SAR ENDOLYSIN REGULATION IN dsDNA PHAGE LYSIS OF GRAM-NEGATIVE HOSTS

A Dissertation

by

GABRIEL FAITH KUTY

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

DOCTOR OF PHILOSOPHY

December 2011

Major Subject: Biochemistry

SAR Endolysin Regulation in dsDNA Phage Lysis of Gram-Negative Hosts

SAR ENDOLYSIN REGULATION IN dsDNA PHAGE LYSIS OF GRAM-NEGATIVE HOSTS

A Dissertation

by

GABRIEL FAITH KUTY

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

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Ryland F. Young
Committee Members,	Michael Benedik
	James C. Hu
	James C. Sacchettini
Head of Department,	Gregory D. Reinhart

December 2011

Major Subject: Biochemistry

ABSTRACT

SAR Endolysin Regulation in dsDNA Phage Lysis of Gram-Negative Hosts.

(December 2011)

Gabriel Faith Kuty, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Ryland F. Young

SAR endolysins are a recently discovered class of muralytic enzymes that are regulated by dynamic membrane topology. They are synthesized as enzymatically inactive integral membrane proteins during the phage infection cycle and then are activated by conformational remodeling upon release from the membrane. This topological duality depends on N-terminal SAR (Signal-Anchor-Release) domains, which are enriched in weakly hydrophobic residues and require the proton motive force to be maintained in the bilayer. The first SAR endolysin to be characterized was P1 Lyz, of phage P1. Its activation requires a disulfide bond isomerization involving its catalytic Cys initiated by a free cysteine thiol from the newly-liberated SAR domain. A second mode of disulfide bond regulation, as typified by Lyz^{103} of the *Erwinia Amylovora* phage ERA103, has been demonstrated. In its membrane bound form, Lyz^{103} is inactivated by a disulfide that is formed between cysteine residues flanking a catalytic glutamate.

A second class of SAR endolysins, typified by R^{21}, the lysozyme of the lambdoid phage 21, does not require disulfide bond isomerization for activation. Rather, these proteins are dependent on the release of the SAR domain for proper folding of the

catalytic cleft. Bioinformatic analysis indicates that the regulatory theme of R^{21} is common in the SAR endolysins of dsDNA phages. Furthermore, bioinformatic study of endolysins of dsDNA phage of Gram-negative hosts revealed two new classes of SAR endolysins that are not homologous to T4 gp*e,* as all SAR endolysins were once thought to be. SAR endolysins were found in nearly 25% of sequenced dsDNA phages of Gram-negative hosts including 933W, which is involved in the release of Shiga toxin from EHEC strain EDL933. An inhibitor study against the SAR endolysin of 933W, R^{933W}, was performed using a custom compound library in a high through-put, *in vivo* lysis assay. Of nearly 8,000 compounds screened, one compound, designated 67-J8, inhibited lysis but not growth. *In vivo* and *in vitro* experiments show that the compound has no effect on R^{933W} activity, accumulation, or secretion. *In vivo* experiments suggest that 67-J8 increases the proton motive force, thereby presumably retaining the SAR domain in the membrane.

DEDICATION

This dissertation is dedicated to my Grammy and Poppy for their endless love, advice, and encouragement. Grammy, you are truly an inspiration and Poppy, you taught me to love life and to reach for the stars. Thank you.

ACKNOWLEDGEMENTS

I want to thank my P.I., Dr. Ry Young, for the opportunity to learn from him and for his guidance. The Young Lab has been the best environment in which to grow as a scientist. The 'round-table' discussions and random brainstorming sessions were absolutely invaluable to me. I would like to thank every member of the Young Lab, past and present, for their input and insightful discussions on this project. I would also like to thank my committee members, Dr. Benedik, Dr. Hu, and Dr. Sacchettini for their constructive comments and advice. A special thanks to Daisy Wilbert for her clerical assistance, general helpfulness, and her ear.

I want to thank my friends in the department with whom I have the privilege of taking this ride. I will never forget our Friday afternoon 'band practice' sessions. I also need to thank my friends outside the department, especially Rachel, who has been such a true friend, the best I could ever ask for.

To my wonderful Mom and Dad: you gave me the best childhood, full of love, support, and encouragement. The lessons you taught me have brought me to this point in my life. I can never thank you enough. To my Grammy and Poppy: you know I couldn't have done this without you! I also thank my Grandma Irene for her love, laughter, and warm smiles. My extended family at Congregation Beth Shalom, especially Carol and Manny Parzen and Susan Miller, have also enriched my life. Last, but certainly not least, is my Tony who has been my rock; thank you for your unconditional love and support.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Phage lysis timing

