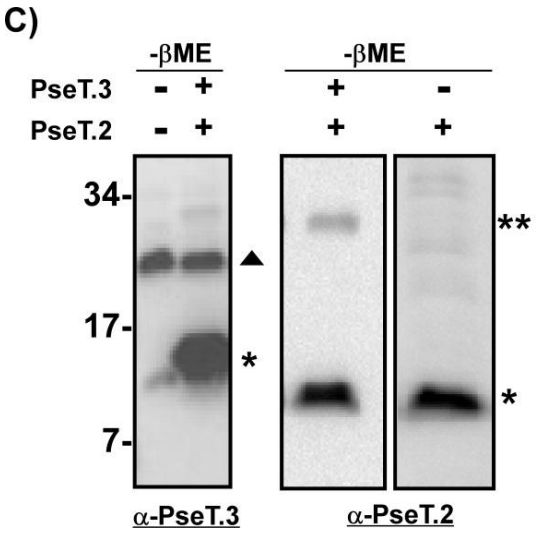
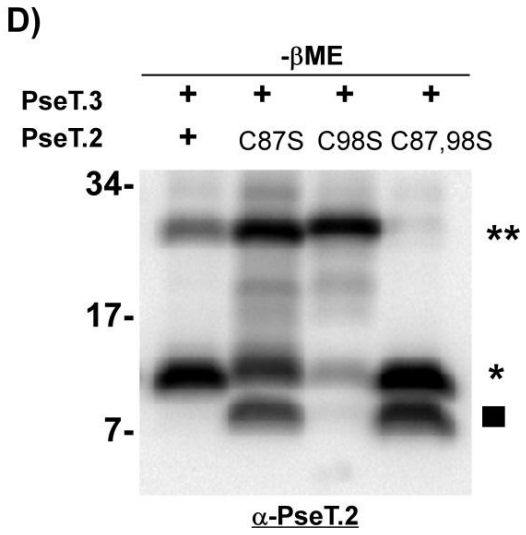



Figure 2.10 The requirement of two cysteines in T4 spanin function. **(A)** Unlike embedded genes in λ, the spanin genes *pseT.3* (i-spanin) and *pseT.2* (o-spanin) in T4 are separated. Additionally, these genes are not located near the T4 holin (*t*) and endolysin (*e*), which are essential for lysis. The inset shows the predicted primary structures of PseT.3 and PseT.2 with their TMD and signal sequence highlighted in gray and light green, respectively. *PseT.3* (black) encodes the 117 aa inner membrane subunit (i-spanin) and *PseT.2* encodes 83aa outer membrane spanin subunit (o-spanin). The o-spanin with the C-terminal helix in T4 is comparatively larger than o-spanin in λ which lacks helical structure. The i-spanin of T4 has a coiled-coil structure (open rectangles), similar to that of the λ i-spanin. Secondary structure predictions were made as described in Fig.2.1. The i-spanin of T4 lacks a periplasmic cysteine residue whereas o-spanin has two cysteine residues at position 87 and 98 (arrows). **(B)** MC4100 lysogenic strains grown in LB supplemented with 10 mM $MgCl_2$ were induced at time=0 and growth was monitored at A550. MC4100 (λ900$Rz_{am}Rz1_{am}$) carry the following plasmids: pRE (-X-) ; pRzRz1 (-■-); pPseT.3PseT.2 (-●-); pPseT.3PseT.$2_{C87S}$ (-□-); pPseT.3PseT.$2_{C98S}$ (-Δ-); pPseT.3PseT.$2_{C87,982S}$ (-○-). **(C&D)** MC4100 (λ900$Rz_{am}Rz1_{am}$) lysogens carrying the indicated allele were prepared and analyzed in the absence of β-mercaptoethanol (βME) as in Fig.2.4. Antibodies against T4 spanins are indicated at the bottom of each panel. The location of monomeric (single asterisk) and dimeric (double asterisk) species of PseT.3 and PseT.2 are shown. Additionally, putative degradation products (square) are indicated to the right of each blot. Filled triangles indicate a background band. The alleles are indicated above each lane. Molecular markers in kDa are indicated to the left.

Figure 2.10. Continued



only in the presence of i-spanin (Fig.2.10 C and D). The dimer fraction was highly enhanced in the single Cys to Ser substitution allele of o-spanin (Fig. 2.10 D). The enhanced dimer due to absence of either cysteine, however, represents the non-functional complex according to the genetic study shown in Fig. 2.10B. It is apparent in Fig.2.10D that the absence of cysteine at position C87 results in an increased amount of degradation product, indicating its susceptibility to degradation. However, the absence of cysteine at position 98 is relatively more stable. It remains unclear how these two cysteines in the o-spanin play their essential role in spanin function and requires further investigation.

## Discussion

### *Rz and Rz1 are essential for λ lysis*

Many genetic, physiological and environmental perturbations of the cell envelope can lead to lysis of the bacterial cell. For example, when cell wall antibiotics block the incorporation of cell wall precursors into the PG in a growing *E. coli* cell, the cell ultimately lyses when it attempts to progress through septation, despite there being no direct effects of the antibiotic on the other two components of the envelope, i.e. the cytoplasmic membrane and outer membrane. Phage lysis is qualitatively different from the lysis derived from such derangements, principally in that it must be effected in a temporally-regulated manner. Compelling arguments have been presented that the

capacity to adjust the timing of lysis, and thus the length of the infection cycle, is an important fitness factor, especially in terms of optimizing phage fecundity in the face of changing availability of the host bacterium (121-124). Moreover, it is obvious that the lysis system should have minimal effect on the intracellular rate of accumulation of progeny virions until the scheduled time of lysis, after which the most rapid and efficient dispersement of the progeny virions should be effected (125). These considerations have led to a wide-spread acceptance of the key role of holins and endolysins in phage lysis, with the former controlling the timing of lysis and the permeabilization of the cytoplasmic membrane, which immediately leads to attack on the PG by the latter. However, it was only recently that the disruption of the OM would also be important for rapid and efficient host lysis has begun to be appreciated. Initially, the *RzRz1* genes of λ were designated as auxiliary lysis genes based on the fact that the *RzRz1* null phenotype was a cation-dependent lysis defect, suggesting that the spanin complex, the components of which are encoded by *RzRz1*, was required only when the OM was artificially stabilized (6). Here we have shown that in fact λ lysis, on a time scale useful to the phage, actually requires functional *RzRz1* genes, even without artificial stabilization of the OM. Moreover, evidence is provided showing the terminal phenotype of *RzRz1* null mutants is the formation of a spherical cell sufficiently fragile that the shearing forces derived from aerobic growth in a shaker flask artifactually complement the lysis defect (Fig. 2.2A and B), which led to the inaccurate designation of *RzRz1* as auxiliary lysis genes (5, 69, 126, 127).

The cation-dependent spherical cell phenotype has also been demonstrated for mutants in the 2CS spanin genes of diverse phages, including the coliphage P2 (55) and the cystovirus PRD1 (69). Although similar experiments described in this chapter needs to be done for a few of these systems, the nearly universal presence of either 2CS or 1CS spanin genes in the genomes of phages that grow on Gram-negative hosts (7) makes it likely that the lessons derived from λ lysis will be generally applicable. This conclusion leads to a model that lysis of the Gram-negative host is a three-step process, mediated by holins, endolysins, and spanins, subverting, in turn, the IM, the cell wall and the OM, respectively. In this three-step process, each step depends on the previous step, although there is no evidence for any direct physical interaction between the holins, endolysins or spanins. In fact, the diversity and interchangeability of holins, endolysins and spanins indicate that direct interaction is unlikely. For example, the λ holin and endolysin functions equally well with the phage P2 LysBC 2CS, which are encoded by an overlapped spanin gene architecture and have no sequence similarity with the lambda proteins (55).

***A covalent dimer of either Rz or Rz1 is required for spanin function and phage lysis***

Initial findings were that both Rz and Rz1 form homodimers linked by intermolecular disulfide bonds involving all three periplasmic Cys residues, however, not essential for spanin function. Here we have reported that that the biological function of the spanin complex requires either the $Rz_{Cys152}$-$Rz_{Cys152}$ or the $Rz1_{Cys29}$-$Rz1_{Cys29}$

linkage. There have been very few documented instances of intermolecular disulfide bonds in prokaryotes, all of which were found in proteins of the Gram-negative envelope (128-131). In only one case, with the VirB7 and VirB9 proteins of the *A. tumefaciens* T-complex transport apparatus, have such linkages been shown to be functionally significant. VirB7 forms disulfide-linked homodimers and also heterodimers with VirB9, both of which are important for the assembly of the transport complex. TraN, the VirB9 equivalent in the F-factor transfer complex, is an OM protein involved in mating pair formation has a Cys-rich C-terminal domain that also seems to be involved in intermolecular disulfide linkages, although neither the linkage partner nor which of the more than 30 conserved Cys residues in TraN are involved has been addressed (130). It thus may be significant that both of these cases involve multiprotein-complexes that connect the IM and OM. From this perspective, it is puzzling that the linkages are homotypic and homo-dimeric between both the i-spanins and o-spanins, rather than a heterodimeric linkage between Rz and Rz1, which could easily be rationalized as reflecting vertical forces operating on the spanin complex during its OM disrupting-function. Why is a covalent linkage required when the i-spanin:o-spanin interaction that defines the ability of the spanin complex to span the periplasm and effect OM disruption is a non-covalent C-terminal-C-terminal interaction? Even more puzzling is the finding that the requirement for the homotypic intermolecular linkage can be satisfied either by an Cys29-Cys29 linkage within an Rz1 dimer or by a Cys152-Cys152 linkage within an Rz dimer. Furthermore, our bioinformatic analysis and the genetic and molecular evidence of T4 spanin suggest that Cys residue near the C-terminal end of either spanin

subunit is fairly conserved among almost all 2CS spanin. We have previously shown that Cys-less versions of the periplasmic domains of Rz and Rz1 interact *in vitro* to form hetero-multimeric rod-like complexes of ~25 nm in length, exactly the dimension of the periplasmic space in *E. coli* (9); moreover, this complex formation is associated with large increase in the total alpha-helical content of the two proteins. The experiments were done with the Cys-less proteins because genetic analysis had shown that none of the Cys residues in either Rz or Rz1 was essential. In view of the results reported here, it will be important to re-visit Rz-Rz1 complex formation *in vitro* using Rz with an intact Cys152 or Rz1 with an intact Cys29. Although Cys-less Rz is dimeric, purified Cys-less Rz1 is monomeric in solution. Repeating the complex formation with the Cys-containing proteins may result in different product complexes; it is worth noting that the rod-like complexes exhibited a rather large variance in its long dimension, a feature that might be due to the lack of a covalent intermolecular linkage to stabilize the subunits in some sort of register. In any case, the current model for how spanins effect OM disruption does not provide any insight about the requirement for having at least one of the two homodimerizing linkages. The idea is that Rz and Rz1 form complexes by C-terminal-C-terminal interactions; these complexes, spanning the periplasm, are constrained to an interstitial space in the murein until liberated by the degradation of the peptidoglycan by the phage endolysin. The liberated Rz-Rz1 complexes are then able to oligomerize laterally, leading to a conformational change that causes fusion of the IM and OM. The simplest notion is that during the OM disruption pathway, lateral forces that would disrupt the homotypic interactions in the complex dominate at some point. It

is also worth noting that loss of even the non-essential disulfide linkage, the Rz Cys99-Cys99, results in slightly slower lysis with a short-lived spherical intermediate, indicating that ideal spanin function requires all three disulfide linkages.

***How are the intermolecular disulfides bonds formed?***

Only in the case of the conjugation protein TraN, which has more than 30 conserved Cys residues, was the *dsb* system implicated in the formation of intermolecular links, although neither the specific Cys residues nor the identity of the covalent partner was determined. TraN was found to be monomeric and unstable in a *dsbA* host, but not in a *dsbC* background (130). Here we have documented the involvement of both DsbA and DsbC in the covalent homodimerization of both Rz and Rz1. Although neither Dsb is required for lysis, there is a substantial defect in the *dsbC* background, in that the lysis pathway of the induced cell goes through an extended spherical period of ~15 min, which, in addition to the normal lysis time of ~50 min, would represent a major retardation of virion release. In contrast, the *dsbA* mutation has a much more subtle morphological phenotype in which the spherical form, indicating reduced function of the spanin complex, persists only for ~10-15 sec; moreover, the mild *dsbA* phenotype is epistatic to *dsbC*. This can be rationalized by the fact that in the absence of *dsbC*, the Rz protein appears to accumulate in a form with an intramolecular disulfide bond between Cys99 and the penultimate residue, Cys152. Such a linkage would certainly abrogate the ability of an Rz-Rz homodimer to span the bulk of the

periplasm and interact with Rz1. This suggests that the normal pathway leading to the Rz-Rz linkage is that DsbA first catalyzes formation of the intramolecular linkage as the Rz protein undergoes export through the translocon. The misfolded, oxidized Rz would then be reduced by DsbC and then undergo either DsbA-catalyzed or spontaneous formation of the Rz-Rz disulfide bonds. In the absence of DsbC, only a severely reduced amount of the covalent Rz-Rz dimer can form, either from the action of a minor reductase, possibly DsbG (132), or from spontaneous disulfide bond isomerization between two interacting oxidized Rz molecules (Fig. 2.4A, lanes 4 & 5). In the absence of DsbA, the covalent Rz-Rz dimer can form spontaneously, presumably because the periplasmic domain is predominantly helical with dominant coiled-coil character; in such a scenario, it is easy to imagine that the Cys99 and Cys152 residues would end up in close apposition in the coiled coil. However, this spontaneous process requires the templating of Rz upon Rz1 in the OM (Fig.2.5B, compare lane 1 and 2). Not surprisingly, Rz1 dimerization exhibits no dependence on DsbC, since its sole Cys cannot form an inappropriate intramolecular disulfide (Fig.2.5A, lane 4). However, as for Rz, the presence of the cognate spanin subunit is required for spontaneous Rz1-Rz1 intermolecular disulfide bond formation (Fig.2.5A, compare lane 1and 2). The simplest interpretation is that the normal pathway is for both proteins to undergo catalyzed covalent dimerization (DsbA and DsbC mediated for Rz; DsbA for Rz1), and then form the dimer-dimer spanin complex (Fig.2.11). Alternatively, if the complex forms between non-covalent dimers, the respective Cys residues find themselves in close

apposition. This allows the formation of intermolecular linkages promoted by the oxidative environment of the periplasm.

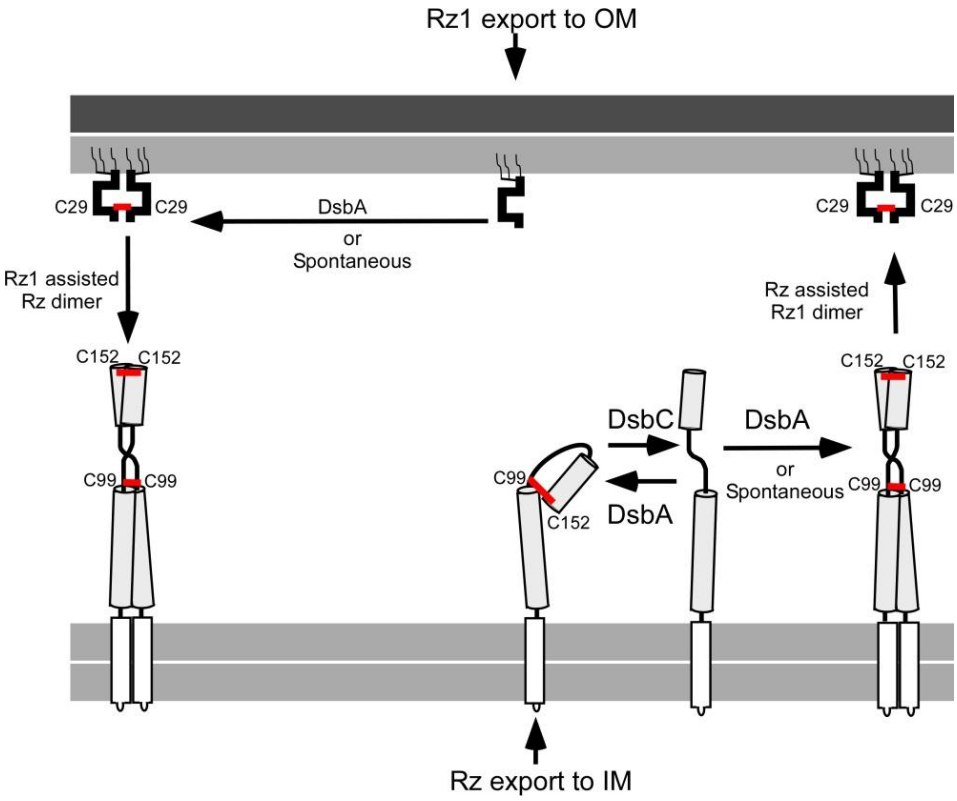



Figure 2.11. Model for intermolecular disulfide bond formation in the spanin complex. Red bars indicate disulfide bonds. Normally, DsbA catalyzes the formation of a non-productive intramolecular disulfide between Cys99 and Cys152 of Rz. DsbC serves to reduce this intramolecular linkage, allowing DsbA to catalyze the formation of intermolecular disulfide bonds in the covalent Rz homodimer. In the presence of Rz1, the disulfide linkages of Rz-Rz covalent homodimer can form spontaneously. For Rz1, DsbA catalyzes the formation of the Rz1-Rz1 disulfide linkage at the single Cys residue, Cys29. In the absence of DsbA, the homodimer disulfide linkage in each subunit can form as long as the cognate subunit is present, presumably as a template.

***Which form of T4 o-spanin is functional?***

In the light of the bioinformatic analysis (Fig.2.9 and Table 2.2), we tested the importance of cysteine residues for T4 spanin function. The results shown in Fig.2.10B indicate that both cysteines play an essential role in spanin function; however, the presence of two populations of o-spanin (larger fraction of monomer and smaller fraction of dimer) *in vivo* is puzzling, especially when the formation of dimer is dependent on the i-spanin. The question arises, which one of these two forms represents the functional protein. The existence of a greater fraction of monomer is tempting to speculate that the monomeric form is the active form. The existence of intramolecular disulfide linkage in the monomeric form is also evident, as single Cys to Ser substitution in the T4 o-spanin results in the enhancement of the dimeric form. However, the genetic complementation assay suggests that the product of the single Cys to Ser substitution allele is the nonfunctional form. Nevertheless, if the monomeric form is functional then substitution of one or both cysteines in the o-spanin has a significant amount of accumulated monomers. Because this form of monomers are not functional implies that the intramolecular disulfide linkage is critical for the function of T4 spanin. In λ at least one spanin subunit is required to be in homodimer state. Why is the T4 spanin different from the λ spanin? This could be revealed by the comparative study of the λ and T4 lysis phenotype. In addition, is the intramolecular disulfide linkage of the T4 o-spanin mediated by host Dsb system? Why is there any o-spanin dimer that is specifically dependent on the i-spanin? Like λ, the i-spanin in T4 possesses coiled-coil elements

which could help them to dimerize. If so, are these sufficient for T4 spanin function? These are some important questions that need to be addressed. In any case, the powerful selections available and the capacity for facile phenotypic detail suggest that the λ and T4 spanin systems could be useful for exploring the biological determinants of inter-molecular and intra-molecular disulfide bond formation.

# CHAPTER III

# MUTATIONAL ANAYSIS OF λ SPANINS

## Introduction

In the previous chapter, we described the molecular pathway for the formation of covalently-linked i-spanin (Rz) and o-spanin (Rz1) homodimers. Homodimers are the functional form of the i-spanin and o-spanin *in vivo*. In addition, these spanin subunits are known to form a complex and then disrupt the OM (8). Our goal is to understand the structure of the complex and its mechanism to disrupt the OM. It is unclear what specific residues are involved in bringing Rz and Rz1 together to form a spanin complex. Furthermore, it is not known how the spanin complex disrupts the OM after complex formation? It is likely that the multiple steps are required between complex formation and OM disruption. To determine the mechanism of spanin-mediated OM disruption, structural study is required. However, purification of spanins is difficult, in part because overexpression of these proteins is toxic to the cell and purifying membrane protein is difficult. Therefore only the soluble domains have been purified and studied *in vitro* (9). An alternate way to analyze spanin function is to perform a comprehensive genetic analysis of *Rz* and *Rz1*. This may locate the residues involved in complex formation and reveal domain-domain interaction. Genetic analysis can also identify elements of secondary structure involved in function and/or stability. Fortunately, phages are genetically tractable, and offer a non-structural approach to address these questions.

Phenotypic analysis might help us identify mutants that are blocked at different steps of spanin function. However, available genetic information is very limited. Part of the reason lies in the embedded arrangement of *Rz* and *Rz1*. In 1999, Zhang et al. conducted a genetic study to establish that *Rz* and *Rz1* were separate genes (6). This was carried out by complementation analysis, which required inactivation of one gene without affecting the other. This was difficult considering the overlap of the nucleotides between *Rz1* and *Rz*. The complexity was resolved after constructing *Rz* and *Rz1* amber mutants in a way that did not affect the translational product in the overlapped region (Fig.3.1). When tested, individual $\lambda Rz_{am}$ or $\lambda Rz1_{am}$ mutants were unable to plate on a non-suppressor host. The defect was rescued when the host was co-infected by both phages. This study concluded that *Rz* and *Rz1* were independent genes. During this study, other interesting phenotypes were observed. First, the $\lambda\ Rz_{Q100am}$ plated normally on both *supE* and *supF* suppressor hosts, indicating that Rz can tolerate a Tyr residue in addition to the normal Gln in the $100^{th}$ codon. The amber codon in *Rz* was located in the linker region between the two predicted periplasmic helices. This was the first evidence suggesting the linker region was insensitive to missense changes. The second interesting suppressor phenotype was that $\lambda\ Rz1_{W38am}$ plated normally on the *supE* host but was defective on the *supF* host. This was a surprising finding considering the substitution was so conservative and the *supF* suppressor is significantly more efficient than the *supE* suppressor.

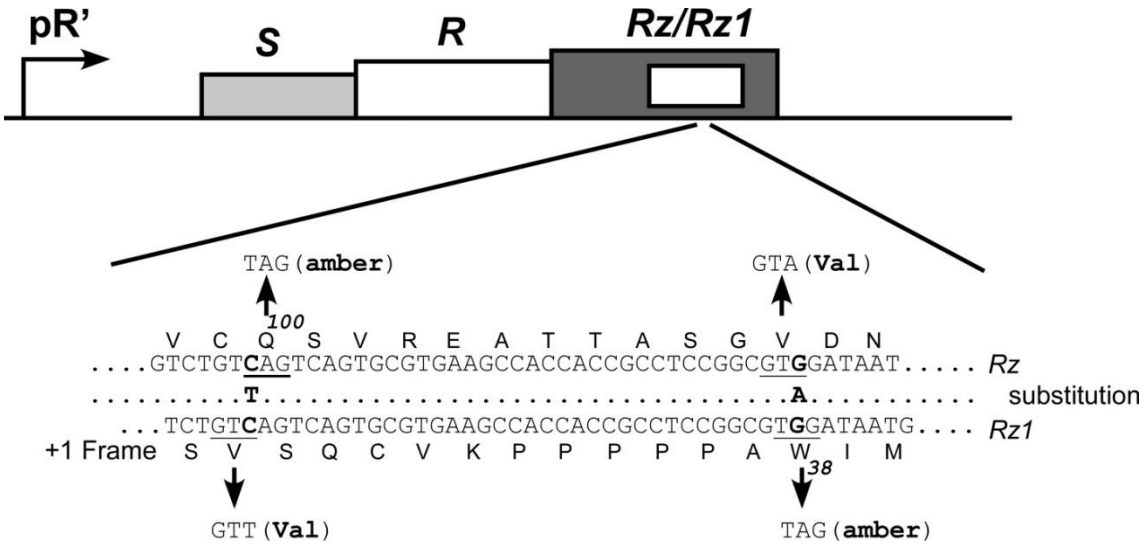



Figure 3.1. Structure of the overlapped region of the $Rz_{Q100am}$ and $Rz1_{W38am}$ genes. Top: a depiction of the λ lysis cassette *SRRzRz1*, which is under the control of the late promoter pR'. The inset shows part of an overlapping nucleotide sequence between *Rz* and *Rz1* with their different translation products. The nucleotides chosen by Zhang et al. (1999) to incorporate the amber codon in *Rz* and/or *Rz1* are shown between the two frames. The $Rz_{Q100am}$ does not affect the overlapping codon for Val in Rz1 and similarly the $Rz1_{W38am}$ does not affect the overlapping codon for Val in Rz.

In 2007, Summer et al. conducted a bioinformatic search for spanins. Researchers grouped the 137 available phage genomes from the RefSeq collection into families based on sequence similarity (7). The alignments of λ Rz and Rz1 families identified multiple conserved residues in the periplasmic domain of each protein, including a PXAWXM motif that spans Trp38 in Rz1. Also conserved in the λ spanin

family were the periplasmic Cys residues, discussed in the previous chapter (see Fig. 2.1). It has been shown that these Cys participate in the disulfide linkage between Rz-Rz and Rz1-Rz1 but how these disulfide linkages are tied to spanin function is unknown.

The material above was the only genetic information available about the spanins of λ at the onset of the work in this dissertation. This chapter describes our effort to extend the genetic analysis of spanins, with four goals: (a) use random mutagenesis to assemble a library of *Rz* and *Rz1* lysis-defective mutants; (b) characterize important mutants in terms of subcellular localization, stability, and function at the molecular level; (c) construct a λ phage with separated *Rz* and *Rz1* genes lacking recombinational similarity ; and (d), use the separated *Rz-Rz1* construct to conduct an inter or intragenic suppressor analysis of Rz-Rz1 interaction. The assembled mutant library will be discussed in terms of spanin function.