In order for a progeny phage to escape a Gram-negative cell, it needs to overcome the three barriers that constitute the envelope: the cytoplasmic or inner membrane (IM), the peptidoglycan (PG) layer, and the outer membrane (OM). For double-stranded DNA (dsDNA) phages, this is accomplished by lysis of the host cell. In addition, the lytic event must be subject to temporal regulation, as the timing of phage lysis is extremely important. If lysis occurs before the end of the eclipse period during which phage components are synthesized and assembled, the infection produces no progeny. If lysis occurs too late, the progeny may miss their opportunity for exponential growth. Based on ecological theory, it has been proposed that there are two major factors contributing to the evolution of lysis timing: host density and host quality (1, 59, 146, 161). Low host density selects for a later lysis time. In this condition, progeny phage would accumulate linearly as long as host biosynthetic capacity and resources allow. High host density selects for a shorter lysis time since the progeny of a phage that lyses sooner will have the opportunity to infect a secondary host sooner, thus exponentially increasing its numbers despite the lower burst size. Host quality is also important, as a low quality host with limited biosynthetic resources would support the production of lower amounts of phage per unit time. In the environment, host density

---

This dissertation follows the style of Journal of Bacteriology.

and quality may vary rapidly between extremes, so the lysis timing system must be able to evolve rapidly to different lysis times. For this reason, it is thought that lysis systems based on holins as the timing determinants have evolved.

Lambda lysis cassette

The first dsDNA lysis system to undergo extensive study was that of the coliphage lambda, making its system the classic lysis paradigm. The lysis system includes four genes, *S*, the holin, *R*, the endolysin, and the *Rz*/*Rz1* spanin genes (Fig. 1.1A). Translation of the *S* gene can occur from one of two translational starts due to a dual start motif (18). Initiation at codon 1 or 3 produces S107 or S105, respectively. S107 is the antiholin and S105 is the holin and are each named according to its amino acid length. A pair of stem loops upstream of the coding sequence allows for production of a 2:1 ratio of S105 to S107 (17, 28). The *Rz*/*Rz1* genes are unique in that *Rz1* is fully imbedded in the *Rz* gene in a +1 register (166). The lysis genes are expressed from the late promoter, pR'. Upon commitment to the vegetative cycle, an anti-terminator, Q, is produced as a delayed early protein. Q allows read-though of terminators by binding to RNA polymerase at the pR' promoter. This prevents termination at terminators between the promoter and the late genes resulting in lysis gene expression.

S107, the antiholin

S105 and S107 differ by 2 amino acids but the proteins have different topologies (Fig. 1.1B) and opposite functions. Both proteins have three transmembrane domains

(TMDs), but the first TMD of S107 contains additional Met1 and Lys2 residues. The additional positive charge of S107 causes the retention of TMD1 in the membrane upon initial insertion resulting in an N-in, C-in topology (152). S105 is the prototype class I holin and has an N-out, C-in topology (50, 148). The N-terminus of S105 retains its formylated Met indicating that its TMD1 is exported to the periplasm faster than *Escherichia coli* deformylase can act (152). Mutation of the S105 start, Met3, results in S107 accumulation and defective lysis suggesting that the 2 additional amino acids block lysis (115). Furthermore, mutation of Lys2 converts S107 into a functional holin (16). The addition of energy poisons reverse S017's ability to block lysis. This suggests that the energized membrane is required for S107 function, possibly through N-terminus retention in the periplasm (16). Interestingly, co-expression of SΔTMD1 with S105 results in delayed lysis reminiscent of S107-S105 co-expression. This result indicates that the absence of TMD1 from the membrane is responsible for S105 inhibition (152). S107 was found to bind S105 as a heterodimer in the membrane indicating that S107 mediated inhibition of S105 is by direct protein-protein interaction (54). Antiholin mediated S105 inhibition is only one level of lambda lysis timing. S105 contains an intrinsic timer dictated by its primary structure.