## Materials and methods

### *Bacterial strains, plasmids, bacteriophages, and culture growth*

The bacterial strains, bacteriophages, and plasmids used in this study are described in Table 3.1. Bacterial cultures were grown and plated in standard LB media or on LB agar, respectively, supplemented with $MgCl_2$ (10 mM), ampicillin (Amp, 100 μg $ml^{-1}$), and kanamycin (Kan, 40 μg $ml^{-1}$) when appropriate. Culture growth was monitored by recording $A_{550}$ as a function of time. Fresh overnight cultures were diluted

300-fold by transferring 100 μl of overnight culture to 30 ml of LB in a 250 ml growth flask and grown with aeration at 30ºC for lysogenic cultures and at 37ºC for non-lysogens. When needed, lysogens were thermally induced at an $A_{550}$ of ~0.25 at 42˚C for 15 min, followed by continued growth at 37˚C. When indicated, isopropyl β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1mM. For a dominance test, a λ*cI857* lysogen carrying a wild type copy of *Rz* and *Rz1* was transformed with the plasmid bearing a mutant copy of *Rz* or *Rz1*and thermally induced as described above.

The synthetic alleles of *Rz* and *Rz1* were synthesized chemically by GenScript (Piscataway, NJ). These synthetic genes were sub-cloned into a pRE vector, which is a derivative of pJF118EH. The $lacI^Q$ and $P_{tac}$ promoter in this vector is replaced with a pR' promoter region of λ. This promoter is transcriptionally activated by λ antiterminator protein Q (45).

Table 3.1. Bacteriophages, strains, plasmids, and primers

| **Bacteriophages** | **Genotypes and relevant features** | **Sources** |
|---|---|---|
| λ900 | λ*Δ(stf tfa)::cat* $cI_{857}$ *bor::kan;* carries $Cam^R$ and $Kan^R$*;* $Rz^+Rz1^+$ | Lab stock |
| λ900$Rz_{Q100am}$ $Rz1^+$ | | Lab stock |
| λ900$Rz^+Rz1_{W38am}$ | | Lab stock |
| λ900$Rz_{Q100am}$ $Rz1_{W38am}$ | | Lab stock |
| $\lambda^{Syn}Rz^{Syn}Rz1$ | λ*Δ(stf tfa)::cat* $cI_{857}$ *bor::kan;* carries synthetic allele of *Rz* and *Rz1* | This study |
| $\lambda^{Syn}Rz^{Syn}Rz1_{I54N}$ | | This study |
| $\lambda^{Syn}Rz_{D65G}{}^{Syn}Rz1$ | | " |
| $\lambda^{Syn}Rz_{L93S}{}^{Syn}Rz1$ | | " |

Table 3.1. Continued

| **Bacteriophages** | **Genotypes and relevant features** | **Sources** |
|---|---|---|
| $\lambda^{Syn}Rz_{K146N}{}^{Syn}Rz1$ | | |
| $\lambda^{Syn}Rz_{G143R}{}^{Syn}Rz1$ | | " |
| $\lambda^{Syn}Rz_{Y147H}{}^{Syn}Rz1$ | | " |
| $\lambda^{Syn}Rz_{Q151R}{}^{Syn}Rz1$ | | " |
| $\lambda Rz_{R91P}Rz1(embd)$ | Rz1 embedded in Rz | " |
| **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 ($\lambda 900Rz_{Q100am}\ Rz1^{+}$) | MC4100 *tonA::Tn10* lysogenized with $\lambda 900\ Rz_{Q100am}\ Rz1^{+}$ | Lab stock |
| MC4100 ($\lambda 900Rz^{+}Rz1_{W38am}$) | MC4100 *tonA::Tn10* lysogenized with $\lambda 900\ Rz^{+}Rz1_{W38am}$ | Lab stock |
| MC4100( $\lambda 900Rz_{Q100am}$) $Rz1_{W38am}$) | MC4100 *tonA::Tn10* lysogenized with $\lambda 900\ Rz_{Q100am}\ Rz1_{W38am}$ | 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 with intact *bor* gene; sensitive to kanamycin | This study |
| RY17299 $lacI^{q1}$ | Derived from MG1655 *ΔtonA* | (45) |
| **Plasmids** | | |
| pRE | Plasmid with the λ later promoter pR' that is transcriptionally activated by λQ | Lab stock |
| pQ | low copy plasmid carrying antiterminator Q under $P_{lac/ara-1}$ promoter | " |
| pRz | pRE carrying Rz with inactivated Rz1 | Lab stock |
| $pRz_{mutX}$ | pRE carrying lysis defective Rz (see Table 3.1) | This study |
| pRz1 | pRE carrying Rz1 | Lab stock |
| $pRz1_{mutX}$ | pRE carrying lysis defective Rz1 (see Table 3.1) | This study |
| pRz/Rz1 | Rz1 downstream of pRz | Lab stock |

Table 3.1. Continued

| **Plasmids** | **Genotypes and relevant features** | **Sources** |
|---|---|---|
| plipoRz/Rz1 | Rz1 downstream of plipoRz | This study |
| pRz/Rz1$_{his}$ | Rz upstream of Rz1 with His-tag at the C-terminal end | This study |
| pRz1$_{his}$ | Rz1 with His-tag at the C-terminal end | (8) |
| pER157 | pBR322*Δtet SRRzRz1 bor::kan* | (6) |
| pAK1 | pER157 with separated synthetic *Rz* and *Rz1* allele | This study |
| **Primers** | **Sequences** | |
| RM_for | TAGCAGCATGATTGCCACGGATGGCAA | |
| RM_rev | AGGCAAATTCTGTTTTATCAGA | |
| SynRzRz1_for | GATTGATGTATGAGCAGAGTCACCGCGATTATCTCCGCTC | |
| SynRzRz1_rev | GAGTTGCCCATCGATATGGGCAAGCTTTCATCCCCTTTCCG | |
| LC_for | CTTTTACACATGACCTTCGTGAAAGCGGGTGGC | |
| LC_rev | GGAAGGTTTTACCAATGGCTCAGGTTGCC | |
| Rz1I54N_for | CTAAATGGAATCAATTCGCCATCGGAAAG | |
| Rz1I54N_rev | CTTTCCGATGGCGAATTGATTCCATTTAG | |
| RzD65G_for | GCTGCGCTCGGTGCAAAATACAC | |
| RzD65G_rev | GTGTATTTTGCACCGAGCGCAGC | |
| RzL93S_for | GCGGCGCCGATCACATATTAAGG | |
| RzL93S_rev | CCTTAATATGTGATCGGCGCCGC | |
| RzG143R_for | ACAACTGGAACGAACCCAGAAGTAT | |
| RzG143R_rev | ATACTTCTGGGTTCGTTCCAGTTGT | |
| RzK146N_rev | GCTCGATGCAAACTACACGAAGG | |
| RzK146N_rev | CCTTCGTGTAGTTTGCATCGAGC | |
| RzY147H_for | AACCCAGAAGCATATTAATGAGC | |
| RzY147H_rev | GCTCATTAATATGCTTCTGGGTT | |
| RzQ151R_for | ATTAATGAGCGGTGCAGATAG | |
| RzQ151R_rev | CTATCTGCACCGCTCATTAAT | |
| RzR91P_for | GCCGCTGGTCGTCCTCGGTTGCACATCA | |
| RzR91P_rev | TGATGTGCAACCGAGGACGACCAGCGGC | |
| LC_seq | CAAAATTCAAAGAAGCGGGCGGAAC | |

***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. Primer pairs used were RM_for and RM_rev (Table 3.1). To maximize single nucleotide changes, a high amount of template DNA (~5μg of plasmid DNA) was used. 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*. Digested fragments were ligated into pRE plasmid using T4 ligase and transformed into XL-1 Blue cells. Each transformant plate with >200 colonies/plate was pooled by slurrying and then the pooled cells were used to prepare the plasmid DNA using the Qiagen spin miniprep kit. The MC4100 (λ) lysogens with $Rz_{Q100am}$ or $Rz1_{W38am}$ were transformed with the mutagenized plasmid pool and selected for growth in the presence of appropriate drug. In order to account for frequency of mutation, 10 random colonies from each plate were grown and tested for a lysis defect. The lysis defect frequency was 60% for *Rz* and 40% for *Rz1*. 20 % of the *Rz* alleles had a silent mutation and remaining 20% were parental-type. In the case of *Rz1*, 30% of the *Rz1* alleles had a missense change (M11I, W38R, and G52V) that behaved as functional and the remaining 30% were parental type. In parallel, to separate the non-mutagenized parental alleles (~40% of *Rz* and ~60% of *Rz1* based on previous frequency) and enrich for lysis-defective mutants, the transformant plate was slurried, diluted 300-fold into 30 ml of LB with 10 mM $MgCl_2$, grown and thermally-induced as above. After 65 min of induction,

the culture was centrifuged at 4000 rpm for 5 minutes to harvest the non-lysed, $Mg^{++}$-stabilized spherical cells. The supernatant contained plasmid released from lysed cells with functional alleles, thus was discarded. Cation-stabilized spherical cells were carefully washed once with LB supplemented with $MgCl_2$ before plasmid extraction. Plasmids from this initial enrichment were used to transform λ lysogens that were either *Rz* or *Rz1* null. Single colonies were picked and screened for a lysis defect by growing them in 5 ml of LB supplemented with $MgCl_2$ and subjecting them to thermal induction. Plasmids of defective *Rz* and *Rz1* alleles were sequenced by Eton Biosciences (San Diego, CA) to determine the mutation responsible for the lysis defect.

***Detection of spanin complex***

The strain RY17299*lacI*$^q$ *ΔtonA* harboring the plasmid pQ and pRz/Rz1-His or its derivative was grown to mid-logarithmic phase in LB broth. A 25 ml culture of IPTG-induced cells were harvested at 50 min after induction by centrifugation at 6,500 x g for 10 min at 4$^o$C using a Thermo Scientific F15S-8x50cy rotor. The harvested cells were then treated with EDTA and lysozyme according to the spheroplasting procedure described elsewhere (133), except that a protease inhibitor cocktail (Sigma-Aldrich, P8849) was added according to the manufacturer's instructions. Our observation was variable; cells expressing wild type spanins did not survive spheroplasting very well compared to cells expressing mutant alleles. Our objective was to detergent-solubilize the entire spheroplast suspension and isolate the complex. Therefore, the spheroplasts in

a HSB (high sucrose buffer) were then lysed using an equal volume of lysis buffer (10mM sodium phosphate, 50mM NaCl, 0.2% Tween). The isolation of the spanin complex from 500 μl volume of lysate using 20 μl of Dynabeads TALON (Invitrogen) was performed as previously described (8, 116). All samples were eluted from the beads using 50 μl of His-elution buffer (300mM Imidazole, 50 mM sodium phosphate, 300mM NaCl, 0.01% tween-20, pH 8.0). The eluted sample was mixed with 50 μl of 2X sample loading buffer containing 200 mM of Beta-Mercaptoethanol (BME) and boiled for 5 min. A normalized amount of sample corresponding to 0.6 A550 units of culture was run on an SDS-PAGE gel and analyzed by Western Blot as described previously (56).

Accumulation of Rz, Rz1, or other mutant products were tested after precipitating the protein by adding of one ml of culture into a 111 μl of 100% trichloroacetic acid (TCA) that was kept in ice. The precipitated protein was washed twice with acetone and air dried prior to mixing with sample loading buffer. The procedure for SDS-PAGE and Western blotting was followed as described in the Chapter II and elsewhere (56).

### *Separation of Rz and Rz1 in λ chromosome by recombination and site-directed mutagenesis*

The synthetic allele of *Rz1* was cloned downstream of the synthetic *Rz* allele in the pRE plasmid. This pair of synthetic spanin alleles was lifted by PCR using the flanking regions that were 50 bp homologous of either side of embedded *Rz/Rz1* on the λ chromosome. The primer pairs used were SynRzRz1_for and SynRzRz1_rev (Table

3.1). The PCR product was used as a megaprimer to replicate the pER157 plasmid (Table 3.1) that contains the entire lambda lysis cassette including a *bor* gene with an insertion of a kanamycin selectable marker. This PCR product, now containing the synthetic alleles of spanins instead of the embedded genes, was called pAK1. The product mixture was DpnI-digested followed by transformation, and plasmid extraction. The replacement of the synthetic spanins was confirmed by sequencing via Eton Bioscience (San Diego, CA). Additionally, the pAK1 plasmid has a downstream *bor::Kan* used for selection. Strain RY17341, lysogenized with the temperature sensitive λ*cI857,* was transformed with the pAK1 plasmid and selected for its growth on an LB-Amp plate. Single colony was grown to a mid-log phase and thermally-induced to allow recombination between the λ chromosome and the plasmid. After >70 minutes of growth, the culture was treated with 1% chloroform twice for 5 minutes at room temperature to obtain a phage lysate. This phage lysate was used to re-lysogenize a parental RY17341 host and selected for growth on an LB-Kan plate. Single lysogens (pre-screened by PCR method, (114)) were tested for its ability to lyse after thermal induction. The lysis cassette of the recombinant phage was PCR amplified using LC_for and LC_rev primers and sequenced by Eton Bioscience (San Diego, CA) using the primer LC_seq to verify the incorporation of the separated synthetic alleles of *Rz* and *Rz1* in the λ chromosome. To incorporate the desired mutation in the separated allele of *Rz* or *Rz1*, pAK1 was used as a template for standard laboratory site-directed mutagenesis prior to recombination with the parental λ chromosome. Primers used to

generate successful mutants are listed in Table 3.1. The lysis-defective λ recombinants were obtained by treating with 1% chloroform following induction.

***Isolation of suppressors against λ Rz and Rz1 mutants***

Several attempts were made to isolate a second site suppressor against multiple mutants of *Rz* and *Rz1*. Initially, a spontaneous suppressor mutant was sought by plating mutant phages on the sensitive host RY17341. In brief, a 100 μl of mutant phage lysate of ~$2x10^3$ pfu/ml was mixed with the 100 μl of an over-day RY17341 culture with added 5mM $MgCl_2$. After 30 min incubation of the mixture at room temperature, it was mixed with soft agar adjusted with 10 mM, 20 mM, 30 mM, 40 mM, or 50 mM of $MgCl_2$ before pouring on the TB plates. The plates were incubated for minimum of 12 hours at 37˚C. No detectable suppressor was identified this way. Later this method was modified by using a phage lysate that was obtained from three rounds of serial passage of replicating mutant phage in a sensitive host. This method was not fruitful either in generating suppressors. In our last attempt, slight modification was done in a plating strategy that was adopted from Zhang et al. (6). The λ mutants do not form plaques on a sensitive host for ~7 hours compare to a wild type. However, prolonged incubation would result in tiny plaques which have obscured detection of suppressors in our earlier method. Therefore, by decreasing the incubation time we could plate higher number of phages, enhancing the suppressor hunt. Briefly, ~$2x10^6$ phages (1000 fold higher than the previous attempt) were plated on a sensitive host by the standard procedure as

described above, except that these plates were incubated for only 6 hours at 37˚C. A few early plaques with wild type morphology were observed. These early plaques were isolated and purified twice before sequencing the lysis cassette.

## Results and discussion

### *Design and implementation of the mutagenesis system*

Individual clones of *Rz* and *Rz1* (pRz and pRz1described in Table 3.1) were available to implement the mutagenesis (8). To obtain a pool of *Rz* and *Rz1* mutants both pRz and pRz1 plasmids were used as a template and subjected to PCR mutagenesis (see Materials and Methods). This pool of mutant DNAs was used to construct a library of medium- copy plasmids with the *Rz* or *Rz1* mutant gene cloned under the control of the λ late promoter pR'. Initially, the non-lytic alleles were enriched by plasmid retention of those that failed to complement the $Rz_{am}$ or $Rz1_{am}$ defect (Fig.3.2 and see Materials and Methods). From the enriched pool, individual clones were screened by thermal induction for failure to support host lysis. Each lysis-defective clone was sequenced to determine the mutation. In the first batch of lysis-defective mutants from the enriched pools, 30% of the *Rz* and 20% of the *Rz1* mutants were single missense alleles. The remaining lysis-defective *Rz* and *Rz1* alleles had either multiple point mutations or frameshifts, including both base insertions and deletions. Overall, the nucleotide changes were 42% transition and 58% transversion mutations, as expected for this type of mutagenesis (134).

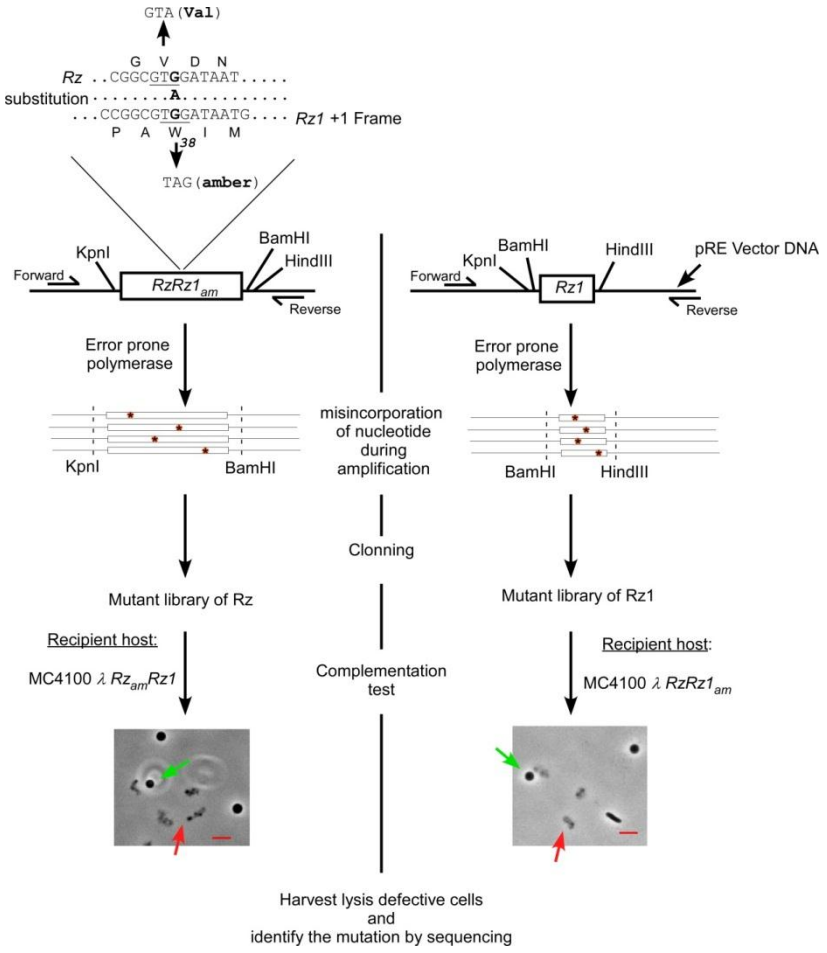



Figure 3.2. Illustration of the major steps in the PCR-based random mutagenesis technique. An individual clone of *Rz* with inactivated *Rz1*is shown at the top left. Inset depicts the amber codon in *Rz1*. A separate clone of *Rz1* is shown at the top right panel. These clones were subjected to PCR amplification using an error-prone polymerase. The forward and reverse primers (see Table 3.1) annealed to the vector DNA flanking the genes are shown by half arrows. Nucleotide changes at random positions during PCR are marked by the red asterisk (*) symbol. The mutant fragments were sub-cloned into the laboratory standard pRE vector using restriction sites shown in the diagram to generate mutant libraries of *Rz* and *Rz1*. The *Rz* mutant library was used to transform a recipient lysogen with a null *Rz* gene. The *Rz1* mutant library was used to transform a recipient lysogen with a null *Rz1*gene. After induction, lysis-defective mutants resulted in spherical cells (green arrows), as opposed to alleles with functional spanin genes, which resulted in lysis (red arrows indicating lysed cell debris). The $Mg^{++}$-stabilized spherical cells were harvested to extract DNA (see Materials and Methods). The mutations (shown in Fig 3.3) were identified by sequencing.

***Lysis-defective mutants of Rz***

Out of a total of 300 sequenced lysis-defective alleles, 64 *Rz* alleles had a single point mutation. Among these, 29 were nonsense changes and 35 were missense changes (Fig.3.3). The remaining 236 lysis-defective alleles had either multiple mutations or single mutation for 21 different positions that were repeatedly isolated (Rz mutants of Table3.2). If it is assumed that all the nonsense mutations accessible by a single base change would be lysis-defective, these results indicated that the mutagenesis had reached >60% saturation (i.e., 29 out of 44 possible nonsense mutants were obtained). It should be noted that repeated mutants were not counted in this measure of saturation. Our high frequency of codon degeneracy indicated that we have reached close to a practical saturation point for the mutation screen. These nonsense mutants were distributed across the entire length of Rz, except for the extreme C-terminal region where three nonsense mutations were not isolated. Among the missense changes, 10 of 35 were proline substitutions, mostly in the predicted periplasmic alpha-helical domains, consistent with the notion that the secondary structure of these domains is crucial for i-spanin function. Overall, the missense mutations were significantly biased towards the proximal helical domain, where 60% (21 of 35) of the missense changes were clustered in only ~30% (residues 43 to 88) of the total Rz sequence. In addition to the Pro substitutions, the amino acid substitutions in the helix region were primarily charge-changes (9 of 21). Charge changes in the context of an important helix would be expected to disrupt protein-protein interactions.

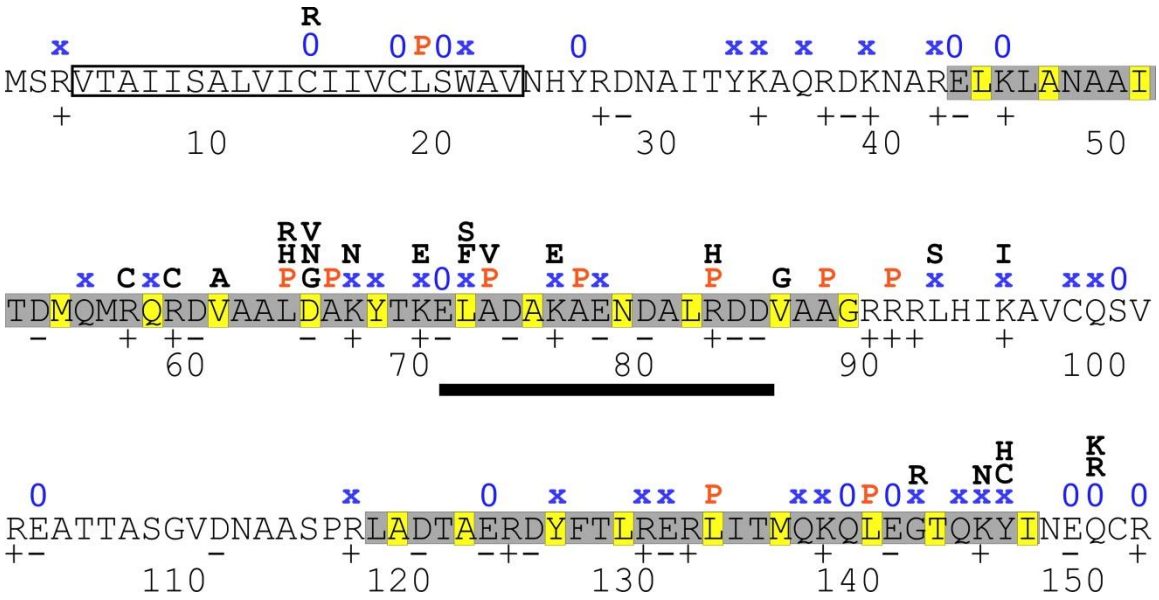



Figure 3.3. Lysis-defective mutants of Rz. Missense changes (bold letters) are indicated above the primary structure with respect to their position indicated below. The predicted topological domains of Rz are: the N-terminal TMD (rectangle) and two predicted periplasmic α-helices (shaded grey), separated by a linker (not shaded). Residues highlighted in yellow represent the "*a*" and "*d*" hydrophobic residues of the heptad repeat (135). The charge of the residues is indicated by + or - signs. Proline residues are shown in red. Nonsense changes (obtained = x, and not obtained = 0) are indicated in blue letters. The epitope region of the Rz antibody is indicated by a black filled line.

Table 3.2. The isolated mutants of Rz and Rz1.

The total number of mutants and their types are indicated in the left end of each row. The last column indicates the repeated number (in parenthesis) of missense and nonsense mutants of Rz and Rz1 for various positions.

| **Type** | **Total** | **Mutations found in Rz and Rz1** | **Repeated mutants** |
|---|---|---|---|
| | | **Rz mutants** | |
| **Missense** | 35 | C14R, L19P, R57C, R59C, V61A, L64H, L64R, L64P, D65G, D65N, D65V, A66P, K67N, K70E, L72F, L72S, A73V, A73P, K76E, A77P, R83P, R83H, V86G, A88P, L93S, K96I, L134P, L141P, G143R, K146N, | Y147C(1), Y147H(1), Q151R(1), Q151K(1), C14R(3), L19P(2), R57C(2), L64H(2), D65G (2) A88P(2), L93S(3), L134P(2) |
| **Nonsense** | 29 | R3St, W21St, Y33St, K34St, Q36St, K39St, R42St, Q55St,Q58St, K67St,Y68St, K70St, K76St, L72St, E78St, L93St, , K96St, C99St, R118St, Y127St, K139St, R131St, E132St, Q138St, K139St, G143St , Q145St, K146St, Y147St, Q36St, Q138St, Q138St, K139St, Q138St, Y127St, Y127St, L93St, W21St, Q145St, R118St, C99St, K146St, Q145St, C99St, Q138St, K96St, R131St, K146St, Q55St, K34St, K34St, K139St, Y127St | Q36St(4), Q55St(3), K67St (2), K70St(2), L93St(2), K96St(3), C99St(2), E132St(3), K139St (3) |
| | | **Rz1 mutants** | |
| **Missense** | 18 | M12R, M12V, G19C, C20R, C20F, C20S, C20Y, P32Q, P33L, P35H, P36Q, P36L, P44S, W46R, W46C, L50R, L50P, I54N | G19C(2), C20S(3), C20Y(2), P36Q(2), L50P(2) |
| **Nonsense** | 7 | K3St, K23St, C20St, C20St, C29St, W38St, W38St, Q41St, Q41St, W46 St | C20St(2), 23St(2),W38St(2), Q41St(3) |

Mutations in the predicted coiled-coil determinants were especially prevalent in the middle of the first helical domain, with 3 of the 4 *a* and *d* positions between residues 62 and 72 accounting for 7 missense alleles. Only three relatively conservative mutations were found in this region: V62A, A73V, and V86G, all of which resulted in changes in the bulk of a hydrophobic side chain. Thus the protein-protein, presumably coiled-coil structure, may require relatively drastic interfacial change to abrogate function. Mutation in one particular position, D65, was repeated five times by two hydrophobic and one uncharged residues: D65G, D65V, and D65N (Fig3.3 and Table 3.3). This indicated that Asp at position 65 which is at the odd position of heptad repeat might play a critical role in the structure and/or function of spanin complex. Overall, these missense changes covered 11 out of 41 conserved residues of the λ Rz family, among which 9 conserved residues belonged to the proximal helix and 2 belonged to the distal helix (compare Fig.3.3 with Fig.6 of Summer et al., (7)). In contrast, the TMD helical domain was relatively silent, with only two missense changes. Importantly, 73 out of 119 possible changes in this region would code for hydrophobic residues which could participate in the trans-membrane helix without major structural changes. This strongly suggests that the TMD of Rz would be less likely to generate non-functional mutants in our screen. The remaining possible changes in this region may lead to charge imbalance which would not be tolerated in the membrane. This is supported by the detection of C14R which is likely prevented from stable membrane localization of Rz. However, the isolation of L19P is somewhat provocative, because proline residues are generally well-tolerated in TMDs (136, 137). There were total of 7 possible changes for proline and we

were only able to obtain one that was lysis defective. In contrast, only three mutations were located in the linker region (i.e., ~6% of the missense mutations in ~20% of the residues (positions 89-117), suggesting that the linker may not contribute a major functional role other than a flexible spacer. The C-terminal 13 residues, likely to contain the domain that interacts specifically with Rz1, were also highly enriched in missense mutations, with 7 mutations in 5 positions. The isolation of Q151R and Q151K as lysis-defective mutations implicates the extreme C-terminus in Rz function and suggests that the lack of nonsense mutations in the last three residues may be happenstance.

***Lysis-defective mutants of Rz1***

Of a total of 200 sequenced *Rz1* lysis-defective mutants, only 25 alleles had a single point mutation. The remaining 175 lysis-defective alleles had either multiple mutations or single mutation for 9 different positions that were repeatedly isolated (Rz1 mutants of Table3.2). Among these, 7 were nonsense changes and 18 were missense changes (Fig. 3.4). Using the criteria explained above, the mutagenesis of *Rz1* was not as close to saturation as in the case of *Rz*, in that 7 out of 18 possible nonsense changes were obtained. In particular, none of the 5 most distal potential nonsense mutants, spanning 14 of the 40 residues of the mature Rz1 sequence, were isolated. Moreover, six of the mutants (M12R, G19C, and all four mutations at C20, are all changes that would abolish secretion or lipoprotein-modification of the Rz1 gene product. Thus only 11 missense changes were isolated in the mature Rz1 sequence, and only 8 of 40 positions

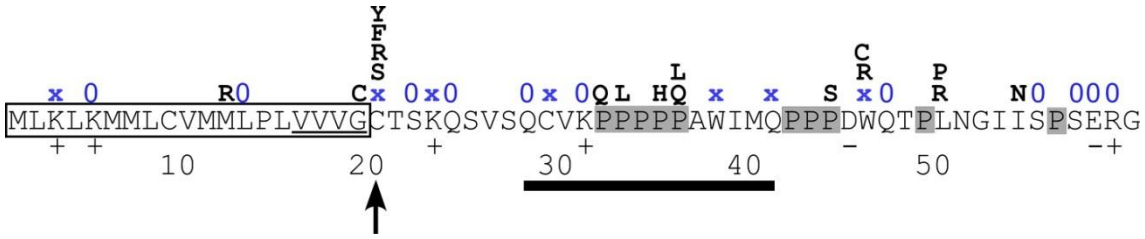



Figure 3.4. Lysis-defective mutants of Rz1. Missense changes (bold letters) are indicated above the primary structure. The N-terminal end of Rz1 consist of a signal sequence (rectangle box in the N-terminus) marked by a lipobox (underlined) and a SPaseII cleavage point (upward arrow at position 20). The polypeptide of a mature form of Rz1 has many proline residues in its periplasmic domain (shaded in grey). The charge of the residues of Rz1 are indicated by + or - signs. The position and total number of nonsense changes indicate the saturation level of the screen. Nonsense changes (obtained = x, and not obtained = 0) are indicated in blue letters. The epitope region of the Rz1 antibody is indicated by a black filled line.

accounted for these changes. The most striking feature of the mutational distribution, especially in contrast to the frequency of mutations to proline in *Rz*, is the prevalence of mutations in Pro codons in *Rz1*. Rz1 is proline-rich, with 10 Pro residues in the mature (40 residue) sequence. Five of these Pro residues gave rise to missense changes, including four Pro residues in a penta-proline stretch (residues 32-36). Within this pentapetide region, no mutation was obtained in the central proline at position 34, consistent with a previous alanine-scanning mutagenesis study that demonstrated that a $Rz1_{P34A}$ retains full lytic function (Joel Berry, unpublished). Taken together, these results suggest that a proline helix secondary structure is not required for Rz1, which

would require three or more consecutive Pro residues to make a turn (138). Other than the existing proline residues, only three other residues in the periplasmic domain give rise to lysis defective mutations: W46, L50 and I54N. The W46C mutations would likely cause the formation of a disabling intermolecular disulfide bond involving C29. Interestingly, although change-from-proline mutations dominate the mutational spectrum of *Rz1*, one of the few other mutations that were isolated is L50P, a mutation that would create a Pro-Pro sequence in the distal region of Rz1. This suggests that the Pro-rich domains of Rz1 have a more specific structural role than just constituting a disordered structure that might confer membrane-disordering character, as suggested previously (72).

***Expression of lysis-defective Rz and Rz1 alleles***

It was important to determine whether the lysis-defective missense mutations in *Rz* and *Rz1* affected accumulation of the protein. Therefore, whole-cell samples were taken at 50 min post-induction and examined by Western blotting. Most of the alleles exhibited wild type protein expression, with a few exceptions (Fig.3.5). For Rz, L72F appeared to be unstable, marked by reduced accumulation and the presence of apparent degradation products. Another allele, K76E, was not detected, but this result is open to question because residue 76 is part of the epitope for the anti-Rz antibody, so the drastic Lys to Glu change could be expected to affect immuno-detection (139). In Rz1, only the accumulation of P33L, P35H, and P36Q appeared diminished, but all these mutations

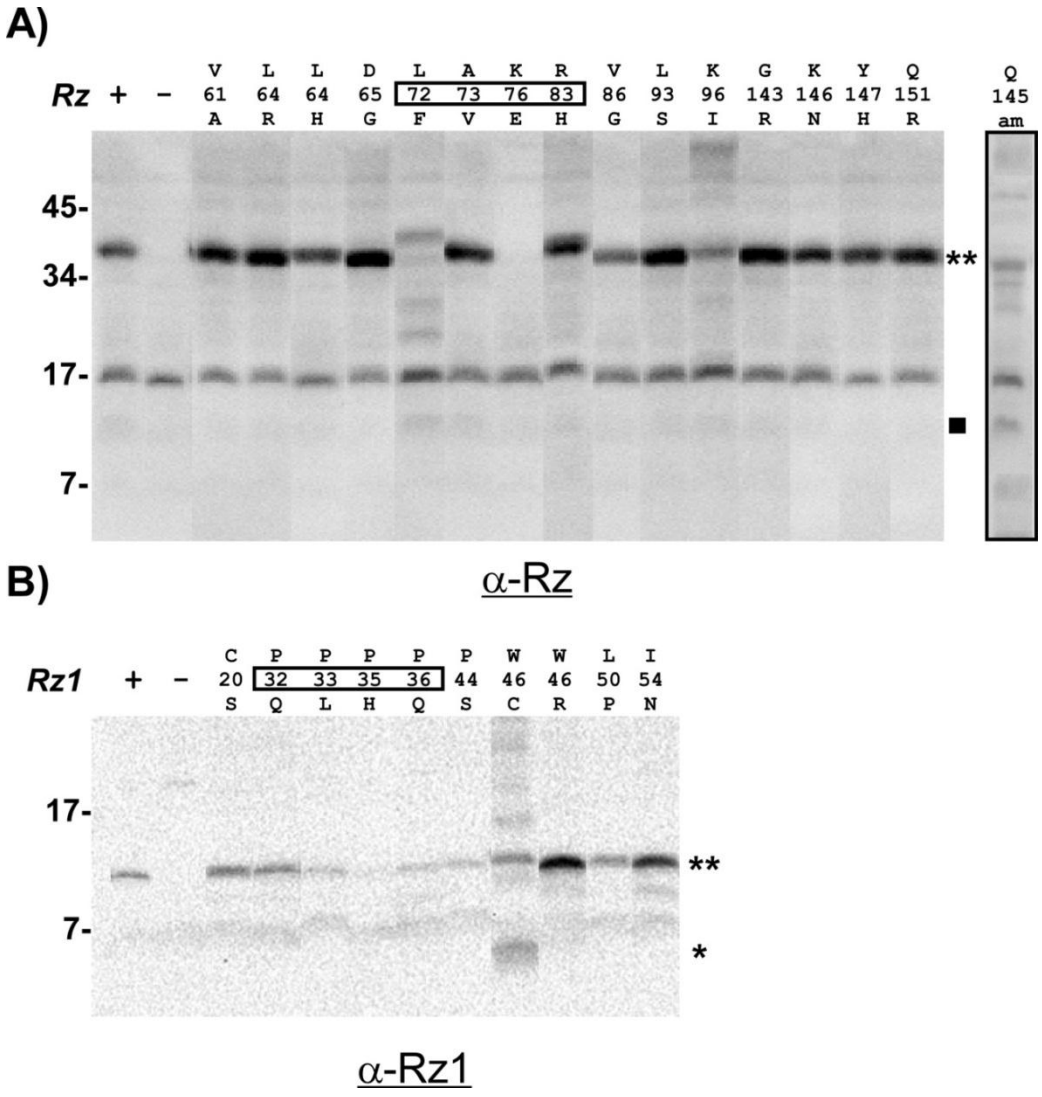



Figure 3.5. Lysis-defective mutants of Rz and Rz1 accumulate to physiological level. TCA precipitates of whole-cell samples taken at 50 min post induction were analyzed by western blot under non-reducing conditions. Representative Rz mutants **(A)** and Rz1 mutants **(B)** are indicated at the top of each panel. Accumulation of protein due to mutation in the epitope region was poorly detected (boxed mutants). A "+" sign indicates wild type and a"-" sign indicates $Rz_{Q100am}$ (in **A**) and $Rz1_{W38am}$ (in **B**). Protein accumulation of $Rz_{Q145am}$ is shown in far right of panel A. The monomer and dimer products are indicated by an asterisk symbol and the putative degradation product is indicated by a filled square on the right.

fall within the Rz1 epitope region, highlighted in Fig.3.5. It is important to notice that almost all Rz and Rz1 mutants appear to accumulate exclusively as disulfide-linked homodimers (double asterisks) with the exception of Rz1W46C, which likely is largely blocked in an internal disulfide-bonded state. Thus, the function of these defective alleles is most likely blocked after their dimerization step.

***Rz mutants form a complex with Rz1***

To test if the various Rz mutants were able to form the spanin complex with wild type Rz1, pull-down experiment using a functional oligohistidine-tagged Rz1 protein was performed, as previously described (8). As shown in Fig.3.6, most of the Rz mutants co-purified with Rz1-His, comparable to wild type (compare lane 1 with the rest). This suggests that these mutants are not defective in forming a Rz-Rz1 complex and are, thus, presumably defective at another, downstream step after complex formation. Only Rz L72F and Rz Q145am, both which were shown to accumulate *in vivo*, did not segregate with the His-tagged Rz1 on immobilized cobalt beads. Instead a ~12 kDa degradation product of Rz (black square in Fig. 3.6) was detected. In addition, the same degradation products were detected with three other Rz mutants: V86G, R91P, and L93S. Rz has been shown to be susceptible to proteolytic degradation in the absence of Rz1, indicating that the interaction between Rz and Rz1 is required to stabilize the complex. It is possible that some of these Rz mutants are unable to form a stable complex with Rz1.

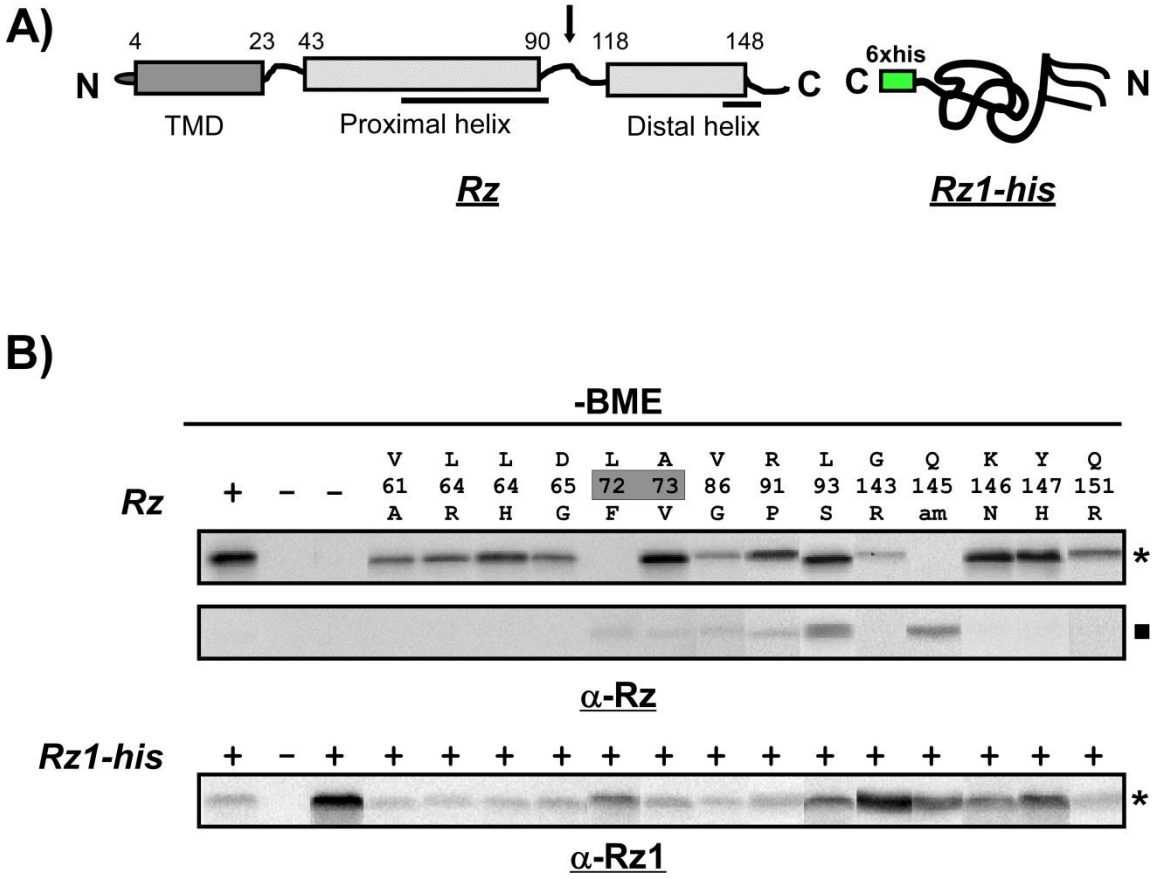



Figure 3.6. Complex formation with lysis-defective Rz mutants. **(A)** Shown are the predicted structures of Rz and Rz1. The position of most amino acid substitutions causing a defective lysis phenotype in Rz is underlined below the stylized secondary structure. The 6XHis-tag in the C-terminal end of Rz1 is represented by a green square. The arrow in the linker indicates the putative cleavage point of Rz in the absence of complex formation (56). **(B)** RY17299 *lacI*$^{q}$ cells carrying the indicated Rz-Rz1$_{his}$ pair were treated with EDTA and lysozyme at 50 min post induction and were solubilized in a detergent before subjecting the lysate to affinity fractionation by magnetic Dynabeads. Proteins bound to the beads were separated by SDS-PAGE under non-reducing condition and analyzed by western blotting with antibodies indicated below each panel. The "+" and "-" symbols indicate the presence and absence of the wild type allele in trans, respectively. The monomer band of Rz and Rz1 are indicated by an asterisk symbol and the cleavage product is indicated by a filled square. The L72F and A73V mutations are in the epitope region.

Overall the degradation product was more apparent when the mutation was in the proximal helix of Rz (residue 72 to 93) or Rz was truncated near at the C-terminal end. We were also surprised by the variability in the amount of isolated Rz1-His, as it is evident from the Western blot shown in Fig.3.6. One interpretation is that the diminished amount of Rz1-His could be due to the occlusion of His-tag epitope after complex formation with Rz. This could be explained based on the observation that a higher amount of Rz1-His was captured by the metal beads when only Rz1-His was expressed (Fig.3.6, lane 3). In cases where Rz1-His was co-expressed with the C-terminal mutants of Rz, we were able to affinity purify relatively higher amount of Rz1-His (lane of G143R to Y147H), in exception to the Q151R mutation at the terminal end of distal helix. It is possible that the diminished or enhanced amount of Rz1-His is dependent on the degree of exposure of His-tag during complex formation. On the other hand, isolation of the degradation product is tempting to speculate that Rz1-His can directly interact with N-terminal half of Rz. However, it cannot be ruled out that these complexes were formed due to an indirect interaction.

***Dominance testing of spanin mutants***

The lysis-defective mutants of *Rz* and *Rz1* were tested for their dominance or recessiveness character. To assign dominance or recessive character, individual mutant alleles of *Rz* or *Rz1* was induced for expression from the medium copy plasmid in the presence of prophage-borne wild type copy of *Rz* or *Rz1*. All of the mutant alleles

indicated in Fig.3.3 and 3.4 were unable to block lysis, indicating they were all recessive to the parental alleles.

***Synthesis of separate, "dis-embedded" Rz and Rz1 genes***

Genetic characterization of λ *Rz* and *Rz1* was limited due to the unique embedded genetic architecture. For most mutations, a point mutation in the overlapped region in *Rz1* would incorporate a change in *Rz* and vice versa. By separating the *Rz* and *Rz1* genes in λ, rigorous genetic studies, including suppression analysis, could be done for the individual spanin subunits. Suppression analysis might, ideally allow mapping the interacting residues of Rz and Rz1. However, separating the parental genes of *Rz* and *Rz1* would be problematic in the context of the phage because of the DNA identity throughout the common DNA sequences, which would make homogenotization inevitable in the context of a selection. Therefore, separate synthetic alleles were generated by making every possible change in the codon without affecting the amino acid sequence of the primary translational products (Fig.3.7A). This manual approach allowed us to separate the 100% overlapping nucleotide region of *Rz* and *Rz1* into two alleles that are now by comparison only ~40% identical and thus incapable of homologous recombination. The parental embedded copy of *Rz/Rz1* in λ chromosome was then exchanged with the tandem, separated synthetic alleles of *Rz* and *Rz1* by recombinatorial approach (Fig.3.7C and see Materials and Methods). The recombinant phage, designated as $\lambda^{Syn}Rz^{Syn}Rz$, exhibited plating behavior indistinguishable from the

Figure 3.7. Synthetic λ i-spanin and o-spanin are comparable to the parental embedded genes. **(A)** A nucleotide alignment of synthetic *Rz* ($^{Syn}$*Rz*), embedded *Rz/Rz1* ($^{Emb}$*Rz/Rz1*), and synthetic *Rz1* ($^{Syn}$*Rz1*) alleles with their translation. The sequence on the top and bottom codes for the same polypeptide. This polypeptide sequence is denoted in the middle. The start codon of all three alleles is underlined. **(B)** A heat map of the nucleotide identity between the overlapping region (nucleotide 208 to 402) of $^{Syn}$*Rz* and $^{Syn}$*Rz1* alleles was generated by a base to base comparison. Blue bar = identical nucleotide, red bar = different nucleotide. **(C)** Construction of λ with separated *Rz* and *Rz1* genes by homologous recombination. Plasmid pAK1 carrying the lysis cassette of λ with *Rz* and *Rz1* synthetic alleles and the *bor::kan* selectable marker was transformed into an *E.coli* λ lysogen, in which the prophage carries the parental embedded *Rz/Rz1*. After thermal induction of the lysogen, a mixture of phage progeny is produced, from which a recombinant phage carrying *bor::kan* allele was selected by re-lysogenization into new host and selection for Kan resistant (*see Materials and Methods*). The recombinant phage that carries the separated synthetic alleles *Rz* and *Rz1* was verified by PCR (inset; compare lane 2 and 3; emb=embedded genes & sep= separated genes). **(D)** The lysis kinetics and plaque morphology of isogenic λ carrying the parental or separated synthetic spanin genes.The culture turbidity was measured after thermal induction of the λ prophage and the plaques were scanned 12 hrs after incubation at 37˚C. Prophages: *λcI857Rz$_{am}$/Rz1$_{am}$* (empty circle, -○-); *λcI857$^{Emb}$Rz/Rz1* (filled circle, -●-): *λcI857$^{Syn}$Rz$^{Syn}$Rz1* (cross, -x-).

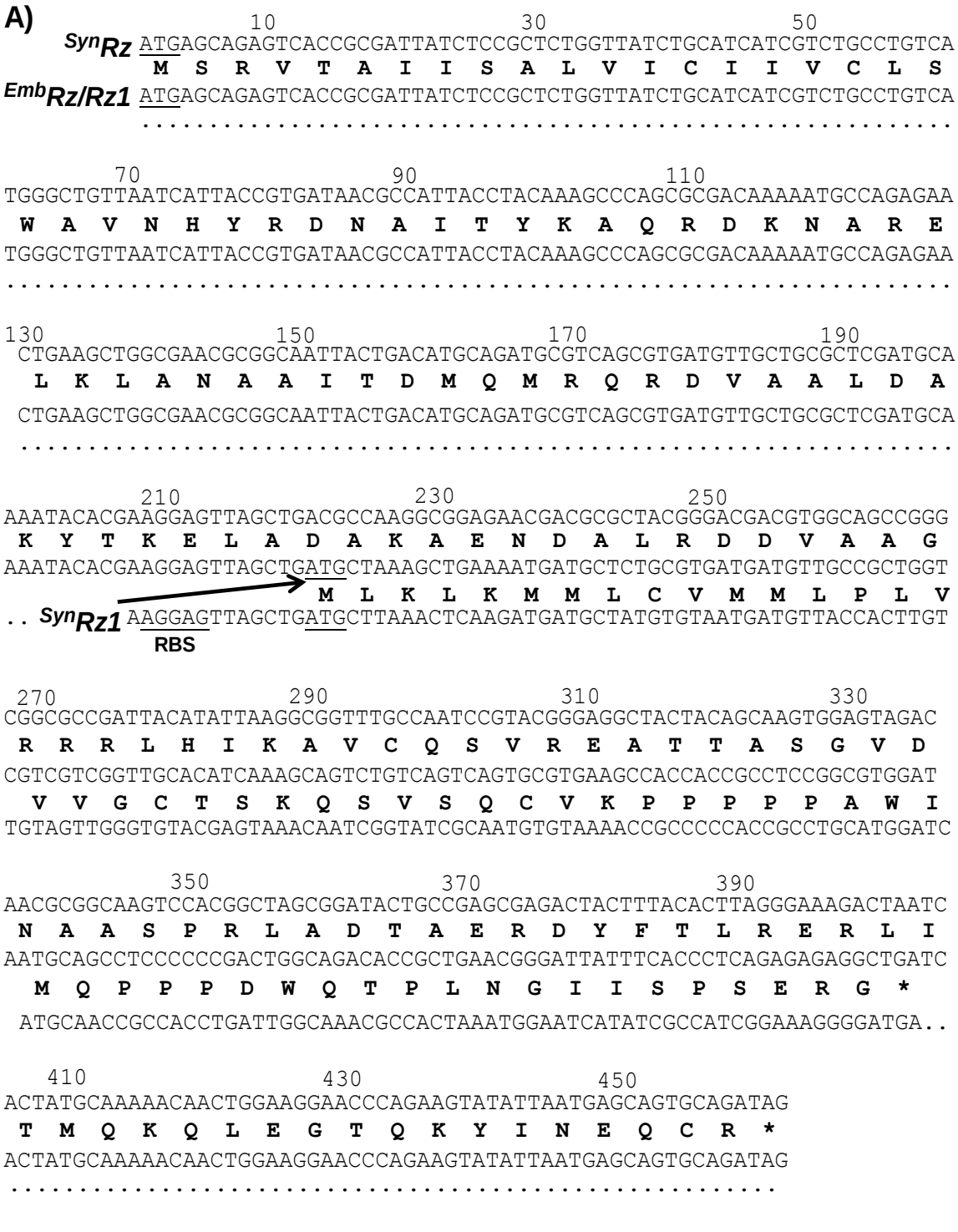



Figure 3.7. Continued

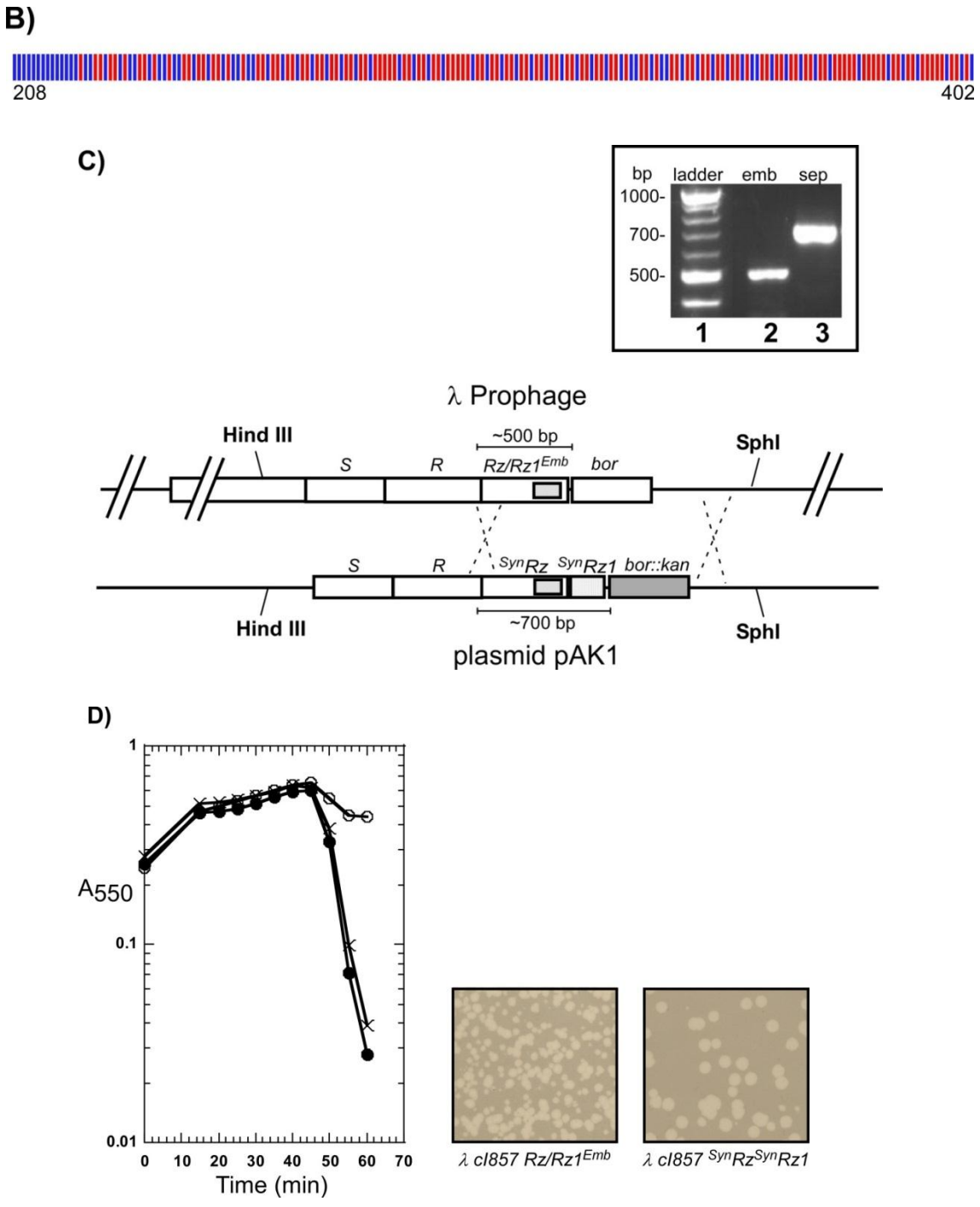



Figure 3.7. Continued

parental (Fig.3.7C). Moreover, the lysis kinetics observed from induction of lysogens carrying the $^{Syn}Rz^{Syn}Rz1$ genes was also identical.

### *C-terminal mutation in Rz is suppressed by missense change near the interface of the cytoplasmic membrane*

Our initial objective of separating the *Rz* and *Rz1*genes in the λ chromosome was to eliminate the complication of genetic characterization of embedded genes. In addition, we envisioned that it would allow us to identify the interacting domains of Rz and Rz1 that participate in the spanin function and give us insights about the mechanism. The idea was to isolate intergenic suppressors for various lysis-defective mutants of Rz or Rz1. Therefore, 7 λ variants with carrying single missense changes in the separated genes; 6 in *Rz* and 1 in *Rz1* (Table 3.3) were subjected to lysogenic induction. Although the mutants do not have an absolute plating defect, the parental and spanin mutant phages can be differentiated by the rate at which plaques develop on plates (6). The first three independent suppressor mutants were obtained as early-developing plaques for the mutant allele $Rz_{Y147H}$ , which has a mutation near the extreme C-terminus. When we purified the suppressor plaques and analyzed them by sequencing thelysis gene regions, all of them had the same intragenic suppressor mutation in *Rz*: S20P. To test whether other mutations elsewhere in the genomes of the suppressor isolates were required for the suppression, the $Rz_{S20P,Y147H}$ double mutant and each of the single mutants were created in the context of pRE and tested for complementation of the lysis defect of $Rz_{am}$ (Fig. 3.8). The double mutant exhibited full lytic function, as did the single $Rz_{S20P}$

mutant, whereas the $Rz_{Y147H}$ mutant exhibited its absolute lysis defect. The S20P change is located at the C-terminal end of the TMD, near the interface between the cytoplasmic membrane and the periplasm. It is surprising that introducing a Pro residue, presumably

Table 3.3. List of λ spanin mutants and suppressor mutations.
Each mutation in column 1 was incorporated into the separated spanins of the synthetic genes. The intragenic compensatory change of S20P in the RzY147H mutant restores its function. "-" means no suppressor has been identified yet.

| **Mutation type** | **Suppressor type/ change** |
|---|---|
| **Rz D65G** | **-** |
| **RzL93S** | **-** |
| **RzG143R** | **-** |
| **RzK146N** | **-** |
| **RzY147H** | **Intragenic / S20P** |
| **RzQ151R** | **-** |
| **Rz1I54N** | **-** |

putting a kink in the helical TMD as it emerges from the bilayer, can suppress the defect of a missense change at the extreme C-terminus of the periplasmic domain, which is presumably disposed near the interface of the periplasm and the outer membrane. Moreover, the functionality of S20P itself is unexpected, considering that a Pro

substitution in the adjacent residue, L19P, has an absolute lysis defect, as do so many other Pro substitutions in the periplasmic domain. Mechanistic interpretation will have to wait until more suppressor mutants are isolated, allele-specificity can be determined and useful biochemical assays are developed. Nevertheless, this result demonstrates that the separated synthetic gene construct can be used with the early plaque-development selection to isolate true suppressors. Although the first suppressor turned out to be an intragenic suppressor, there is no reason to expect that intergenic suppressors will not also be obtainable.

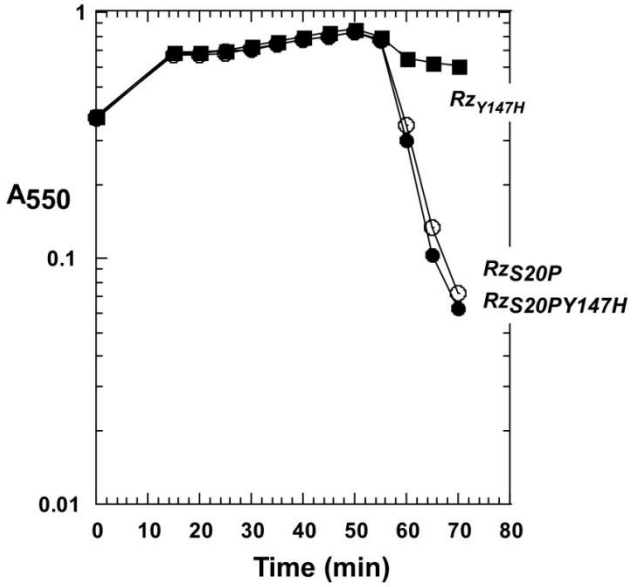



Figure 3.8. Intragenic suppression of $Rz_{Y147H}$. MDS-12(λ$Rz_{am}Rz1_{am}$) lysogens were thermally induced with concomitant expression of $Rz_{S20P}$ (solid circle, -●-), $Rz_{Y174H}$ (soild square, -■-), or double mutant $Rz_{S20P,Y147H}$ (open circle, -○-) from the plasmid in trans. The culture turbidity was measured as described above.

# CHAPTER IV

# MEMBRANE FUSION BY λ SPANINS DURING PHAGE LYSIS*

## Introduction

Phage lysis, the most common cytolytic event in the biosphere, has been extensively studied in phage λ, where four genes encoding five proteins (Fig. 4.1A) effect a three-step lytic process that releases the progeny virions (8, 140). The infection cycle suddenly terminates when the S105 holins, small membrane proteins encoded by gene *S*, are redistributed into large two-dimensional foci, resulting in the formation of micron-scale holes in the cytoplasmic membrane (41). This event, called holin "triggering", occurs at a time specific to the allelic state of *S* and is temporally regulated by the proportion of a second *S* product, the antiholin S107 (34, 125). The R endolysin is then able to escape through the holes to attack the PG. Since the PG layer confers shape and mechanical integrity to the cell, holin and endolysin function was long thought to be necessary and sufficient for lysis (40, 141) . However, recent genetic and physiological studies revealed that two other λ proteins, Rz and Rz1, are also required (Fig. 4.1B) (10). Rz and Rz1 are a type II integral membrane protein (N-in, C-out) and an OM lipoprotein, respectively (51, 52, 66). Interacting by the C-termini of their periplasmic

---

* Reprinted with permission from "Membrane fusion during phage lysis" by Rajaure M., Berry J., Kongari R., Cahill J., and Young R., 2015. *Proc Natl Sci Acad Sci USA*, doi:10.1073/pnas.1420588112 by HighWire Press.

domains, Rz and Rz1 form a complex spanning the entire periplasm, designated as the spanin complex to denote its topology in the envelope. Accordingly, Rz and Rz1 are designated as the *i*nner membrane (i-spanin) and *o*uter membrane (o-spanin) subunits of the spanin complex (Fig. 4.1A, B) (56). Experiments with GFP-Rz chimeras and biochemical analysis of envelope proteins indicate that the spanin complexes accumulate in the envelope throughout the morphogenesis period of the infection cycle (9, 10). The available data indicate that, after destruction of the PG by the endolysin, the spanin complex functions to disrupt the OM. In the absence of spanin function, the infection cycle terminates in a spherical cell form, in which the IM has been lethally permeabilized by the holin, the PG has been destroyed by the endolysin, but the OM is intact (Fig. 4.1C, previously published in (10)).

The λ *Rz* i-spanin and *Rz1* o-spanin genes have a unique genetic architecture, with the latter entirely contained within the former in the +1 reading frame (Fig. 4.1A). To our knowledge, these are the only genes in biology that share the same DNA in different reading frames and are both required for the same biological function (7). Primary structure analysis of Rz indicates that its periplasmic domain has a mostly alpha-helical structure, with two helical domains separated by a central unstructured hinge region. In contrast, the mature Rz1 o-spanin is only 40 aa and is predicted to be unstructured, largely because of its high proportion of proline residues (10 of 40). These predictions were supported by CD analysis of the purified periplasmic domains (9). When mixed, the two periplasmic domains formed complexes with a 1:1 subunit ratio and underwent a conformational change involving a dramatic increase in alpha-helicity.

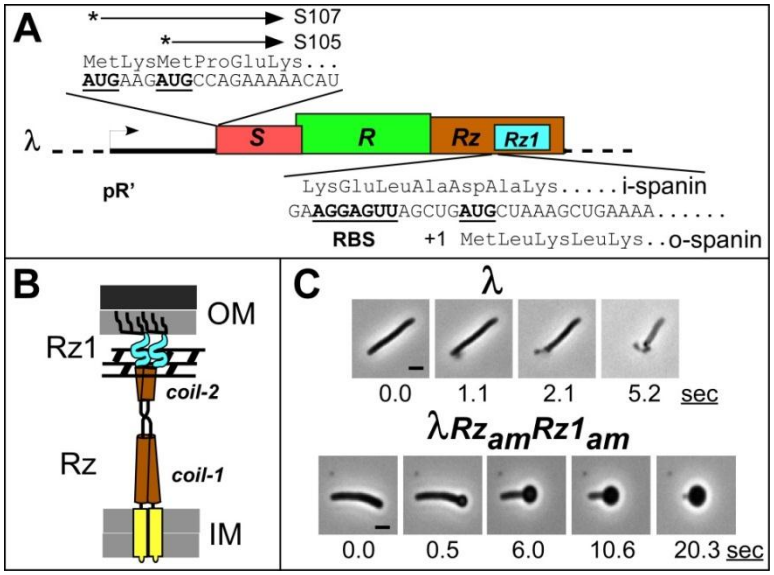



Figure 4.1.The lysis genes, topology and features of spanin complex, and lysis morphology. **(A)** The lysis cassette of paradigm phage λ. The λ lysis genes encoding the holin (red), endolysin (green), i-spanin (brown), and o-spanin (cyan), are located downstream of the late gene promoter pR'. λ *S* gene has a dual-start motif, encoding both the holin (S105) and antiholin (S107) (34). Both start codons are bold and underlined. The 2CS spanin gene of λ is the representative of the embedded gene architecture. The insets show the translation frame of Rz1 in +1 frame of Rz. The strong ribosome binding site (RBS) for Rz1 translation is underlined. The start codon for Rz1 is bold and underlined. **(B)** Spanin topology. Cartoon models for the topology of the 2CS spanin (Rz = i-spanin; Rz1 = o-spanin). The Rz i-spanin, embedded in the cytoplasmic membrane by an N-terminal TMD (yellow), has coiled-coil alpha helical periplasmic domains (brown) separated by a linker region. The mature Rz1 o-spanin (cyan) is attached to the OM via the 3 fatty acyl groups decorating the N-terminal Cys residue. Both the i-spanin and o-spanin form covalent homodimers linked by intermolecular disulfide bonds in the periplasmic domains (56). The i-spanin/o-spanin complex is shown connecting the IM and OM through the meshwork of the PG (black hatching). **(C)** Spanin-null lysis phenotype (adapted from Berry et al. (10) with permission[*]) Representative time-series images of induced λ lysogens, either wt (top) or $Rz_{am}Rz1_{am}$ (bottom), are shown. Time (in sec) is from first detectable change in cell morphology at the onset of the lysis process, ~ 50 min after induction. Scale bar 1 μm.

[*] Panel C of this figure is reprinted with permission from "The spanin complex is essential for lambda lysis" by Berry J., Rajaure M., Pang T., and Young R.,2012. *J. Bacteriol.* 194:5667-5674, 2012 by American Society for Microbiology.

Imaged by negative-stain EM, the complexes were found to be oligomeric rod-like bundles of ~25 nm length, corresponding to the measured width of the lateral periplasm. The complexes are thought to form by interactions at the C-termini of both the subunits, a notion supported by yeast two-hybrid analysis of the analogous i-spanin and o-spanin genes of phage T7 (67).

The fundamental question is how the Rz-Rz1 complex achieves disruption of the OM barrier. It was proposed that spanin complexes interacted with holins to transduce the micron-scale lesions formed in the cytoplasmic membrane directly to the OM (69). However Berry et al. (31) showed that spanin function was also required for lysis even if the endolysin is secreted via the general secretory pathway, rather than via the holin-mediated holes. Thus spanin function is necessary for lysis after degradation of the murein layer, irrespective of the presence of holin lesions in the cytoplasmic membrane. We have proposed that spanins remove the OM as a topological barrier by fusing it with the cytoplasmic membrane (8). Indeed, membrane fusogenic activity has been reported for Rz1 and suggested as a component of lambda lysis, based on studies with purified protein and overexpression of Rz1 *in vivo* (72). However, these results have been challenged (8).

Here we report experiments designed to interrogate the lytic mechanism of the phage spanins. The results are discussed in terms of the membrane fusion model, the regulation of spanin function, and the potential for leveraging spanins for the study of the fundamental process of membrane fusion.

## Materials and methods

### *Bacterial growth, induction, and time-lapse phase-contrast microscopy*

Bacterial strains used in this study are listed in Table 4.1. Bacteria were grown in standard LB media or on LB agar, supplemented with appropriate antibiotics, as indicated: ampicillin (Amp, 100 μg/ml), kanamycin (Kan, 40 μg/ml), chloramphenicol (Cam, 10μg/ml). D-Ala auxotrophic strains were supplemented with 150μM of D-Ala as indicated. Cultures were started with a 1:300 dilution of a fresh overnight culture in 25ml of media in a 250 ml flask and grown with aeration either at 30$^{o}$C for lysogenic cultures or 37$^{o}$C for non-lysogens. Lysogens were thermally induced at $A_{550}$~0.25 by shifting to 42$^{o}$C for 15 min and then continued growth at 37$^{o}$C until lysis. Cultures induced for the fusion assay (see below) were chemically induced by adding isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Phase-contrast microscopy was conducted as described previously (10), except that an EC Plan-Neofluor 100X/1.3 oil Ph3/ ∞/0.17 objective was used to capture the images. For imaging, a 1ml sample of growing culture was harvested by centrifugation at 12,000 rpm for 1min and immediately re-suspended in 100uL of LB, of which 1 μl was spotted on the ethanol washed glass slide and covered with the 22x22mm coverslip prior to mounting on the stage for the inverted microscope objective. Cells of growing cultures were imaged every 30 min or as indicated.

Table 4.1. Bacteriophages, strains, and plasmids

| Bacteriophages | Genotypes and relevant features | Sources |
|---|---|---|
| λ900 | λ*Δ(stf tfa)::cat* $cI_{857}$ *bor::kan;* carries $Cam^R$ and $Kan^R$*;* $Rz^+Rz1^+$ | Laboratory stock |
| λ900$Rz_{Q100am}$ $Rz1^+$ | Q100am nonsense allele of *Rz* | (6) |
| λ900$Rz^+Rz1_{W38am}$ | W38am nonsense allele of *Rz1* | (6) |
| λ900$Rz_{Q100am}$ $Rz1_{W38am}$ | | (6) |
| λ901 | λ900 $S_{am7}$ nonsense allele of *S* holin gene | Laboratory stock |
| λ901$Rz_{Q100am}$ $Rz1^+$ | | (8) |
| λ901$Rz^+Rz1_{W38am}$ | | (8) |
| λ901$Rz_{Q100am}$ $Rz1_{W38am}$ | | (8) |
| ***E. coli* strains** | | |
| MC1000 | $araD_{139}$ *Δ(ara-leu)*$_{7679}$ *galU galK Δ lac* | (142) |
| MB2159 | MC1000 $dadX_{EC}$*::frt* $alr_{EC}$*::frt* | (142) |
| MC1000 (λ901) | Lysogen carrying λ901prophage | This study |
| MC1000 (λ901$Rz_{Q100am}Rz1_{W38am}$ ) | | This study |
| MB2159 (λ901) | | This study |
| MB2159 (λ901$Rz_{Q100am}$ $Rz1^+$) | | This study |
| MB2159 (λ901$Rz^+Rz1_{W38am}$) | | This study |
| MB2159 (λ901$Rz_{Q100am}$ $Rz1_{W38am}$) | | This study |
| RY17299 $lacI^{q1}$ | Derived from MG1655*Δ tonA* | (45) |
| **Plasmids** | | |
| pQ | λ *Q* cloned under $P_{lac/ara-1}$ promoter in a low copy number plasmid pZS-24* | (143) |
| pRE | medium copy vector carrying Q-dependent pR' (λ late promoter). | (45) |
| pRz | *Rz* cloned in pRE with $Rz1W_{38am}$ null mutation | (8) |
| pimRz1 | $Rz1_{T21DS22D}$ cloned in pRE | (8) |
| pRz/ GFP | *gfp* cloned downstream of *Rz* in pRz | This study |
| pimRz1/GFP | *gfp* cloned downstream of imRz1 in pRz | This study |
| $pRz_{R91P}$ | pRz carrying R91P allele of *Rz* | This study |

Table 4.1. Continued

| Bacteriophages | Genotypes and relevant features | Sources |
|---|---|---|
| $pRz_{Q151R}$ | pRz carrying Q151R allele of *Rz* | This study |
| $pRz1_{I54N}$ | pmRz1 carrying I54N allele of *imRz1* | This study |
| $pRz_{R91P}$/GFP | pRz/GFP carrying R91P allele of *Rz* | This study |
| $pRz_{Q151R}$/GFP | pRz/GFP carrying Q151R allele of *Rz* | This study |
| pmCherry/ $imRz1_{I54N}$ | pmCherry/imRz1 carrying I54N allele of *imRz1* | This study |
| $pGFP_{A206K}$ | *gfp* cloned downstream of the pR' promoter | Laboratory stock |
| pmCherry | *mCherry* cloned downstream of the pR' promoter | Laboratory stock |
| pmCherry/Rz | *mCherry* cloned upstream of *Rz* in pRz | This study |
| pmCherry/imRz1 | *mCherry* cloned upstream of $Rz1_{T21DS22D}$ allele in pimRz1 | This study |

***D-Ala starvation/depletion***

To impose D-Ala starvation, an overnight culture grown in supplemented D-Ala was washed twice with plain LB prior to a 1:300 dilution into 25ml of LB amended to 10% sucrose. For D-Ala depletion experiments, the overnight culture diluted 1:300 into fresh LB media and 150μM D-Ala. At A550 ~ 0.25, the culture was harvested by centrifugation at 6,500 x g for 5 min at room temperature using a Thermo Scientific F15S-8x50cy rotor, washed three times with LB to remove the remaining D-Ala in the media and resuspended in fresh media (without sucrose).

***Strain and plasmid construction***

Strain MB2159, carrying complete deletions of *alr* and *dadX*, and its isogenic parental MC1000 were generous gifts from Michael Benedik (142). Both of these strains were lysogenized with phage λ901, λ901$Rz_{Q100am}$ $Rz1^{+}$, λ901$Rz^{+}Rz1_{W38am}$ , or λ901$Rz_{Q100am}$ $Rz1_{W38am}$ (Table4.1) as described elsewhere (6). Phages λ901 and its derivatives carry the temperature-sensitive repressor *cI857* and contain deletion substitutions of non-essential genes; *stf::cam* and *bor::kan*. Lysogens were thus selected by plating on LB-kan plus D-Ala whenever necessary. Candidate lysogens were tested for a single copy prophage as described elsewhere (114).

The plasmids used in this study were constructed by standard molecular biology techniques. The plasmid of pBR322 derivative engineered with a λ late promoter pR' was used to regulate the expression of *Rz*, *imRz1*, *GFP*, and *mCherry*. In the paired clone *GFP* and *Rz*& *mCherry* and *imRz1* were cloned 30 and 22 nucleotides apart with their own Shine-Dalgarno sequence for translation. The expression of all individual or paired clones was dependent on the activity of the antiterminator Q of λ. In a non-lysogenic strain the antiterminator Q was supplied in trans from a plasmid pQ, a low-copy plasmid with a IPTG inducible $P_{lac/ara-1}$ promoter (8).

***Spheroplast fusion assay and fluorescence microscopy***

The host for all spheroplast experiments was RY17299, which is MG1655 *lacI*$^{q1}$ *tonA::*Tn*10* (Table 4.1). This host was transformed with pQ, a low copy plasmid carrying the λ *Q* late activator gene under *lac* control (Table 4.1). The plasmid pQ is compatible with the plasmids carrying the spanin genes (*Rz* and *imRz1*) and fluorescent protein genes (*gfp* and *mCherry*) (see description above) cloned under the control of the λ late promoter, pR'.

For Rz-imRz1 fusion experiments, logarithmically growing cells were induced at A550~ 0.25. At t=55 min, the cultures were placed on ice for 10 minutes to stop growth. Approximately 26 ±1 A550 units of cells from each culture were harvested by centrifugation at 6,500 x g for 10 min at 4$^{o}$C using a Thermo Scientific F15S-8x50cy rotor. Pellets were resuspended, washed twice in cold HSB (0.75M sucrose, 10mM Tris-acetate, and pH7.8) and finally resuspended in 0.5 ml HSB. The concentrated cells from two cultures (e.g., one expressing GFP and Rz and another expressing mCherry and imRz1) were combined to make 1 ml final volume and transferred into a small glass beaker on ice. This mixture was first supplemented with 10 μl of 5 mg/ml DNase (Sigma-Aldrich) and 50 μl of 10 mg/ml egg-white lysozyme (0.1 mg per 10 A550 units; Sigma-Aldrich) with gentle mixing, followed by incubation in ice for 2-4 minutes. For spheroplasting, the lysozyme-treated cells were then treated with 2.1 ml of 1.15 mM EDTA (twice the volume of initial resuspension) by slow swirling in ice bath as described elsewhere (133). The EDTA treatment was adjusted to the flow rate of ~250

μl/min. All preparation steps were carefully done in ice after harvesting the cells. For testing the effect of energy poisons on spanin-mediated spheroplast fusion, the spheroplasting protocol was changed so that KCN was added at a final concentration of 10 mM simultaneously with the standard addition of EDTA. We have noticed that the treatment of cell with higher amounts of lysozyme caused spheroplast clumping, an effect that has been described elsewhere (133). We also noticed that spheroplasting of cells above A550 ~ 0.3 was generally unsuccessful. After ~10 min incubation of the spheroplast mixture in ice, 1 μl of sample was placed onto a coverslip (24x50 mm, thickness 0.16 mm), gently covered with a pre-made agarose pad (0.6% wt/vol in 0.75M sucrose) of ~ 5x5 mm dimensions and immediately imaged by an inverted epifluorescence microscope (Eclipse TE2000-E; Nikon) using a 100X objective (Plan Fluor, numerical aperture 1.40, oil immersion) and standard filter sets. Microscopy was done at room temperature. Images were acquired using a CCD camera (Cascade512; Photometrics). Image analysis was done using NIS-Elements imaging software. For each experiment, snapshots of 10-20 different frames were captured before repeating the process. Each frame was analyzed individually and number of spheroplasts emitting red, green, or both were manually scored. The percentage of fused spheroplast was calculated using the basic formula [% = (number of fused cells/ total)*100; total = red + green + both]. Similarly, % of adhesion point = (n/ total)*100; where n= number of visible interface between red and green as shown in the inset of Fig.3B. For example, if a single red cell is clustered with two green cell then n=2 and total cell number = 3.

### *Protein accumulation*

To assess the accumulation of spanin subunits, 1 ml samples of induced cultures were added to ice-cold tricholoroacetic acid (10% final concentration) for the precipitation of the total cellular protein. Total precipitated proteins were analyzed by Western blot as previously described (56).

## Results

### *In vitro spheroplast fusion by λ spanins*

To address how the spanin complex effected OM disruption, we were first stimulated by electron-microscopic images of the products of spanin-mediated lysis in the absence of holin function (Joel Berry, unpublished). In this case, endolysin activity was secreted directly to the periplasm by the host *sec* system, leaving the IM intact. In the absence of spanin function, endolysin-mediated degradation of the PG did not result in lysis. Instead, spherical cells were formed, bound by the OM. In the presence of spanins, however, lysis was observed and the lysis debris was dominated by highly vesicularized membranous structures. This led us to posit that spanins achieve destruction of the OM by a topological solution: fusion of the IM and OM (Fig.4. 2A), analogous to the vesicular fusions central to Golgi trafficking, neurotransmitter release, and viral fusion (76).

To test this model, we developed a spheroplast fusion assay (Fig. 4.2B). First, we constructed two plasmids engineered to express either *Rz* or *Rz1* at physiological levels and also a fluorescent protein to label the spheroplast cytoplasm: GFP for *Rz* and mCherry for *Rz1*. The wt allele was used for the *Rz* construct but for *Rz1*, a missense allele, *imRz1*, was constructed in which the parental Thr21-Ser22 codons were substituted by Asp codons. Asp residues in these positions, immediately distal to the lipoylated Cys, prevent mature lipoproteins, including Rz1, from being sorted to the OM by the host Lol system (8, 144). We hypothesized that spheroplasting these cells by standard techniques (i.e., treatment with EDTA and lysozyme; (133)) would expose the periplasmic domains of the Rz i-spanin and the ectopically localized imRz1 o-spanin. Mixtures of such spheroplasts should allow spanin complex formation between these exposed periplasmic domains, which, according to our model, would lead to fusing of the spheroplast membranes and mixing of the labeled cytoplasms (Fig. 4.2B). Accordingly, when the two induced cultures were mixed at a 1:1 ratio and converted to spheroplasts, we found ~10% of the product spheroplasts were labeled with both GFP and mCherry (Fig. 4.3B-II). This meant that at least 20% of the input cell mixture participated in spheroplast fusion events, since each doubly-labeled spheroplast required at least one input singly-labeled cell. Most of the fusion events had occurred by the time of the first microscopic examination after spheroplasting (Fig.4. 3A and Fig 4.4C). However, occasionally, red and green-labeled spheroplasts could be visualized in the process of fusion, which occurred only between spheroplasts already adhered together and was completed in the time range of 1-5 min (Fig.4. 4A).

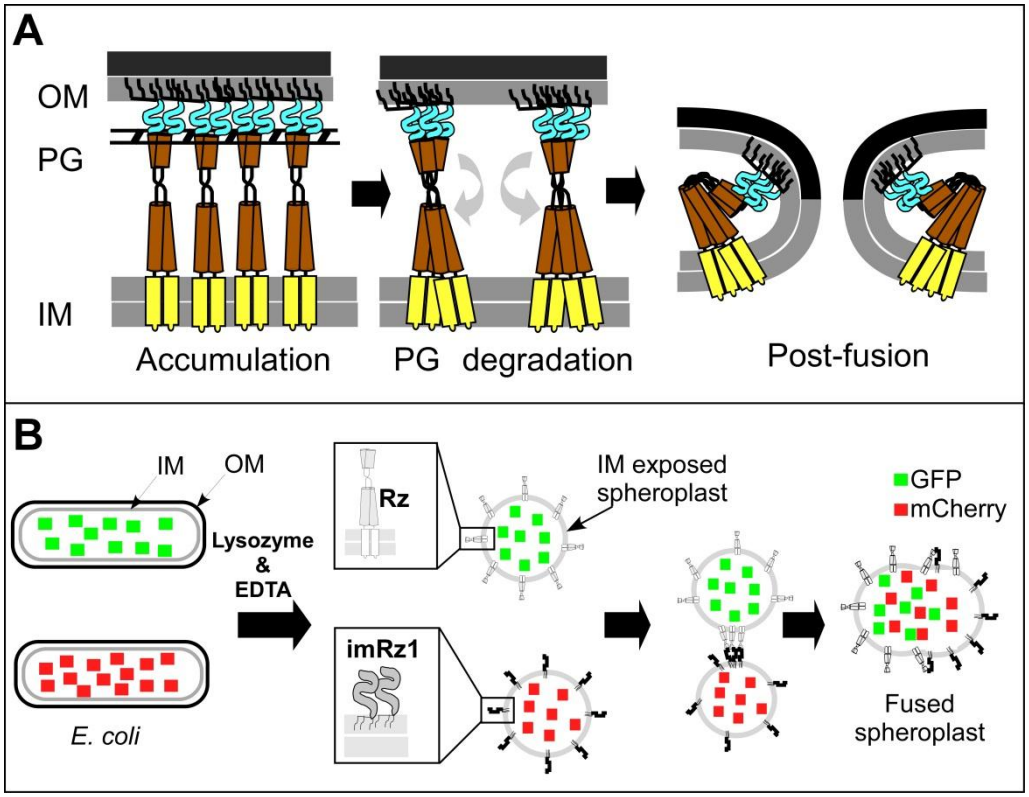



Figure 4.2 Membrane fusion model and spheroplast fusion assay **(A)** Membrane fusion model for Rz-Rz1 lytic function. During the infection cycle, spanin complexes connecting the IM and OM accumulate within the lacunae formed by PG cross-linking. Endolysin-mediated destruction of the PG liberates the spanins to oligomerize and undergo conformational changes to bring the opposing membranes into contact and stimulate IM-OM fusion. **(B)** Fusion assay. Shown are schematics of spheroplasts prepared from rod shaped cells expressing either Rz/GFP (top) or imRz1/mCherry (bottom). Spanin-mediated fusion events result in mixed cytoplasmic contents, labeled with both GFP and mCherry.

The double-labeled spheroplasts were larger and more irregular in shape than the spheroplasts with only a single label (Fig. 4.4C). No doubly-labeled spheroplasts were produced when either one or both plasmids were substituted with a vector lacking a

spanin subunit (Fig. 4.4B and Fig. 4.5A). However, fusion events were obtained, albeit a much lower level, when both GFP and mCherry-labeled cells produced Rz (~3.5% doubly-labeled spheroplasts) or imRz1 (~1% doubly-labeled spheroplasts) (Fig. 4.5B). Occasionally, less than 0.5% of fusion occurred in our controls lacking functional spanin in the plasmid. This could have arisen due to harsh treatment of cells with EDTA and lysozyme which makes the spheroplasts fragile. The optimization of spheroplast for fusion assay will be re-visited in near future and is discussed in Chapter V.

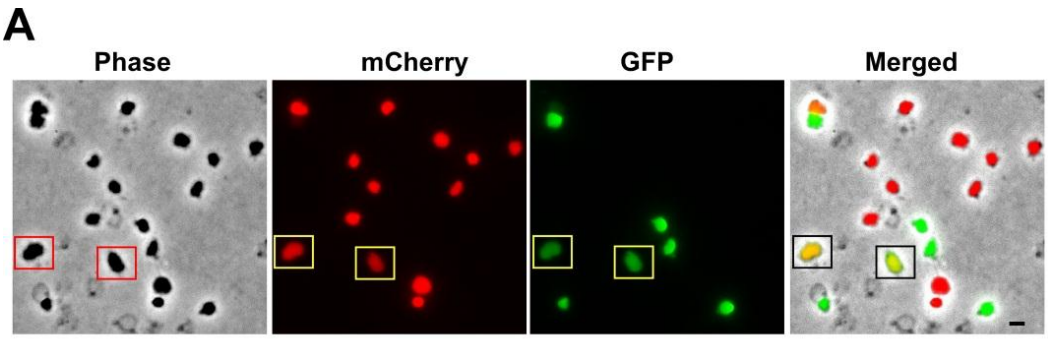



Figure 4.3. Qualitative and quantitative analysis of fusion events. **(A)** Fluorescence images of the fused spheroplasts (boxes) indicate the co-localization of the GFP and mCherry. **(B)** Quantification of fusion events. (Upper graph) The percentage of labeled spheroplasts obtained with different combinations of *Rz* and *imRz1* alleles labeled at the bottom. Red bar: spheroplasts labeled with mCherry; green bar: spheroplasts labeled with GFP; yellow bar, spheroplasts labeled with both fluorescent proteins. (Lower graph) The percentage of spheroplast-spheroplast adhesions (see Methods and inset of B). The adhesion fraction increases to ~24% when the fusion is blocked by mutations in both the Rz and Rz1 periplasmic domains (see text and inset of B). The dashed line represents the background level of adhesions in the absence of spanin protein. Values indicate the mean ± SD of three replicates (set I-V) and two replicates (set VI). Spheroplast count; n= 533 ± 23 for each replicate. *, $P < 0.05$; **, $P < 0.005$ versus negative control (one tailed t-test).

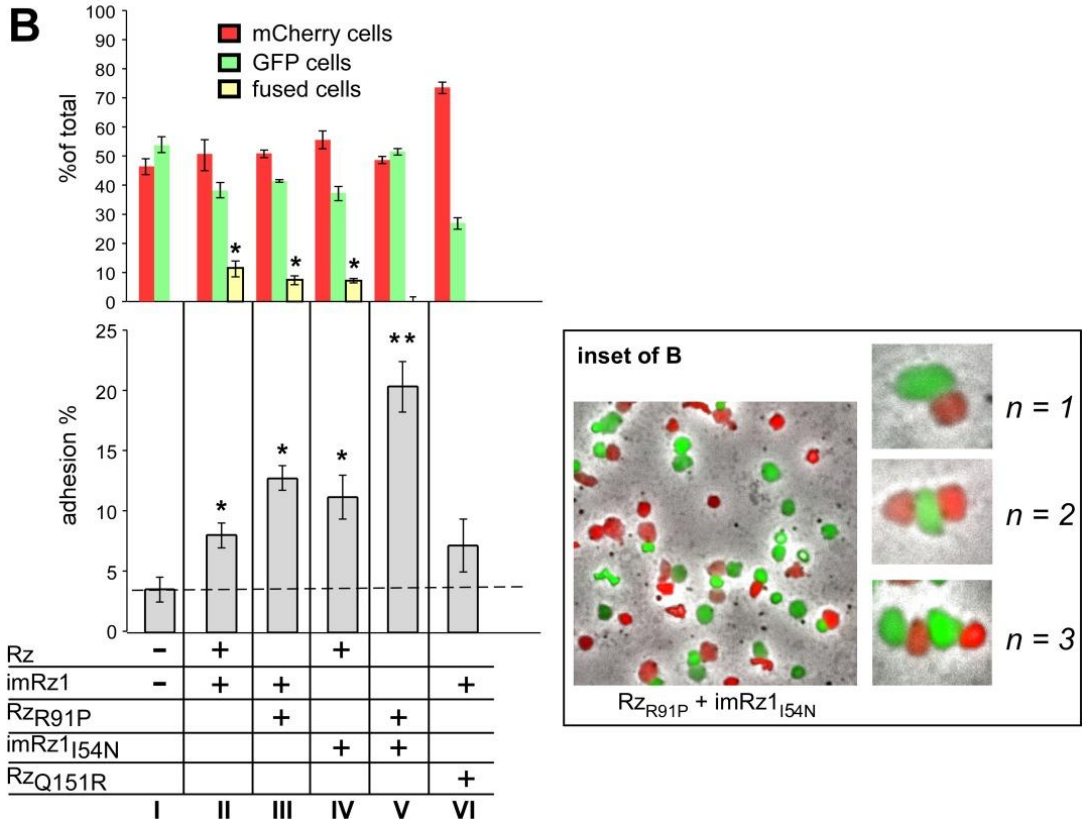



Figure 4.3. Continued

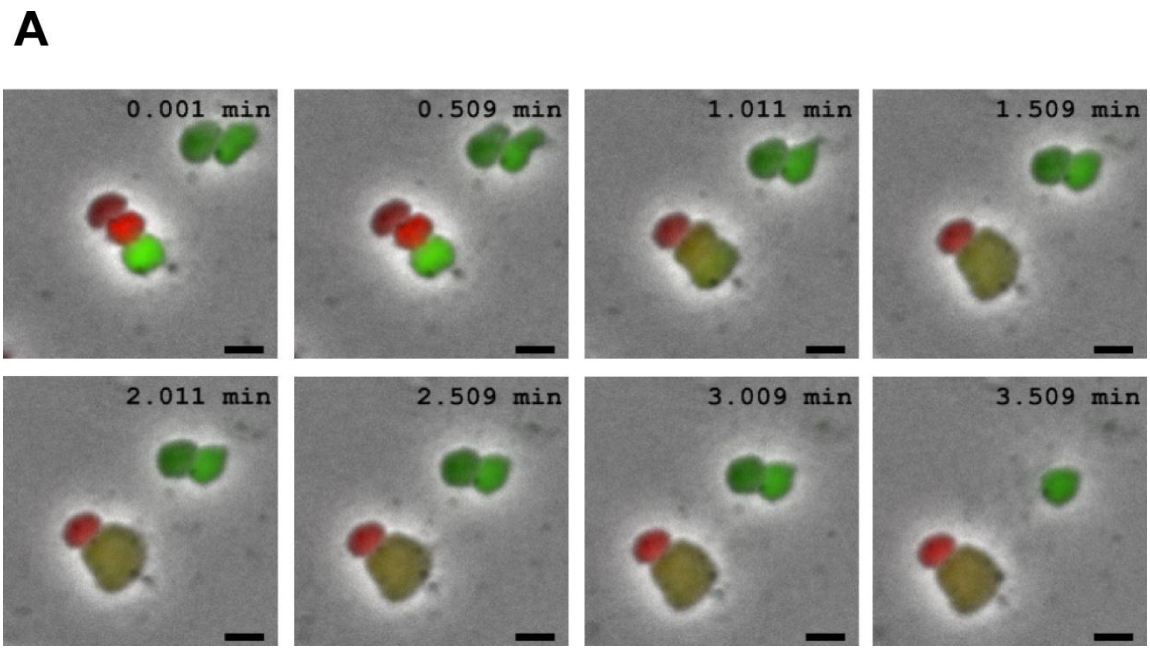



Figure 4.4. Micrographs of spheroplast fusion **(A)** Complete fusion of adhered spheroplasts expressing Rz and imRz1 (previous page). An area containing two adhered spheroplasts of different fluorescent markers was located under the microscope at the beginning of imaging process. Images were captured in 30 sec interval. In this and other visualized fusion events between spheroplasts that have expressed Rz/GFP and mCherry/imRz, fusion occurred in less than 1.5 min. Time is indicated in the top. **(B)** Images of adhered spheroplasts lacking Rz and imRz1 were treated similarly as of panel A and images were captured for 15 minutes. For simplicity, the frames of 2 minutes interval are shown. **(C)** A collage of representative images of fused spheroplast (boxed) that are larger in size and more irregular in shape. Scale bar = 5μm.

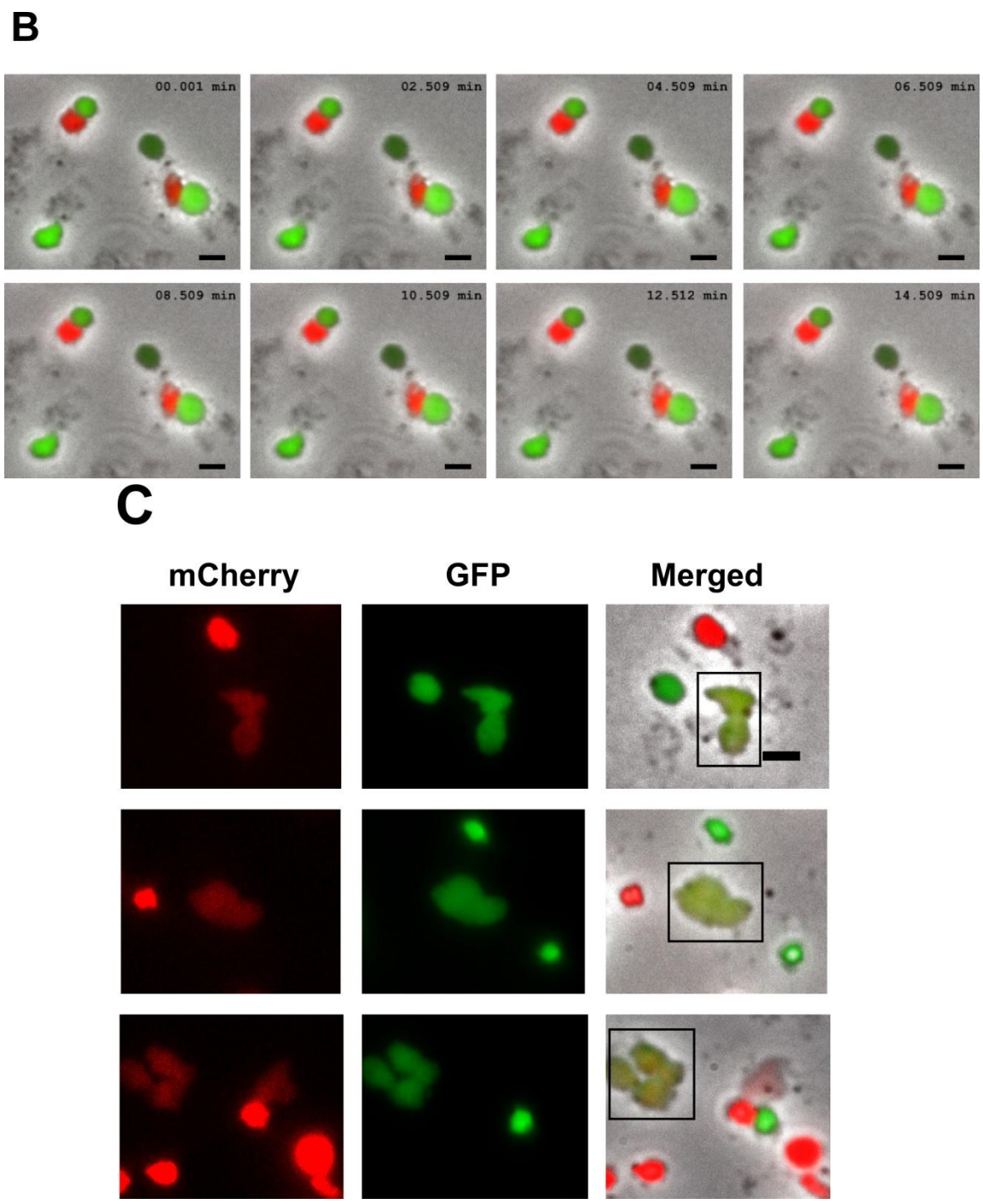



Figure 4.4 Continued

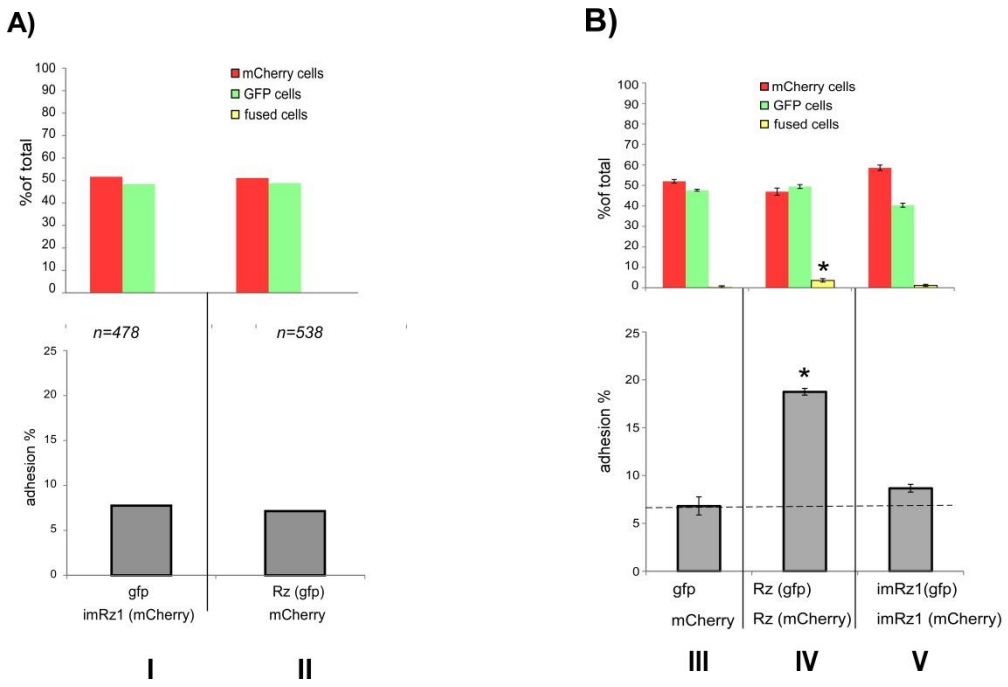



Figure 4.5. Spheroplast fusion requires both spanin subunits **(A)** (Upper graph) The percentage of labeled spheroplasts obtained from fusion assay lacking either *Rz* (lane I) or *imRz1* (lane II) allele labeled at the bottom. (Lower graph) The percentage of spheroplast adhesion in the absence of either *Rz* or *imRz1*allele. Spheroplast count; n=478 and n=538. **(B)** Quantification of homotypic fusion events by spanin subunits. (Upper graph) The plasmids used to generate the labeled cells were: I, pGFP and pmCherry; II, pRz/GFP and pmCherry/ Rz; and III, pimRz1/GFP and pmCherry/imRz1. (Lower graph) The percentage of spheroplast-spheroplast adhesions with the dashed line indicating the background of adhesion obtained in the absence of spanin proteins (see text for details). Values indicate the mean ± SD of two replicates. Spheroplast count; n= 527±23 for each replicate. *, $P < 0.05$ versus negative control (one-tailed t-test). Red, green and yellow bars represent fraction of spheroplasts labeled with mCherry, GFP or both fluorescent proteins, respectively.

***Lysis-defective λ spanins fail to fuse spheroplasts in vitro***

By several criteria, *in vitro* spanin-mediated spheroplast fusion had similar properties to spanin-mediated OM disruption *in vivo*. First, spheroplast fusion was insensitive to the energization of the cytoplasmic membrane, as judged by the result that the presence of 10 mM KCN, shown in previous work to collapse the membrane potential and trigger holin function, had no effect on the yield of fused spheroplasts (Fig.4.6).

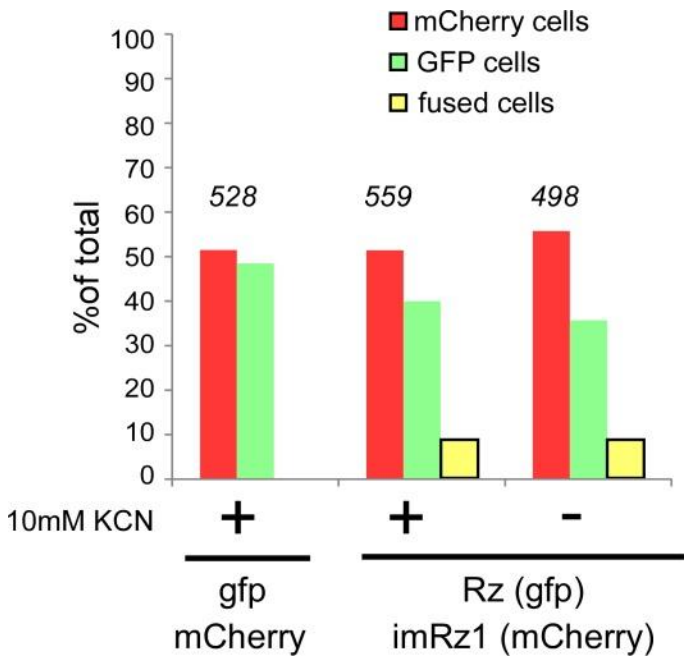



Figure 4.6 Effect of KCN on spanin-mediated spheroplast fusion. The fraction of fused spheroplasts, as judged by the fluorescence of mCherry, GFP or both fluorescent proteins, is indicated. The total number of spheroplast counted for each set is indicated at the top. Presence or absence of 10mM KCN is indicated by (+) or (-) sign at the bottom of the graph.

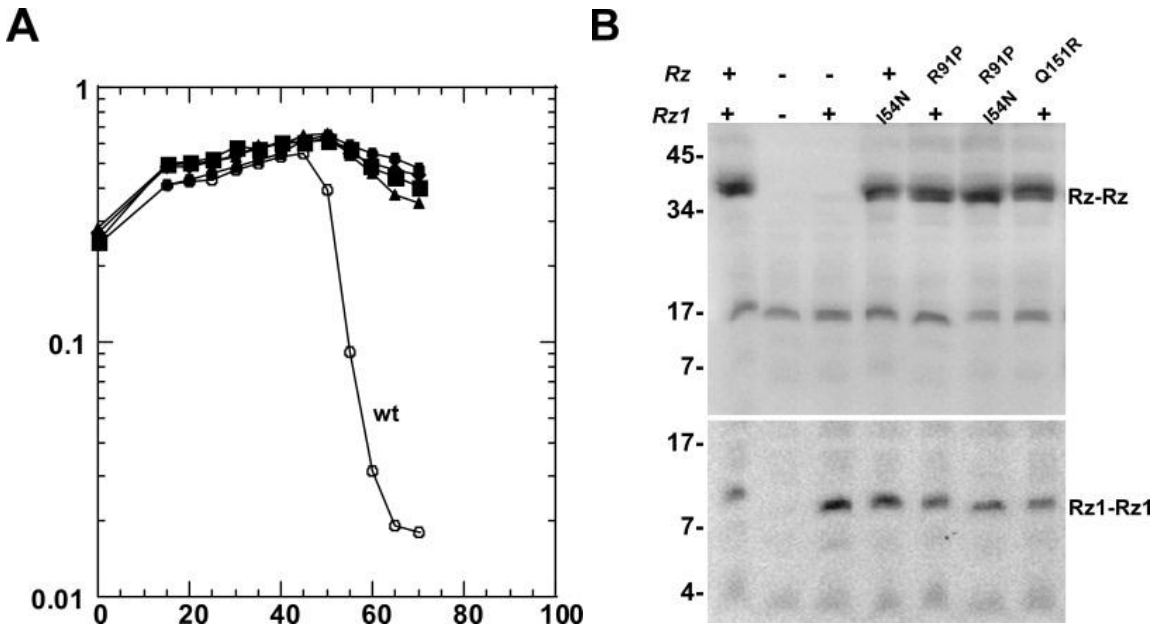



Figure 4.7 Missense mutations that block spanin function (also see Fig.4.3B) **(A)** Cells lysogenic for phages carrying the following spanin alleles were grown in LB supplemented with 10mM MgSO4, induced at t=0 and monitored for culture mass ($A_{550}$): wt (open circle; ○), $Rz_{am}Rz1_{am}$ (filled circle; ●), *Rz* alleles with missense changes R91P or Q151R (filled triangle and diamond ;▲ and ♦), and *Rz1* allele with missense change I54N (filled square;■). (C) Expression of *Rz* and *Rz1* alleles. **(B)** Whole-cell samples were precipitated in cold 10% TCA, subjected to SDS-PAGE without reducing agent and analyzed by Western blot, as described in Methods. Both Rz and Rz1 were detected as covalent homodimers (as indicated on the right). A plus sign indicates the expression of the wt allele while a minus sign indicates the null (nonsense or *amber*) allele. The corresponding changes in the mutant alleles are indicated. The upper and lower sections are gels probed by anti-Rz and anti-Rz1 antibody, respectively.

In addition, spheroplast fusion was sensitive to the allelic state of the spanin subunit. Three non-functional spanin missense alleles, all of which confer absolute lysis defects *in vivo* without affecting accumulation of the gene product (Fig. 4.7), were tested

alone and in combination for fusion *in vitro*: I54N in Rz1; R91P and Q151R in Rz (Fig. 4.3B III – VI). Fusion was not affected by the introduction of either I54N into *imRz1* or R91P into *Rz*, suggesting that the fusion of cytoplasmic membranes *in vitro* is less demanding than the fusion of the IM and OM *in vivo*. Nevertheless, when both missense changes were present, fusion was abolished (Fig. 4.3B, compare set V with III and IV). Although fusion was not detected with this combination of alleles, there was an increase in spheroplast adhesion, defined as the number of visible interfaces between differentially-labeled spheroplasts (inset of Fig.4.3B). Indeed, in the double mutant experiment, spheroplast adhesions were elevated compared to others (Fig. 4.3B lower panel, compare set V with the rest). This suggests that the mutant I54N o-spanin and R91P i-spanin subunits retained the ability to form a complex but were blocked in the actual membrane fusion step. Interestingly, spheroplast adhesion was also elevated when Rz was surface exposed from both types of cells (Fig.4.5B). This is likely due to the ability of the periplasmic domain of Rz to form coiled-coil bundles (9). When we tested another i-spanin allele, $Rz_{Q151R,}$ we found that fusion was abolished irrespective of the allelic state of *imRz1* (Fig. 4.3B, compare set VI with the rest). With this *Rz* allele, persistence of adhesions could not be reliably assessed because the number of surviving GFP-labeled spheroplasts was significantly reduced, suggesting that the Q151R substitution at the extreme C-terminus of Rz is detrimental to the membrane. In any case, taken together, these results support the notion that the spanin complex is membrane-fusogenic, consistent with our model that the last step in the lysis pathway is spanin-mediated fusion between the IM and OM.

***The peptidoglycan meshwork inhibits spanin lytic function***

Since the spanin complexes form throughout the morphogenesis phase of the λ infection cycle, there must be some regulation that prevents premature membrane fusion, which would obviously destroy the integrity of the cytoplasmic membrane. The intact PG is a meshwork of glycosidic chains cross-linked by oligopeptides, forming covalent lacunae of ~ 2nm diameter (145). The simplest notion is that during the infection cycle, each spanin complex connecting the IM and OM is restricted to the lacuna of the intact PG within which it is formed (Fig. 4.2A, Accumulation panel, black hatching) and thus prevented from lateral diffusion. Liberation of these complexes by endolysin-mediated degradation of the cell wall would permit oligomerization of the complexes, which we have suggested to be the source of the free energy needed for membrane fusion (Fig. 4.2A, PG Degradation panel). To test the idea that it is the meshwork of the PG that blocks spanin function, we examined spanin-mediated lysis in conditions where the murein was depleted in an endolysin-independent fashion. To deplete the level of PG, we used an *E. coli dadX alr* strain. The *dadX* and *alr* genes encode the alanine racemases necessary for conversion of L-Ala to D-Ala, which is essential for the biosynthesis of PG (142, 146, 147). Removal of D-Ala from the medium blocks net PG synthesis and causes the growing auxotrophic cells to gradually lose their rod-like morphology (Fig. 4.8), starting with a localized bulging phenotype similar to cells treated with β-lactam antibiotics or vancomycin (148, 149). To prevent endolysin function, we constructed a λ$S_{\text{am}}$ lysogen of this strain. The $S_{am}$ null holin allele blocks endolysin function by

restricting the muralytic enzyme to the cytoplasm but has no effect on the expression of the spanin genes or spanin complex formation (8).

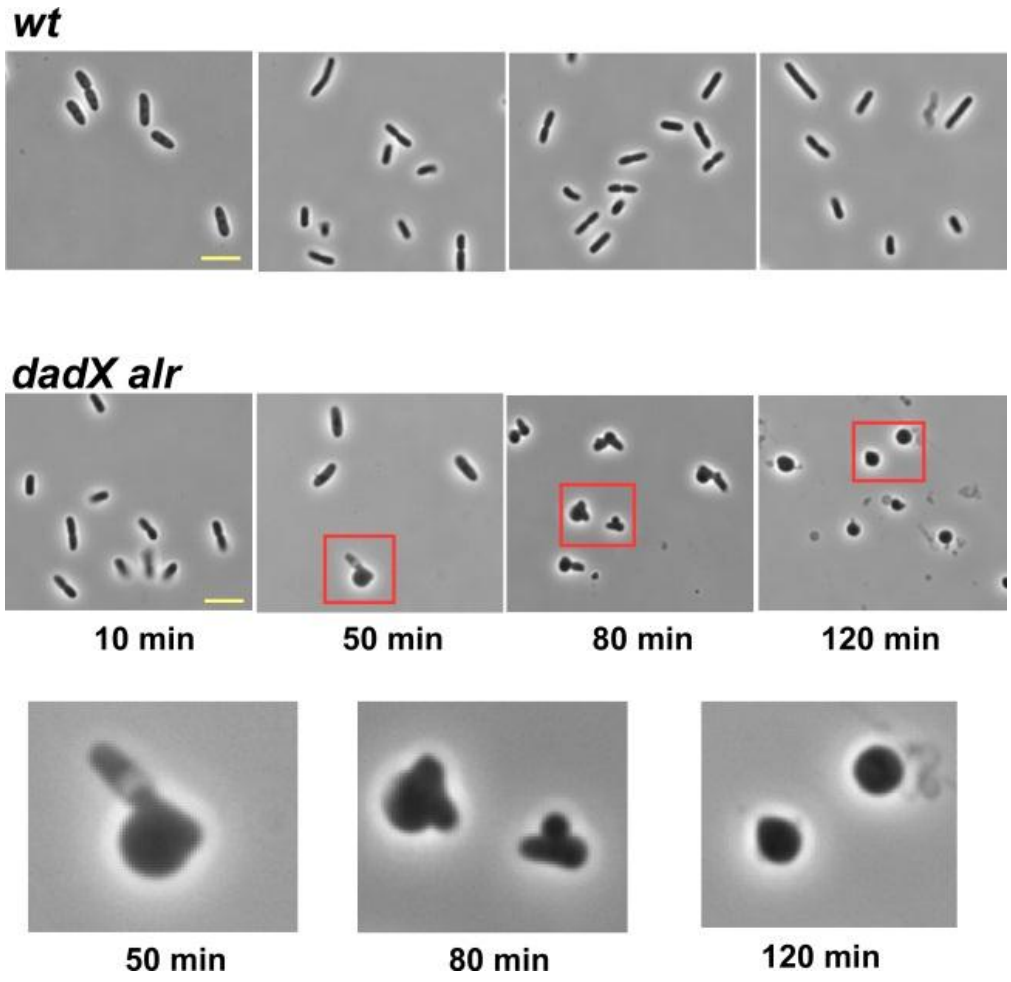



Figure 4.8 Morphological changes after depletion of D-Ala. Uninduced logarithmic cultures of *E. coli* $\lambda S_{am7}$ lysogenic cells, either prototrophic (wt; above) or auxotrophic (*dadX alr*; below and boxed insets) for D-Ala were suddenly depleted of D-Ala at t=0 (see methods). The morphology of the cells was monitored by phase-contrast microscopy at indicated times. Scale bar (yellow) = 5μm.

When we induced this lysogen after starving for D-Ala, dramatic lysis of the spherical cells was observed beginning ~ 15 min after induction (Fig. 4.9A; filled squares). This lysis required both the i-spanin Rz (open squares) and the o-spanin Rz1 (open triangles), and was blocked by re-addition of D-Ala (open circles) at the time of induction. By comparison, lysis was not observed in cells prototrophic for D-Ala (filled triangles, crosses). In another experiment, at the time of induction, we abruptly depleted D-Ala from the culture (Fig. 4.9B). Under these acute conditions, in the absence of spanin function, the cells undergo morphological changes characteristic of cell wall defects (Fig. 4.9C, upper panels) and some eventually lyse beginning at ~ 100-110 min (Fig. 4.9B, open squares). In contrast, in the presence of the functional spanins, lysis occurs much earlier, at ~45-50 min, and more rapidly (Fig. 4.9B, filled squares). Importantly, many of the lysed cells retained their rod-shaped morphology, suggesting that the spanin-mediated lysis occurred even before the murein was sufficiently depleted to cause gross morphological changes (Fig. 4.9C, lower panels). These experiments support the notion that the PG network acts as a negative regulator of spanin lytic function.

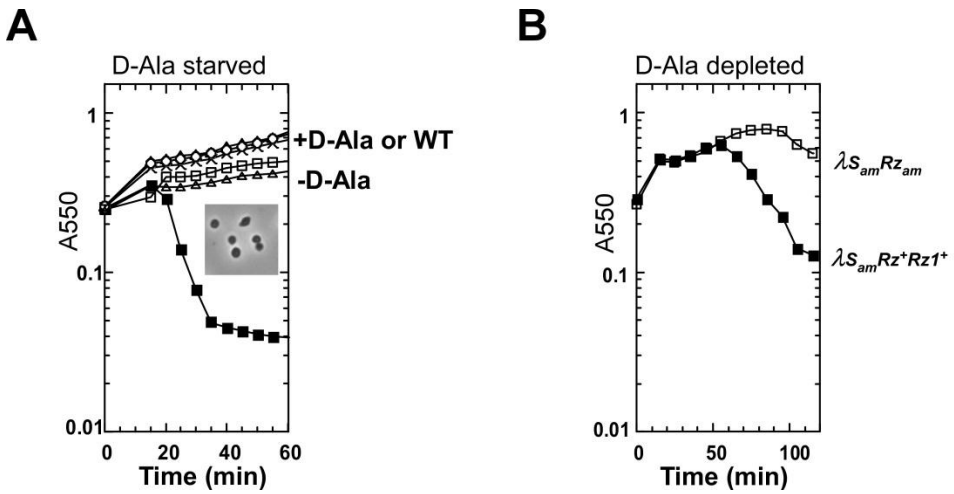



Figure 4.9 Spanin function is negatively regulated by the PG. **(A)** Cells lysogenic for $\lambda S_{am}$ $Rz^{+}Rz1^{+}$ and auxotrophic for D-Ala were grown in LB in the absence of D-Ala, induced at t=0 and monitored for culture mass ($A_{550}$). The induced culture underwent sudden lysis beginning at t= ~15 min (filled squares, ■). This lysis was blocked in the wt host (D-Ala prototroph; filled triangles; ▲) or in the D-Ala auxotroph supplemented with 150μM D-Ala (open circles; ○). Lysis was also blocked in the starved auxotrophs if either *Rz* (open squares; □) or *Rz1* (open triangles; Δ) was absent. In these conditions, the cells maintained their spherical morphology (see inset for image of $Rz^{-}$ cells; $Rz1^{-}$ cells were identical). **(B)** D-Ala auxotrophic cells lysogenic for $\lambda S_{am}$ were induced and suddenly depleted of D-Ala at time = 0 (See Methods). Onset of lysis was observed ~ 60 min after induction of $\lambda S_{am}Rz^{+}Rz1^{+}$ (filled squares; ■), whereas only a slight decrease in $A_{550}$ was observed for the isogenic $Rz_{am}$ prophage (open squares; □) **(C)** Morphology of cells from the experiment of panel B were imaged by phase-contrast microscopy at indicated sampling times after induction. Arrowheads indicate swollen morphology of the D-Ala-deprived cells in the top images and rod-shaped cells that have undergone lysis in the bottom images. Scale bar (yellow) = 5 μm.

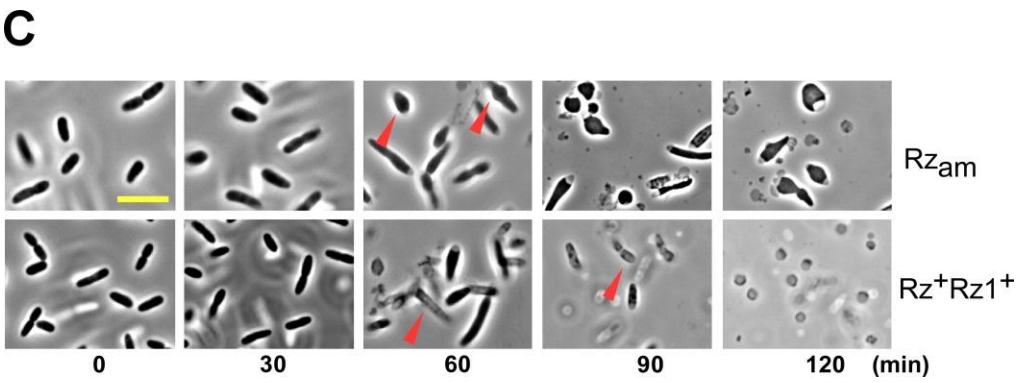



Figure 4.9. Continued

## Discussion

### *Membrane fusion as the last step in lysis*

Here we have presented results that phage spanins can mediate efficient membrane fusion in spheroplasts. Moreover, spheroplast fusion by the Rz-Rz1 system of λ is blocked by missense mutations that also block the last step of λ lysis, without affecting accumulation of the spanin subunits. These results strongly support the model that the last step in lysis of Gram-negative hosts is topological, rather than degradative (8, 150). In other words, the final barrier to phage progeny release, the OM, is circumvented through fusion with the IM.

In addition, the experiments with the D-Ala prototroph indicate that the spanin-mediated lytic step, which we suggest is IM-OM fusion, is not directly coupled to any

other phage lytic function, including endolysin-mediated degradation of the cell wall. Instead, the spanin complexes can cause fusion upon partial depletion of the PG imposed by starvation for D-Ala, even when the PG is sufficiently intact to preserve the rod-shaped cell morphology. These results support our interpretation of the dramatic "local blow-out" morphology characteristic of λ lysis (10). We have suggested that endolysin molecules escaping through the micron-scale IM holes formed by the triggered holin population initially degrade the patch of the PG over the hole. This interpretation is based on the observation that, in the absence of spanin function, the envelope deforms beginning at a confined point and is then transformed into a spherical shape by progressive deformation along the long axis of the cell, presumably reflecting diffusion of the muralytic endolysin molecules away from the initial hole. Thus, the lytic blow-out occurs when sufficient spanin complexes are liberated within the degraded area of the PG. It would be essential to avoid premature fusion before the gaps in the cell wall are large enough to allow unimpeded passage of the progeny virions. From fluorescence images of cells expressing a functional GFP-tagged Rz and wt Rz1, the distribution of spanin complexes appears to be relatively uniform throughout the envelope (10). Thus, using quantitative Western blotting, it should be possible to quantify the density of spanin complexes in the envelope at the time of triggering and thus establish an estimate of the number of these complexes needed to effect catastrophic IM-OM fusion. Moreover, IM-OM fusion events and/or vesicularization should be directly visualizable using cryo-electron microscopy and tomography, as were the micron-scale holin lesions (42).

***Mechanism of membrane fusion***

The results presented here allow inferences to be drawn about the mechanism of IM-OM fusion. First, the fact that *Rz-Rz1* mutations that block lysis also block fusion *in vitro* without affecting accumulation or subcellular localization of the spanin subunits suggests that the *in vitro* spheroplast fusion and the IM-OM fusion are mechanistically similar, if not identical. In particular, the combination of the $Rz_{R91P}$ i-spanin and $Rz1_{I54N}$ o-spanin alleles, despite having no effect on accumulation, not only blocks both lysis and spheroplast fusion but also results in the accumulation of adhered spheroplasts. This suggests that these alleles are defective in a final step in membrane fusion, beyond bringing the target membranes into contact. Indeed, studies of other membrane fusion systems have suggested that, after the bilayers are brought together by a protein complex, one or both of the opposed leaflets must be perturbed to initiate the fusion event (74, 76, 80). In general, opposing membranes are brought into close proximity, often by a protein with propensity to form coiled-coil structures. For example, in Golgi vesicular fusion, SNARE proteins anchored in both membranes undergo coiled-coil interactions that bring the bilayers into apposition (106, 151). Moreover, in some fusion systems, a protein domain with intrinsic ability to destabilize a membrane leaflet is required; for example, the fusion peptide p15 of orthorheovirus has a polyproline helix shown to be essential for viral entry (152). The spanin system has features reminiscent of both these examples, in that the i-spanin Rz has conserved coiled-coil helical domains that occupy nearly all of the soluble segment. As noted previously by Bryl et al.(132),

the mature o-spanin Rz1 has 25% proline content, including a conserved run of five proline residues, which could confer membrane-destabilizing character. Indeed, the lower level of fusion detected in the homotypic situations (i.e., with Rz or imRz1 in both cells; Fig. S3) suggests that both proteins have some fusogenic character. This observation suggests an attractive rationale for the evolution of the 2CS spanins in which ancestral phages of Gram-negative hosts had less efficient primordial versions of either the i-spanin or the o-spanin; a phage that acquired both might have a significant fitness advantage, after which evolutionary pressure might result in the interaction of the two subunits and the clustering of their genes.

Despite these functional homologies, the fact that in the case of the $Rz_{I54N}$ and $Rz1_{R91P}$ lysis-defective alleles each retain undiminished spheroplast fusion activity but exhibit a synthetic defect (i.e., the presence of both blocks fusion), suggests that IM-OM fusion is more demanding than IM-IM fusion. This is not surprising, considering our previous finding that in the absence of the PG the OM is capable of withstanding the internal osmotic pressure of the cell (10). This conclusion has been recently confirmed by other approaches (149). Thus the OM has significantly more tensile strength than the IM, which is rapidly destroyed in conditions of osmotic imbalance (133, 153). In any case, the facile genetics of λ and phage in general should provide a powerful platform for dissecting the molecular basis of membrane fusion.

***Energetics of spanin function***

Since spanin-mediated fusion occurs in the context of the periplasm where nucleotide phosphate energy carriers are not available, it is of interest to consider the source of the free energy for spanin-mediated membrane fusion. Although the i-spanin has a TMD that spans the IM and thus possesses a cytoplasmic N-terminus, the cytoplasmic domain is not large enough to be involved directly in ATP interactions. Another potential source of energy is the membrane potential, which is ~0.2 V in *E. coli* under standard conditions (154). However, spanins have been shown to function normally not only after holin triggering, which collapses the membrane potential (41), but also in the absence of holin function under conditions where the PG is degraded by a secreted endolysin (8). In the latter case, the energized state of the cytoplasmic membrane is undisturbed. The result reported here, showing that cyanide had no effect on the spanin-mediated spheroplast fusion, is consistent with that finding. By extension, since the energization of the IM is irrelevant to spanin function, so is energy transduction from the IM to the OM via the TonB-dependent and TolQ-dependent pathways (155). We favor the notion that the free energy for fusion results from conformational changes at the secondary, tertiary and quaternary levels after liberation of the spanins from the entrapping meshwork of the PG. Dramatic increases in complex size and alpha-helical content have been demonstrated for the formation of equimolar complexes of sRz and sRz1, the purified soluble periplasmic domains of the i-spanin and o-spanin subunits (9). Nothing is yet known about the behavior of the soluble domain of the u-spanin. The

presence of predicted β-strands instead of the helical and coiled-coil structures that are characteristics of the i-spanin, suggests the conformational change would be completely different.

***Evolutionary implications***

It is striking that both the i-spanin/o-spanin system and the u-spanin system both effect membrane fusion, despite the dissimilar topology and structure. This would appear to be a clear case of convergent evolution to a required function, suggesting that the topological solution to the problem of the OM is at least highly favored, if not required. In this perspective, it should be noted that the phages of the Gram-positive mycolata face a similar barrier to dispersal of the progeny phages: the mycolic-acid rich outer layer. Like all dsDNA phages, the lytic systems used by these phages are dependent on holins and endolysins for attacking the cytoplasmic membrane and PG, respectively (156, 157). However, instead of encoding spanins, these phages employ a second degradative enzyme, designated as LysB, which has been shown to hydrolyze mycolic acid esters(156, 158, 159). Thus in these cells, the holin effects the release of the endolysin and LysB. Obviously, fusion of the cytoplasmic membrane and the mycolate layer would not be structurally feasible, since the outer barrier is not a comparable bilayer. Nevertheless, the question arises why a similar enzymatic alternative would not be an option for phages of Gram-negative hosts? In fact, periplasmic phospholipases that could attack the OM have been described (160). We

suggest that the answer might lie in the "all or nothing" and rapid kinetics of lysis. Both types of spanin systems allow pre-positioning of the lytic proteins in the envelope throughout the infection cycle, without detrimental effect on the health of the host. Upon holin triggering, which marks the end of biosynthesis in the host, there is sudden degradation of the PG by the endolysins, leading to immediate local membrane fusion and thus then rapid and violent expulsion of the cytoplasmic contents. In contrast, it is difficult to conceive how a phospholipases could bring about localized damage to membranes, given the rapid lateral diffusion of phospholipids (161). Nevertheless, at least ~15% of phages of Gram-negative hosts do not encode either the o-spanin subunit of a 2CS spanin or a u-spanin, based on the absence of any genes containing the lipobox signal required for lipoprotein processing (7, 65). It is possible that in these cases, there is an enzyme for OM disruption or, more likely, a spanin that connects the OM and IM without using an OM lipoprotein determinant.

***Next steps***

There is immense diversity among spanins, including 2CS spanin genes that have different genetic architectures from the embedded structure of lambda *Rz-Rz1* (Fig. 4.1A and Fig.1.6 in Chapter I) ) (7). The massive evolutionary diversity should empower bioinformatic analysis once domains can be tentatively assigned to functions in the membrane fusion pathway and, potentially, will permit domain-interaction analysis through the construction of hybrid spanins. Combined with, the facile genetics of λ and

phage in general, the spanin system should provide a powerful platform for dissecting the molecular basis of membrane fusion. In addition, some phages of Gram-negative hosts, instead of having an i-spanin/o-spanin gene pair, encode a unimolecular spanin (u-spanin), that effects OM disruption during lysis (see Fig. 1.4 in Chapter I) (7). The prototype u-spanin is *gp11* of the paradigm phage T1. Gp11 has an OM lipoprotein signal at the N-terminus and a C-terminal domain that spans the IM (inset of Fig.S7). The periplasmic domain of *gp11* lacks the characteristic features of the 2CS spanins, including intermolecular disulfide bonds and predicted coiled-coil helical structure (9, 56) and is instead predicted to be dominated by β-strand structures (7). In work to be described elsewhere, *gp11* has also been shown to promote spheroplast fusion, which should allow comparative genetic and biochemical studies to focus on the molecular features essential for membrane fusion (Rohit Kongari, unpublished).

Finally, it should be noted that the capacity for spheroplast fusion may allow the development of entirely new approaches to synthetic and systems biology. There is nothing inherently lethal about spanin function in *E. coli*, since live, stable fusion spheroplasts can be generated efficiently. Thus large DNA molecules or even chromosomes that would be intractable for transfer from one *E. coli* cell to another could potentially be introduced into a novel cytoplasm if the donor and recipient cells were equipped with the appropriate i-spanin and redirected o-spanin pair. Given the rich diversity and ubiquity of phages in Gram-negative hosts, there should be no barrier to finding appropriate spanins for any bacterial species, even if the λ Rz-Rz1 proteins are not efficiently expressed. Moreover, both spanin systems could be used directly in

Gram-positive hosts. In fact, native Rz1 could be used in *B. subtilis*, for example, because there is no Lol system and the mature Rz1 lipoprotein should be automatically displayed on the outside of the cytoplasmic membrane. Thus spanin-mediated bacterial "hybridoma" construction may be possible, limited only by the ability to form stable spheroplasts.

# CHAPTER V

# CONCLUSIONS AND FUTURE DIRECTIONS

The results presented here further develop a new perspective on phage lysis. Classically, it was thought that λ lysis required only the holin and endolysin (140). Degradation of the PG was thought to be sufficient to cause lysis of Gram-negative hosts, because the OM was not considered to be a significant physical barrier for phage release. Recently, this notion was shown to be incorrect and that in fact phage λ requires the spanins, Rz and Rz1, for disrupting the OM. That is, disruption of the OM is the third step of lysis (5, 6, 8). Not surprisingly, spanins are ubiquitous in the genomes of dsDNA phages that grow on Gram-negative hosts (7). It was unexpected to find that there are two different types of spanins, a 2CS spanin and a 1CS spanin, both of which are functionally equivalent in acting to disrupt the OM (7). However, the mechanism of spanin function in both types remained unknown. The work that served as the foundation of this dissertation was focused on understanding the role of the λ Rz and Rz1 in OM disruption (8, 9, 68). The work presented in this dissertation provides further insight into the mechanism of λ Rz and Rz1 function and has provided a framework for mechanistic studies. The key results are summarized below, along with perspectives of what to do next.

***Essentiality of the spanins for lambda lysis***

Initially, the *Rz* and *Rz1* genes of λ were designated as auxiliary lysis genes because both genes were dispensable for lysis in liquid cultures unless the OM was stabilized by divalent cations. The conditional phenotype assigned to *RzRz1* was fundamentally changed when Joel Berry began observing lysis by video-microscopy (10, 68). One conclusion was that the presence of divalent cations had nothing to do with the lysis process other than enhancing the stability of the OM. This also implied that the OM has sufficient tensile strength to resist the osmotic imbalance following the collapse of the IM and PG. This finding led to a hypothesis that mechanical shearing forces present in the shaker flasks must be complementing the $Rz_{am}Rz1_{am}$ lysis defect. In Chapter II, we re-investigated the stability of spherical cells which resulted from inducing *RzRz1* null mutants and found that Rz and Rz1 are absolutely required for the disruption of the OM in liquid cultures, if shearing forces are avoided (Fig.2.2). In addition, we found that other physiological parameters such as temperature and growth phase can also influence the stability of the spherical cell resulting from spanin-defective lysis (Fig.2.2).

Although the requirement for spanin function has been demonstrated for lambda in infections of *E. coli* in complex medium, it is not known whether the requirement is maintained so strictly in other conditions, especially minimal media or in scenarios designed to be similar to the enteric milieu. In addition, although spanins appear to be found in nearly all phages of Gram-negative hosts, lambda is an unusually "moderate" phage in terms of its near total lack of interference in host metabolism and gene

expression. It could be that, phages that are more active in terms of abolishing host gene expression and destroying the host chromosome might not face as robust an OM barrier and thus might not require spanin function to achieve phage release. Formally, it would be useful to know if the requirement for spanins would be the same. It is important to note that there are other classes of endolysins enzymatically distinct from the lambda transglycosylase (162). Finally, in going forward, the requirement for spanin function in lysis should be established for other phages of Gram-negative hosts, especially for phages with the spanin genes in the two other architectures: overlapped and separated. It should be noted that no spanin genes were identified as essential, in terms of the amber mutant hunts that defined all the other essential genes of any of the genetically-facile paradigm phage systems of *E. coli* and *Salmonella*.

### *Requirement for intermolecular disulfide linkage in at least one subunit of the spanin complex*

Another key finding prior to the beginning of this work was that both Rz (i-spanin) and Rz1 (o-spanin) exist as a disulfide linked homodimers. These disulfide linkages were all homotypic: C99 to C99 and C152 to C152 in the i-spanin and C29 to C29 in the o-spanin. Chapter II highlights a further investigation toward the importance of multiple disulfide linkages for spanin function. Our genetic analysis along with observations of lysis morphology indicated that the disulfide linkages responsible for homodimerization of the i-spanin and the o-spanin are important for function (Fig.2.3, Fig. 2.7, and Fig. 2.8). The intriguing finding was that out of three native disulfide

linkages in the spanin complex, one must be near the C-terminal end of either the i-spanin **or** the o-spanin. It became apparent from this experiment that a disulfide linkage in the i-spanin can compensate for the absence of a disulfide linkage in the o-spanin or vice versa. This is a surprising and provocative result that suggests that at some point in the lysis pathway, there comes into play lateral forces that would disrupt the spanin heterotetrameric complex along its homotypic interface (Fig. 5.1). To our knowledge, intermolecular disulfide bonds have not been shown to be essential for any protein function in prokaryotes, much less a requirement that allows either subunit of a complex to have the covalent linkage.

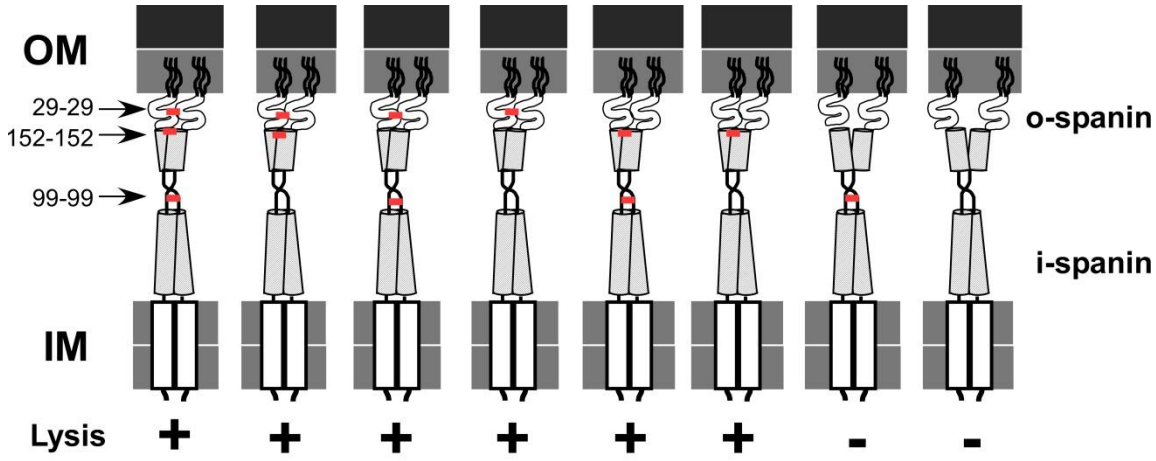



Figure 5.1 A disulfide linkage in the heterotypic interface is necessary for spanin function. Far left is the wild type spanin complex with three disulfide bonds (indicated by red). Seven out of eight complexes lack either one, two, or all disulfide linkages by Cys to Ser substitution in different combinations. The six functional versions are indicated by a"+" sign in the bottom. The position of disulfide linkages is indicated in the left and the i-spanin and o-spanin are indicated in the right.

Crystal or solution structures for covalent homodimers of sRz and sRz1 (soluble periplasmic domains of each spanin subunit; (9)) should be the next goal in unraveling the mystery of the disulfide bond requirement. This will require a re-examination of the conditions for overproduction of the two periplasmic subunits in the cytoplasmic over-expression system used for production of the Cys-less variants. The major challenge will be the formation of the intermolecular disulfide bonds, since there is no natural disulfide bond-formation system in the *E. coli* cytoplasm. Since the Cys-less sRz homo-dimerizes spontaneously *in vitro*, it is possible that the intermolecular disulfide bonds could be formed by simple exposure of the purified wild type sRz homodimer to oxidizing conditions. The sRz1 disulfide bond may be more problematical; however, this laboratory has had success using over-expression strains engineered to promulgate disulfide bond formation in the cytoplasm (163, 164). A biochemical/biophysical approach, including scanning calorimetry, aimed at comparing the formation and stability of the homodimers and hetero-tetrameric complexes by the parental and Cys mutant proteins could reveal the molecular basis for the disulfide bond requirement.

***Covalent maturation of the i-spanin and o-spanin***

We investigated the molecular pathway required for the covalent maturation of both i-spanin and o-spanin. Our results indicated that the i-spanin undergoes an extended conformation with two intermolecular disulfide linkages from an intramolecular disulfide-linked hairpin conformation (Fig.2.4 and Fig.2.5). The covalent maturation

process is assisted by host Dsb system. On the other hand, homodimerization of the o-spanin is dependent on the host Dsb system as well as on the spontaneous oxidative process, after the i-spanin provides a template for the correct alignment of two o-spanin molecules (Fig.2.6). Overall, the hetero-tetramer complex consists of a i-spanin dimer and a o-spanin dimer, constituting the final complex that interconnects the IM and OM (8).

The results described here suggest that the spanin complex assumes two states: the extended state and the hairpin state. TEM images of the sRz-sRz1 complex are the only structures that are currently available. These structures support the notion that spanins forms higher order oligomers, once liberated from the constraints of the PG. Single particle analysis indicated the average size of the spanin complex to be ~25 nm in length and ~4 nm in width, although there was some heterogeneity in the particle size. Since the sRz-sRz1 complexes were genetically altered to be thiol free, the TEM structures may not be reflective of the native spanin supramolecular complex. Structural study is of highest priority (See above).

***Occurrence of cysteines in the periplasmic domain of spanins***

A bioinformatic analysis suggested that in general, 2CS spanins contain at least one cysteine residue either in i-spanin, o-spanin, or both. However, the positions of cysteine residues were not conserved among different families (Table 2.2, Fig.2.9, and also see Fig.6 of Summer et al. (7)). This stands in stark contrast to u-spanins, which do

not contain any periplasmic Cys residues and thus lack disulfide linkage(s). We extended our investigation to other spanin equivalents to learn if the presence of and requirement for disulfide linkages among spanins is a general phenomenon. The system chosen was the *pseT.3-pseT.2* (separated i-spanin/o-spanin) genes of phage T4. The T4 i-spanin and o-spanin, respectively have zero and two Cys residues, both of which are required for spanin lytic function (Fig. 2.10 B). Moreover, SDS-PAGE analysis of samples done in reducing and non-reducing conditions indicated that the T4 o-spanin forms primarily an intramolecular disulfide bond, with only a fraction of the o-spanin protein in the disulfide-linked homodimer form (Fig. 2.10). Thus, the T4 spanin system maintains the requirement for Cys residues established for lambda, but the details of the requirement are different, perhaps not surprising in view of the fact that the T4 o-spanin has an additional C-terminal predicted alpha-helical domain, compared to lambda Rz. It remains to be determined whether the fraction of intermolecular disulfide bond-linked dimers represents the functional spanins, which may be in equilibrium via disulfide-bond isomerization with the o-spanin species carrying intra-molecular disulfide bond.

In general spanins are small proteins that are highly diverse. The diversity of spanins, along with the prevalence of cysteines in various positions of 2CS, could serve as a broad template to study the mechanism of disulfide bond isomerization. In light of what is already known about the **intra**molecular disulfide bond formation and isomerization in the protein folding pathway of other proteins (165, 166), spanins is going to be more enlightening in terms of **inter**molecular disulfide bond formation. Furthermore, it might help us understand the necessity of covalent linkages in spanins.

Moreover, the presence of cysteine residue in the potential spanin candidate could be an additional criterion during annotation of phage genome. In any case, structural studies on the T4 spanin subunits and complexes are warranted, as described above for the lambda proteins.

***Important domains and residues for spanin function***

In Chapter III, we analyzed a collection of Rz and Rz1 mutants to identify the residues and domains that are critical for spanin function. From this analysis, we learned that the periplasmic domain of Rz and Rz1 are important for function (Fig.3.3 and Fig.3.4). The analysis identified important spanin sub-domains: the predicted helical domains in Rz, the polyproline stretch in Rz1, and C-terminal ends of Rz and Rz1 that are important for function. Each single missense mutant that was defective for lysis was, however, recessive. There are two alternative explanations for this: 1) the spanin complex works as a higher order oligomer and within this oligomer, a mutant copy is unable to have dominant effect for function 2) a few functional complexes of spanin that are not poisoned by a mutant copy is sufficient for lysis. In addition, the Rz mutants that have been tested so far had the ability to form a complex with Rz1 (Fig.3.6). One interesting finding during the study of complex formation between Rz and Rz1 was that Rz1 can form a complex with the degradation product representing the N-terminal end of Rz. It is puzzling that none of the lysis-defective mutants of Rz are defective for complex formation *in vivo* and are dominant negative. This suggests that most of these

mutants may interfere with spanin function in steps following complex formation. It is possible that there are multiple steps involved in spanin function.

Analyzing the Rz1 mutants for the ability to form a complex with wild type Rz should be of highest priority. In addition, it is crucial to purify the sRz and sRz1 mutant protein separately and assess complex formation *in vitro*.

Our spanin mutant library is only partially saturated. Screening of more mutants could help us identify other important residues or domains that are critical for complex formation and spanin function. Recently, we have learned that using video-microscopy to observe single-cell lysis morphology, especially the kinetics of morphological transformations, can reveal partial-defective phenotypes of spanin. This method can be used to analyze site-directed mutants that do not block lysis, in order to identify residues and domains important for function.

***Suppressor mutants could resolve various steps of spanin function***

A method was developed to study spanin function by suppressor analysis. The idea was to use suppressors to map the interacting intermolecular and intramolecular interfaces. This approach was initially handicapped by the embedded genetic architecture of the λ spanins. To overcome this, we separated the *Rz* and *Rz1* genes in the λ chromosome, creating synthetic parental alleles in which the DNA nucleotide identity between the central region of *Rz* and the *Rz1* gene was abolished (Fig.3.7 and also see Chapter III for details). This allowed us to incorporate a mutation either in *Rz* or

*Rz1* without affecting both genes. As a first result, we were able to isolate a spontaneous suppressor for a Y147H mutation, which is near the C-terminal end of Rz. The compensatory mutation for this change was the intragenic change S20P, located near the periplasmic end of the membrane-anchoring TMD of Rz (Fig.3.8). Based on Western blotting and pull-down experiments, the $Rz_{Y147H}$ mutant exhibits normal covalent homodimer formation and forms the hetero-tetrameric complex with Rz1(Fig.3.5 and Fig.3.6). This suggests that the Y147H mutant is blocked at a downstream step in spanin function. The simplest interpretation of the intragenic suppression by S20P is that Y147H is compromised in forming a collapsing hairpin structure brining C-terminal end of Rz into close contact with the N-terminal end of the periplasmic domain, and that S20P restores that interaction.

Recently, an intergenic suppressor has been isolated for a I54N mutation which is at the C-terminal end of Rz1. The compensatory mutation for this change was RzR37H which is near the N-terminal domain Rz (J. Cahill, unpublished). This finding indicates that Rz1 has a secondary interaction site in addition to the primary interaction site at the C-terminal end of Rz. Although, it is important to perform a biochemical test to validate these results, the separation of i-spanin and o-spanin in the λ chromosome has allowed us to generate a suppressor library. Therefore, an attempt should be made to collect more suppressors against various mutants of *Rz* and *Rz1*. A large library of compensatory changes for different mutations could help us resolve interaction interfaces between different steps of spanin function. This in turn could lead us in understanding the spanin fusion mechanism.

***Regulation of spanin function by the peptidoglycan***

The hetero-tetrameric spanin complexes, unless they frequently dissociate at the heterotypic interface, would be trapped in the PG meshwork, unable to diffuse beyond its "home" lacuna formed by the polysaccharide chains and oligopeptide cross-links (Fig.4.2). It has been established that the degradation of PG is required for spanin function (8). In Chapter IV, we investigated the extent of PG depletion that is required for spanin function. By taking an advantage of host racemase mutants, we were able to deplete the PG without rapid degradation by lysozyme. Our results, although qualitative at this point, indicate that spanins do not require complete degradation of PG. Lysed cells which were depleted of PG precursor appear rod-shaped and translucent, which we attribute to spanin function (Fig.4.8). This reaffirms that spanin function is inhibited by PG and it also implies that spanin function does not require the global degradation of PG but can occur if the lacunae in the PG are sufficiently enlarged.

To understand the extent of PG depletion that is required for spanin function, structural study of the purified murein sacculus is needed. The murein sacculus is a giant macromolecule which makes it relatively easy to purify and scan the surface of the sacculus by electron microscopy and other techniques (31). We have observed that prolonged depletion of D-Ala makes the racemase mutant cell spherical, presumably due to the unavailability of PG precursor. A comparative structural study of peptidoglycan after lysis by spanins due to D-Ala depletion is required. In this case, cells will retain their rod shape (Fig. 4.9), which could help us understand the relation between change in

the structure of PG due to D-Ala depletion and spanin function. It is unclear how large of a gap in PG is required for spanin function.

***Membrane fusion by spanins***

The following points support the membrane fusion model of spanin function: **(a)** structural features that are reminiscent of fusogenic proteins such as predicted coiled-coils of i-spanins and high proline content in o-spanins (Fig.2.1), **(b)** the structural rearrangement due to interaction between the soluble domains of the i-spanin and o-spanin (9), **(c)** an extended confirmation of the Rz-Rz1 complex *in vivo* (8), **(d)** genetic evidence suggesting the possibility of assuming a hairpin structure (Chapter III), and **(e)** the requirement of minimal gap in the PG meshwork for spanin-mediated lysis (Chapter IV) and membrane vesicularization observed by Joel Berry. In Chapter IV, we developed a new method to test the membrane fusion model (Fig. 4.2). Using this method, we were able to demonstrate membrane fusion mediate by spanins (Fig.4.3, Fig.4.4). We have also shown that fusion can be blocked using i-spanin and o-spanin missense mutations that confer absolute defects in lysis but which do not affect accumulation or localization in the envelope *in vivo*. This suggests that the fusion capacity and the lytic pathway are linked.

The spheroplast fusion assay described in Chapter IV is at an early stage. We have been able to demonstrate that spanins can mediate fusion (Chapter IV), albeit the fusion efficiency is ~10%. In theory, if every single cell is potent for fusion then why is

fusion not happening in a higher percentage? In our initial trial, when we tried to fuse already spheroplasted cells, we did not observe fusion. We think this has to do with the fate of the spanin subunits during spheroplasting. Our simplest interpretation is that the i-spanin and o-spanin are free to diffuse laterally in the bilayer after removing PG. This causes the majority of spanin subunits to undergo lateral oligomerization. In the meantime, if i-spanins and o-spanins do not interact to form complexes, then membrane bilayers are not brought into close proximity for fusion. It is possible that by this time the i-spanin and o-spanin assumes a post-fusion conformation and can no longer promote fusion. Fusion is successful when the mixture of cells expressing i-spanins and o-spanins are spheroplasted together. The intuition here is that fusion is more efficient when i-spanins and o-spanins form a complex before the complete degradation of PG. It is, however, difficult to observe if fusion between two cells can occur prior to the completion of spheroplast formation. A microfluidic-based approach could be adopted to investigate how fusion occurs during spheroplasting because a microfluidic platform provides number of advantages during the study of microbial cells in a controlled condition (167). A microfluidic device could allow us to pre-position cells designed for fusion in an isolated chamber and the progression of fusion between two cells could be monitored in real time. Our attempt to regenerate fused spheroplast into a hybrid cells has not been fruitful, primarily due to the presence of EDTA and lysozyme in the resuspended buffer. A potential benefit of microfluidics is that it provides a better opportunity to exchange buffer containing EDTA and lysozyme used for spheroplasting with nutrient media, if desired to regenerate fused spheroplast or hybrids more

efficiently as discussed in Chapter IV. This will allow us to study the evolution of hybrid cells due to transfer of large DNA, if possible, between different bacterial species.

The spheroplast fusion assay can provide us with a better understanding about the topological changes that takes place during merging of membrane bilayers. It has been clear that diverse biological fusion reaction proceeds through the formation of a stalk-type intermediate called hemifusion diaphragm (HD) (168, 169). This key intermediate represents a lower energy state which is connected to the subsequent transient state in the fusion pathway. The transient state is the formation of fusion pore during the radial expansion of HD (170). However, the pathway for the formation of fusion pore from HD remains unclear. Several molecular modeling techniques have been used to describe this pathway (86, 171) and there are limited experimental designs to directly observe and characterize the HDs of biological membranes (172). To our knowledge, arresting the transient state of fusion pore opening is difficult. The inability to study this intermediate stage prevents us from completely understanding membrane fusion. The spheroplast fusion assay can be broadened to track intermediate processes by using of mutants that block different stages of fusion. Techniques such as Cryo-EM may benefit us in capturing different stages of topological changes that takes place between expansion of HD and opening of the fusion pore.

The results obtained from the spheroplast fusion assay described in Chapter IV are extrapolated to suggest the IM and OM fuse during phage lysis. However, spheroplast fusion is exclusively between two IM bilayers. Therefore, fusion between spheroplasts and membrane vesicles composed of OM constituents, also known as outer

membrane vesicles (OMVs), would be an interesting next step forward to reconstitute the fusion between IM and OM during phage lysis. OMVs have been successfully generated and could be enriched for this purpose (173).

In our experience, the spheroplast fusion assay is extremely delicate. Occasionally, we have observed a background level of fusion that is <0.5% in our negative controls. At this point we do not understand the precise nature of the low level of background fusion. Several factors that affect the behavior of spheroplasts, such as aggregation and spontaneous lysis, are described by Osborn et al. (133). The delicate nature of spheroplasts could be the main factor causing background fusion. Therefore, it is imperative to optimize the assay in order to find the conditions in which spheroplasts are most stable. Some stabilization techniques can be adopted from those that are described elsewhere (174).

***Coda***

In addition to the questions that are raised here, other questions will arise as we continue to study spanins and membrane fusion mechanism. Overall, the work presented in this dissertation demonstrates that phage lysis involves membrane fusion. Considering that globally $10^{23}$ lysis events happen per second, lysis by membrane fusion could be among the most frequent fusion process in nature.

## REFERENCES


1. Hendrix RW (2002) Bacteriophages: evolution of the majority. *Theor Popul Biol.* 61(4):471-480.

2. Suttle CA (2007) Marine viruses--major players in the global ecosystem. *Nat Rev Microbiol.* 5(10):801-812.

3. Rodriguez-Valera F, et al. (2009) Explaining microbial population genomics through phage predation. *Nat Rev Microbiol.* 7(11):828-836.

4. Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann ML, & Brussow H (2003) Phage as agents of lateral gene transfer. *Curr Opin Microbiol.* 6(4):417-424.

5. Young R, Way J, Way S, Yin J, & Syvanen M (1979) Transposition mutagenesis of bacteriophage lambda - new gene affecting cell lysis. *J. Mol. Biol.* 132(3):307-322.

6. 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(4-5):659-667.

7. Summer EJ, et al. (2007) *Rz* / *Rz1* lysis gene equivalents in phages of Gram-negative hosts. *J. Mol. Biol.* 373(5):1098-1112.

8. 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(2):341-351.

9. 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(10):1967-1977.

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

11. Freed EO (2004) Mechanisms of enveloped virus release. *Virus Research* 106(2):85-86.

12. Jahn R & Scheller RH (2006) SNAREs--engines for membrane fusion. *Nat Rev Mol Cell Biol.* 7(9):631-643.

13. Oursel D, et al. (2007) Lipid composition of membranes of Escherichia coli by liquid chromatography/tandem mass spectrometry using negative electrospray ionization. *Rapid Commun Mass Spectrom.* 21(11):1721-1728.

14. Miura T & Mizushima S (1968) Separation by density gradient centrifugation of two types of membranes from spheroplast membrane of Escherichia coli K12. *Biochim Biophys Acta* 150(1):159-161.

15. Papanastasiou M, et al. (2013) The Escherichia coli peripheral inner membrane proteome. *Mol Cell Proteomics.* 12(3):599-610.

16. Okuda S & Tokuda H (2011) Lipoprotein sorting in bacteria. *Annu Rev Microbiol.* 65:239-259.

17. Sankaran K & Wu HC (1994) Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J. Biol. Chem.* 269(31):19701-19706.

18. Hara T, Matsuyama S, & Tokuda H (2003) Mechanism underlying the inner membrane retention of Escherichia coli lipoproteins caused by Lol avoidance signals. *J. Biol. Chem.* 278(41):40408-40414.

19. Tokuda H & Matsuyama S (2004) Sorting of lipoproteins to the outer membrane in E. coli. *Biochim Biophys Acta* 1694(1-3):IN1-9.

20. Raetz CR & Whitfield C (2002) Lipopolysaccharide endotoxins. *Annual review of biochemistry* 71:635-700.

21. Leive L (1974) The barrier function of the gram-negative envelope. *Annu N Y Acad Sci.* 235(0):109-129.

22. Schindler M & Osborn MJ (1979) Interaction of divalent cations and polymyxin B with lipopolysaccharide. *Biochemistry* 18(20):4425-4430.

23. Schnaitman CA & Klena JD (1993) Genetics of lipopolysaccharide biosynthesis in enteric bacteria. *Microbiol Rev.* 57(3):655-682.

24. Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev.* 67(4):593-656.

25. Braun V & Wolff H (1970) The lethal λ *S* gene encodes its own inhibitor. *EMBO J.* 14(2):387-391.

26. Suzuki H, et al. (1978) Murein-lipoprotein of Escherichia coli: a protein involved in the stabilization of bacterial cell envelope. *Mol Gen Genet.* 167(1):1-9.

27. Matias VR, Al-Amoudi A, Dubochet J, & Beveridge TJ (2003) Cryo-transmission electron microscopy of frozen-hydrated sections of Escherichia coli and Pseudomonas aeruginosa *J. Bacteriol.* 185(20):6112-6118.

28. Mitchell P (1961) Approaches to the analysis of specific membrane transport. *Biological Structure and Function,* eds Goodwin TW & Lindberg O (Academic Press, New York), pp 581-603.

29. Heppel LA (1967) Selective release of enzymes from bacteria. *Science* 156(3781):1451-1455.

30. Holtje JV (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of Escherichia coli. *Microbiol Mol Biol Rev.* 62(1):181-203.

31. Vollmer W, Blanot D, & de Pedro MA (2008) Peptidoglycan structure and architecture. *FEMS Microbiol Rev.* 32(2):149-167.

32. Gan L, Chen S, & Jensen GJ (2008) Molecular organization of Gram-negative peptidoglycan. *Proc Natl Acad Sci USA* 105(48):18953-18957.

33. Young R (2014) Phage lysis: three steps, three choices, one outcome. *J. Microbiol.* 52(3):243-258.

34. 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.

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

36. Chang CY, Nam K, & Young R (1995) *S* gene expression and the timing of lysis by bacteriophage λ *J. Bacteriol.* 177:3283-3294.

37. Bienkowska-Szewczyk K, Lipiˈnska B, & Taylor A (1981) The R gene product of bacteriophage λ is the murein transglycosylase. *Mol Gen Genet.* 184(1):111-114.

38. Grundling A, Blasi U, & Young R (2000) Biochemical and genetic evidence for three transmembrane domains in the class I holin, lambda S. *J. Biol. Chem.* 275(2):769-776.

39. Graschopf A & Blasi U (1999) Molecular function of the dual-start motif in the lambda S holin. *Mol Microbiol.* 33(3):569-582.

40. Wang IN, Deaton JF, & Young R (2003) Sizing the holin lesion with an endolysin-βgalactosidase fusion. *J. Bacteriol.* 185(3):779-787.

41. White R, et al. (2011) Holin triggering in real time. *Proc Natl Acad Sci USA* 108(2):798-803.

42. Dewey JS, et al. (2010) Micron-scale holes terminate the phage infection cycle. *Proc Natl Acad Sci USA* 107:2219-2223.

43. Savva CG, et al. (2014) Stable micron-scale holes are a general feature of canonical holins. *Mol Microbiol.* 91(1):57-65.

44. Pang T, Park T, & Young R (2010) Mutational analysis of the S21 pinholin. *Mol Microbiol.* 76(1):68-77.

45. Park T, Struck DK, Deaton JF, & Young R (2006) Topological dynamics of holins in programmed bacterial lysis. *Proc Natl Acad Sci USA* 103(52):19713-19718.

46. Pang T, Park T, & Young R (2010) Mapping the pinhole formation pathway of S21. *Mol Microbiol.* 78(3):710-719.

47. Sun Q, et al. (2009) Regulation of a muralytic enzyme by dynamic membrane topology. *Nat Struct Mol Biol.* 16(11):1192-1194.

48. Kuty GF, Xu M, Struck DK, Summer EJ, & Young R (2010) Regulation of a phage endolysin by disulfide caging. *J. Bacteriol.* 192(21):5682-5687.

49. Taylor A, Benedik M, & Campbell A (1983) Location of the Rz gene in bacteriophage lambda. *Gene* 26(2-3):159-163.

50. Taylor A (1971) Endopeptidase activity of phage lamba-endolysin. *Nat New Biol.* 234(48):144-145.

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

52. Taylor A, Kedzierska S, & Wawrzynow A (1996) Bacteriophage lambda lysis gene product modified and inserted into Escherichia coli outer membrane: Rz1 lipoprotein. *Microb Drug Resist.* 2(1):147-153.

53. Casjens S, Eppler K, Parr R, & Poteete AR (1989) Nucleotide sequence of the bacteriophage P22 gene 19 to 3 region: Identification of a new gene required for lysis. *Virology* 171(2):588-598.

54. Ziermann R, Bartlett B, Calendar R, & Christie GE (1994) Functions involved in bacteriophage P2-induced host cell lysis and identification of a new tail gene. *J. Bacteriol.* 176(16):4974-4984.

55. Markov D, et al. (2004) P2 growth restriction on an rpoC mutant is suppressed by alleles of the Rz1 homolog lysC. *J. Bacteriol.* 186(14):4628-4637.

56. Berry JD, Rajaure M, & Young R (2013) Spanin function requires subunit homodimerization through intermolecular disulfide bonds. *Mol Microbiol.* 88(1):35-47.

57. Ahern SJ, Das M, Bhowmick TS, Young R, & Gonzalez CF (2014) Characterization of novel virulent broad-host-range phages of Xylella fastidiosa and Xanthomonas. *J. Bacteriol.* 196(2):459-471.

58. Hoiczyk E & Hansel A (2000) Cyanobacterial cell walls: news from an unusual prokaryotic envelope. *J. Bacteriol.* 182(5):1191-1199.

59. Omata T & Murata N (1984) Isolation and characterization of three types of membranes from the cyanobacterium (blue-green alga) Synechocystis PCC 6714. *Archives of Microbiol.* 139(2-3):113-116.

60. Schmidt W, Drews G, Weckesser J, & Mayer H (1980) Lipopolysaccharides in four strains of the unicellular cyanobacterium Synechocystis. *Archives of Microbiol.* 127(3):217-222.

61. Hoffmann C, Leis A, Niederweis M, Plitzko JM, & Engelhardt H (2008) Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci USA* 105(10):3963-3967.

62. Lupas A, Van Dyke M, & Stock J (1991) Predicting coiled coils from protein sequences. *Science* 252(5009):1162-1164.

63. O'Shea EK, Klemm JD, Kim PS, & Alber T (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* 254(5031):539-544.

64. Berger B, et al. (1995) Predicting coiled coils by use of pairwise residue correlations. *Proc Natl Acad Sci USA* 92(18):8259-8263.

65. Narita S & Tokuda H (2010) Sorting of bacterial lipoproteins to the outer membrane by the Lol system. *Methods Mol Biol.* 619:117-129.

66. 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):1-8.

67. Bartel PL, Roecklein JA, SenGupta D, & Fields S (1996) A protein linkage map of *Escherichia coli* bacteriophage T7. *Nat Genet.* 12(1):72-77.

68. Berry J (2010 (b)) The Final Step in Phage Lysis: The Role of the Rz-Rz1 Spanin Complex in the Disruption of the Outer Membrane. *Doctoral dissertation*.

69. Krupovič M, Cvirkaitė-Krupovič V, & Bamford DH (2008) Identification and functional analysis of the *Rz*/*Rz1*-like accessory lysis genes in the membrane-containing bacteriophage PRD1. *Mol Microbiol.* 68(2):492-503.

70. Fletcher G & Earhart CF (1980) Alterations in the cell envelope of Escherichia coli late in bacteriophage T4 infection. *Biochim Biophys Acta* 596(2):210-222.

71. Kucharczyk K, Laskowska E, & Taylor A (1991) Response of Escherichia coli cell membranes to induction of lambda cl857 prophage by heat shock. *Mol Microbiol.* 5(12):2935-2945.

72. Bryl K, Kedzierska S, Laskowska M, & Taylor A (2000) Membrane fusion by proline-rich Rz1 lipoprotein, the bacteriophage lambda Rz1 gene product. *EMBO J.* 267(3):794-799.

73. Merlini L, Dudin O, & Martin SG (2013) Mate and fuse: how yeast cells do it. *Open Biol.* 3(3):130008.

74. Harrison SC (2008) Viral membrane fusion. *Nat Struct Mol Biol.* 15(7):690-698.

75. Bonifacino JS & Glick BS (2004) The mechanisms of vesicle budding and fusion. *Cell* 116(2):153-166.

76. Jahn R, Lang T, & Sudhof TC (2003) Membrane fusion. *Cell* 112(4):519-533.

77. Sollner TH (2004) Intracellular and viral membrane fusion: a uniting mechanism. *Curr Opin Cell Biol* 16(4):429-435.

78. Fuhrmans M, Marelli G, Smirnova YG, & Muller M (2015) Mechanics of membrane fusion/pore formation. *Chem Phys Lipids* 185:109-128.

79. Wickner W & Schekman R (2008) Membrane fusion. *Nat Struct Mol Biol* 15(7):658-664.

80. Chernomordik LV & Kozlov MM (2008) Mechanics of membrane fusion. *Nat Struct Mol Biol.* 15(7):675-683.

81. Frolov VA, Shnyrova AV, & Zimmerberg J (2011) Lipid polymorphisms and membrane shape. *Cold Spring Harb Perspect Biol* 3(11):a004747.

82. Chernomordik LV & Kozlov MM (2003) Protein-lipid interplay in fusion and fission of biological membranes. *Annu Rev Biochem.* 72:175-207.

83. Markin VS & Albanesi JP (2002) Membrane fusion: stalk model revisited. *Biophys J.* 82(2):693-712.

84. Kozlovsky Y & Kozlov MM (2002) Stalk model of membrane fusion: solution of energy crisis. *Biophys J.* 82(2):882-895.

85. Martens S, Kozlov MM, & McMahon HT (2007) How synaptotagmin promotes membrane fusion. *Science* 316(5828):1205-1208.

86. Kuzmin PI, Zimmerberg J, Chizmadzhev YA, & Cohen FS (2001) A quantitative model for membrane fusion based on low-energy intermediates. *Proc Natl Acad Sci USA* 98(13):7235-7240.

87. Kozlov MM & Chernomordik LV (1998) A mechanism of protein-mediated fusion: coupling between refolding of the influenza hemagglutinin and lipid rearrangements. *Biophys J* 75(3):1384-1396.

88. Chernomordik LV & Kozlov MM (2005) Membrane hemifusion: crossing a chasm in two leaps. *Cell* 123(3):375-382.

89. Jackson MB & Chapman ER (2006) Fusion pores and fusion machines in Ca2+-triggered exocytosis. *Annu Rev Biophys Biomol Struct* 35:135-160.

90. White JM, Delos SE, Brecher M, & Schornberg K (2008) Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol* 43(3):189-219.

91. Earp LJ, Delos SE, Park HE, & White JM (2005) The many mechanisms of viral membrane fusion proteins. *Curr Top Microbiol Immunol* 285:25-66.

92. Barocchi MA, Masignani V, & Rappuoli R (2005) Opinion: Cell entry machines: a common theme in nature? *Nat Rev Microbiol* 3(4):349-358.

93. Hong W (2005) SNAREs and traffic. *Biochim Biophys Acta* 1744(2):120-144.

94. Uemura T, et al. (2004) Systematic analysis of SNARE molecules in Arabidopsis: dissection of the post-Golgi network in plant cells. *Cell Struct Funct.* 29(2):49-65.

95. Weimbs T, et al. (1997) A conserved domain is present in different families of vesicular fusion proteins: a new superfamily. *Proc Natl Acad Sci USA* 94(7):3046-3051.

96. Sutton RB, Fasshauer D, Jahn R, & Brunger AT (1998) Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 A resolution. *Nature* 395(6700):347-353.

97. Sollner T, Bennett MK, Whiteheart SW, Scheller RH, & Rothman JE (1993) A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell* 75(3):409-418.

98. Malsam J, Kreye S, & Sollner TH (2008) Membrane fusion: SNAREs and regulation. *Cell Mol Life Sci* 65(18):2814-2832.

99. Parlati F, et al. (1999) Rapid and efficient fusion of phospholipid vesicles by the alpha-helical core of a SNARE complex in the absence of an N-terminal regulatory domain. *Proc Natl Acad Sci USA* 96(22):12565-12570.

100. Collins KM & Wickner WT (2007) Trans-SNARE complex assembly and yeast vacuole membrane fusion. *Proc Natl Acad Sci USA* 104(21):8755-8760.

101. Ebine K, et al. (2011) A membrane trafficking pathway regulated by the plant-specific RAB GTPase ARA6. *Nat Cell Biol* 13(7):853-859.

102. Fasshauer D, Otto H, Eliason WK, Jahn R, & Brunger AT (1997) Structural changes are associated with soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptor complex formation. *J Biol Chem* 272(44):28036-28041.

103. Fasshauer D (2003) Structural insights into the SNARE mechanism. *Biochim Biophys Acta* 1641(2-3):87-97.

104. Zhao M, et al. (2015) Mechanistic insights into the recycling machine of the SNARE complex. *Nature*.

105. Podbilewicz B (2014) Virus and cell fusion mechanisms. *Annu Rev Cell Dev Biol.* 30:111-139.

106. Weber T, et al. (1998) SNAREpins: minimal machinery for membrane fusion. *Cell* 92(6):759-772.

107. Carr CM & Kim PS (1993) A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* 73(4):823-832.

108. Chan DC, Fass D, Berger JM, & Kim PS (1997) Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 89(2):263-273.

109. 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. *J Biol Chem* 287(5):3403-3414.

110. Cuff JA, Clamp ME, Siddiqui AS, Finlay M, & Barton GJ (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics* 14(10):892-893.

111. McDonnell AV, Jiang T, Keating AE, & Berger B (2006) Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* 22(3):356-358.

112. Miller JH (1992) *A short course in bacterial genetics: A laboratory manual and handbook for Escherichia coli and related bacteria* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

113. Hendrix RW (1983) *Lambda II* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.) pp vii, 694 p., 693 p. of plates (691 folded).

114. 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(25):5765-5766.

115. Rietsch A, Belin D, Martin N, & Beckwith J (1996) An in vivo pathway for disulfide bond isomerization in Escherichia coli. *Proc Natl Acad Sci USA* 93(23):13048-13053.

116. Tran TA, Struck DK, & Young R (2005) Periplasmic domains define holin-antiholin interactions in t4 lysis inhibition. *J. Bacteriol.* 187(19):6631-6640.

117. Bernhardt TG, Roof WD, & Young R (2002) The *Escherichia coli* FKBP-type PPIase SlyD is required for the stabilization of the E lysis protein of bacteriophage ϕX174. *Mol Microbiol.* 45(1):99-108.

118. Kadokura H & Beckwith J (2010) Mechanisms of oxidative protein folding in the bacterial cell envelope. *Antioxid Redox Signal.* 13(8):1231-1246.

119. Vertommen D, et al. (2008) The disulphide isomerase DsbC cooperates with the oxidase DsbA in a DsbD-independent manner. *Mol Microbiol.* 67(2):336-349.

120. Kadokura H, Katzen F, & Beckwith J (2003) Protein disulfide bond formation in prokaryotes. *Annu Rev Biochem.* 72:111-135.

121. Wang IN (2006) Lysis timing and bacteriophage fitness. *Genetics* 172(1):17-26.

122. Bull JJ, Pfennig DW, & Wang IN (2004) Genetic details, optimization and phage life histories. *Trends Ecol Evol.* 19(2):76-82.

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

124. Abedon ST (1989) Selection for bacteriophage latent period length by bacterial density: a theoretical examination. *Microb Ecol.* 18:79-88.

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

126. Lim JA, Shin H, Kang DH, & Ryu S (2012) Characterization of endolysin from a Salmonella Typhimurium-infecting bacteriophage SPN1S. *Res Microbiol.* 163(3):233-241.

127. Schmidt C, Velleman M, & Arber W (1996) Three functions of bacteriophage P1 involved in cell lysis. *J. Bacteriol.* 178(4):1099-1104.

128. Uson I, Patzer SI, Rodriguez DD, Braun V, & Zeth K (2012) The crystal structure of the dimeric colicin M immunity protein displays a 3D domain swap. *J Struct Biol.* 178(1):45-53.

129. Nagano K, et al. (2005) Trimeric structure of major outer membrane proteins homologous to OmpA in Porphyromonas gingivalis. *J. Bacteriol.* 187(3):902-911.

130. Klimke WA, et al. (2005) The mating pair stabilization protein, TraN, of the F plasmid is an outer-membrane protein with two regions that are important for its function in conjugation. *Microbiol.* 151(Pt 11):3527-3540.

131. Spudich GM, Fernandez D, Zhou XR, & Christie PJ (1996) Intermolecular disulfide bonds stabilize VirB7 homodimers and VirB7/VirB9 heterodimers during biogenesis of the Agrobacterium tumefaciens T-complex transport apparatus. *Proc Natl Acad Sci USA* 93(15):7512-7517.

132. Shao F, Bader MW, Jakob U, & Bardwell JC (2000) DsbG, a protein disulfide isomerase with chaperone activity. *J. Biol. Chem.* 275(18):13349-13352.

133. Osborn MJ & Munson R (1974) Separation of the inner (cytoplasmic) and outer membranes of Gram-negative bacteria. *Methods Enzymol* 31(A):642-653.

134. Hanson-Manful P & Patrick WM (2013) Construction and analysis of randomized protein-encoding libraries using error-prone PCR. *Methods Mol Biol.* 996:251-267.

135. Newman JR, Wolf E, & Kim PS (2000) A computationally directed screen identifying interacting coiled coils from Saccharomyces cerevisiae. *Proc Natl Acad Sci USA* 97(24):13203-13208.

136. Brandl CJ & Deber CM (1986) Hypothesis about the function of membrane-buried proline residues in transport proteins. *Proc Natl Acad Sci USA* 83(4):917-921.

137. Ulmschneider MB & Sansom MS (2001) Amino acid distributions in integral membrane protein structures. *Biochim Biophys Acta* 1512(1):1-14.

138. Vila JA, Baldoni HA, Ripoll DR, Ghosh A, & Scheraga HA (2004) Polyproline II helix conformation in a proline-rich environment: a theoretical study. *Biophys J.* 86(2):731-742.

139. Correia BE, et al. (2014) Proof of principle for epitope-focused vaccine design. *Nature* 507(7491):201-206.

140. Young R & Wang IN (2006) Phage Lysis. *The Bacteriophage*, eds Calender R & Abedon ST (Oxford University Press, Inc., Oxford), Second Ed, pp 104-126.

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

142. Strych U, Penland RL, Jimenez M, Krause KL, & Benedik MJ (2001) Characterization of the alanine racemases from two Mycobacteria. *FEMS Microbiol Lett.* 196(2):93-98.

143. Gründling A, Manson MD, & Young R (2001) Holins kill without warning. *Proc Natl Acad Sci USA* 98(16):9348-9352.

144. Yamaguchi K, Yu F, & Inouye M (1988) A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* 53(3):423-432.

145. Demchick P & Koch AL (1996) The permeability of the wall fabric of Escherichia coli and Bacillus subtilis. *J. Bacteriol.* 178(3):768-773.

146. Glauner B, Holtje JV, & Schwarz U (1988) The composition of the murein of Escherichia coli. *J. Biol. Chem.* 263(21):10088-10095.

147. Strych U & Benedik MJ (2002) Mutant analysis shows that alanine racemases from Pseudomonas aeruginosa and Escherichia coli are dimeric. *J. Bacteriol.* 184(15):4321-4325.

148. Huang KC, Mukhopadhyay R, Wen B, Gitai Z, & Wingreen NS (2008) Cell shape and cell-wall organization in Gram-negative bacteria. *Proc Natl Acad Sci USA* 105(49):19282-19287.

149. Yao Z, Kahne D, & Kishony R (2012) Distinct Single-Cell Morphological Dynamics under Beta-Lactam Antibiotics. *Mol Cell* 48(5):705-712.

150. To KH & Young R (2014) Probing the structure of the S105 hole. *J. Bacteriol* 196(21):3683-3689.

151. Rothman JE & Wieland FT (1996) Protein sorting by transport vesicles. *Science* 272(5259):227-234.

152. Moussa SH, Kuznetsov V, Tran TA, Sacchettini JC, & Young R (2012) Protein determinants of phage T4 lysis inhibition. *Protein Sci.* 21(4):571-582.

153. Birdsell DC & Cota-Robles EH (1967) Production and ultrastructure of lysozyme and ethylenediaminetetraacetate-lysozyme spheroplasts of *Escherichia coli*. *J. Bacteriol.* 93:427-437.

154. Felle H, Porter JS, Slayman CL, & Kaback HR (1980) Quantitative measurements of membrane potential in *Escherichia coli*. *Biochemistry* 19(15):3585-3590.

155. Braun V & Endriss F (2007) Energy-coupled outer membrane transport proteins and regulatory proteins. *Biometals.* 20(3-4):219-231.

156. Gil F, et al. (2008) The lytic cassette of mycobacteriophage Ms6 encodes an enzyme with lipolytic activity. *Microbiol.* 154(Pt 5):1364-1371.

157. Catalao MJ, Gil F, Moniz-Pereira J, & Pimentel M (2011) Functional analysis of the holin-like proteins of mycobacteriophage Ms6. *J. Bacteriol.* 193(11):2793-2803.

158. Payne K, Sun Q, Sacchettini J, & Hatfull GF (2009) Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. *Mol Microbiol.* 73(3):367-381.

159. Gil F, et al. (2010) Mycobacteriophage Ms6 LysB specifically targets the outer membrane of Mycobacterium smegmatis. *Microbiol.* 156(Pt 5):1497-1504.

160. Beveridge TJ & Kadurugamuwa JL (1996) Periplasm, periplasmic spaces, and their relation to bacterial wall structure: novel secretion of selected periplasmic proteins from *Pseudomonas aeruginosa*. *Microb Drug Resist.* 2(1):1-8.

161. Sackmann E, Trauble H, Galla HJ, & Overath P (1973) Lateral diffusion, protein mobility, and phase transitions in *Escherichia coli* membranes. A spin label study. *Biochemistry* 12(26):5360-5369.

162. Young R (2013) Phage lysis: do we have the hole story yet? *Curr. Opin. Microbiol.* 16(6):790-797.

163. Bessette PH, Aslund F, Beckwith J, & Georgiou G (1999) Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm. *Proc Natl Acad Sci USA* 96(24):13703-13708.

164. Levy R, Weiss R, Chen G, Iverson BL, & Georgiou G (2001) Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones. *Protein Expr Purif.* 23(2):338-347.

165. Creighton TE (1988) Toward a better understanding of protein folding pathways. *Proc Natl Acad Sci USA* 85(14):5082-5086.

166. Hatahet F, Boyd D, & Beckwith J (2014) Disulfide bond formation in prokaryotes: history, diversity and design. *Biochim Biophys Acta* 1844(8):1402-1414.

167. Rusconi R, Garren M, & Stocker R (2014) Microfluidics expanding the frontiers of microbial ecology. *Annu Rev Biophys* 43:65-91.

168. Chernomordik LV & Zimmerberg J (1995) Bending membranes to the task: structural intermediates in bilayer fusion. *Curr Opin Struct Biol* 5(4):541-547.

169. Risselada HJ, Bubnis G, & Grubmuller H (2014) Expansion of the fusion stalk and its implication for biological membrane fusion. *Proc Natl Acad Sci USA* 111(30):11043-11048.

170. Kozlovsky Y, Chernomordik LV, & Kozlov MM (2002) Lipid intermediates in membrane fusion: formation, structure, and decay of hemifusion diaphragm. *Biophys J* 83(5):2634-2651.

171. Markvoort AJ & Marrink SJ (2011) Lipid acrobatics in the membrane fusion arena. *Curr Top Membr* 68:259-294.

172. Nikolaus J, Stockl M, Langosch D, Volkmer R, & Herrmann A (2010) Direct visualization of large and protein-free hemifusion diaphragms. *Biophys J* 98(7):1192-1199.

173. Bernadac A, Gavioli M, Lazzaroni JC, Raina S, & Lloubes R (1998) Escherichia coli tol-pal mutants form outer membrane vesicles. *J. Bacteriol.* 180(18):4872-4878.

174. Tabor CW (1962) Stabilization of Protoplasts and Spheroplasts by Spermine and Other Polyamines. *J. Bacteriol.* 83(5):1101-1111.