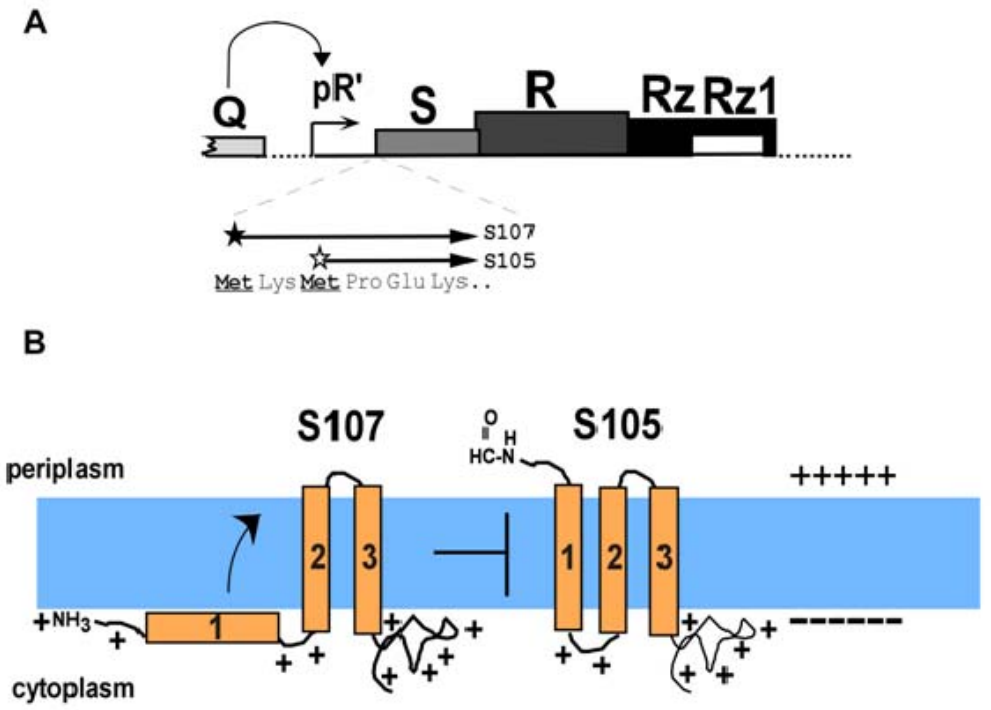

FIG. 1.1. Phage lambda lysis cassette and holin topologies. (A) Bacteriophage lambda uses four genes to beget lysis. *S*, the holin, *R*, the endolysin, and *Rz/Rz1* of which *Rz1* is fully imbedded in *Rz* in the +1 reading frame. The lysis genes are under the control of the late promotor, pR' which is activated by the anti-terminator, Q. Two proteins are expressed from a dual start motif in the *S* gene: S107, and antiholin and S105, the holin. S107 contains two additional positively charged amino acids than S105. (B) The additional Met and Lys of S107 give it an N-in, C-in topology as opposed to an N-out, C-in topology of S105.

S hole formation

According to the classical lysis paradigm of phage lambda, holin accumulates in the membrane until a genetically programmed time upon which holin triggers to form holes in the membrane (162). Triggering refers to the instant when holin-mediated disruption of the membrane occurs, monitored as cessation of cell growth and respiration (161). Gründling *et al.* (53) measured flagellar rotation speed as a direct read-out of the

proton motive force (PMF) in cells expressing S105. Rotation speed was constant until just before lysis suggesting that the accumulation of S105 does not compromise membrane integrity or PMF until triggering. The accumulation of holin was monitored by White *et al.* using green fluorescent protein (GFP) tagged S105 (151). During the morphogenesis period, S105-GFP molecules accumulate uniformly throughout the membrane until suddenly redistributing to a punctate pattern with the fusion protein in aggregates designated as 'rafts' (151). Presumably, the formation of rafts relates to holin triggering. Fluorescence recovery after photobleaching (FRAP) experiments show that the pre-raft state is mobile whereas the triggered state is immobile. *In vivo* chemical crosslinking shows that pre-triggering S105 accumulates in a homodimer state. S105 crosslinking in the membranes of triggered, lysed cells was in higher order oligomers (51). Although raft formation and hole formation have yet to be correlated physically, most lysis-defective S105 mutants do not form rafts and do not oligomerize past a dimer stages. Based on these results and extensive genetic analysis, the current model for hole formation is that during late gene expression, the holin accumulates uniformly in the membrane as mobile homodimers. At a critical concentration of S105, the dimers oligomerize to nucleate large rafts. Genetic evidence suggests that this oligomerization is allele specific in that single amino acid changes can drastically alter lysis time (153). Rafts are thought to be leaky for ions leading to a local depolarization of the PMF that results in hole formation. Recent cryo-electron microscopy work by Dewey *et al*. sheds light on the physical nature of the hole (37). Cells expressing S105 show gaps in the IM ranging in size from 88 nm to 1.2 μm. Cryo-electron tomography revealed that the holes

were irregularly shaped and lipid-free. The average diameter of the holes was 340 nm making them the largest reported membrane lesions in all biology.

Endolysins

Phages use an endolysin, a muralytic enzyme, to degrade the PG. The PG of the Gram-negative cell wall is composed of alternating N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) sugars connected by β–1, 4 glycosidic linkages (46) (Fig 1.2). Sugar strands composed of 3-mers to 14-mers of alternating MurNAc-GlcNAc units are crosslinked by tetrapeptide chains composed of L-Ala-D-Glu-DAP-D-Ala. The PG forms a single, disorganized layer around the cell. This is in contrast to previous models that suggested that the cell wall was more regular and layered.

Endolysins cleave a variety of bonds within the PG (Fig 1.2). Muraminidases such as T4 lysozyme (47, 73) and lambda R (13) cleave the β-1,4 glycosidic linkages between the MurNAc and GlcNAc residues while glucosaminidases (31) cleave the glycosidic linkages between GlcNAc and MurNAc residues. Amidases cleave the amide bonds between the MurNAc residues and the first L-Ala residue of the peptide crosslinks (66) and peptidases cleave the peptide bonds of the amino acid crosslinks (92).

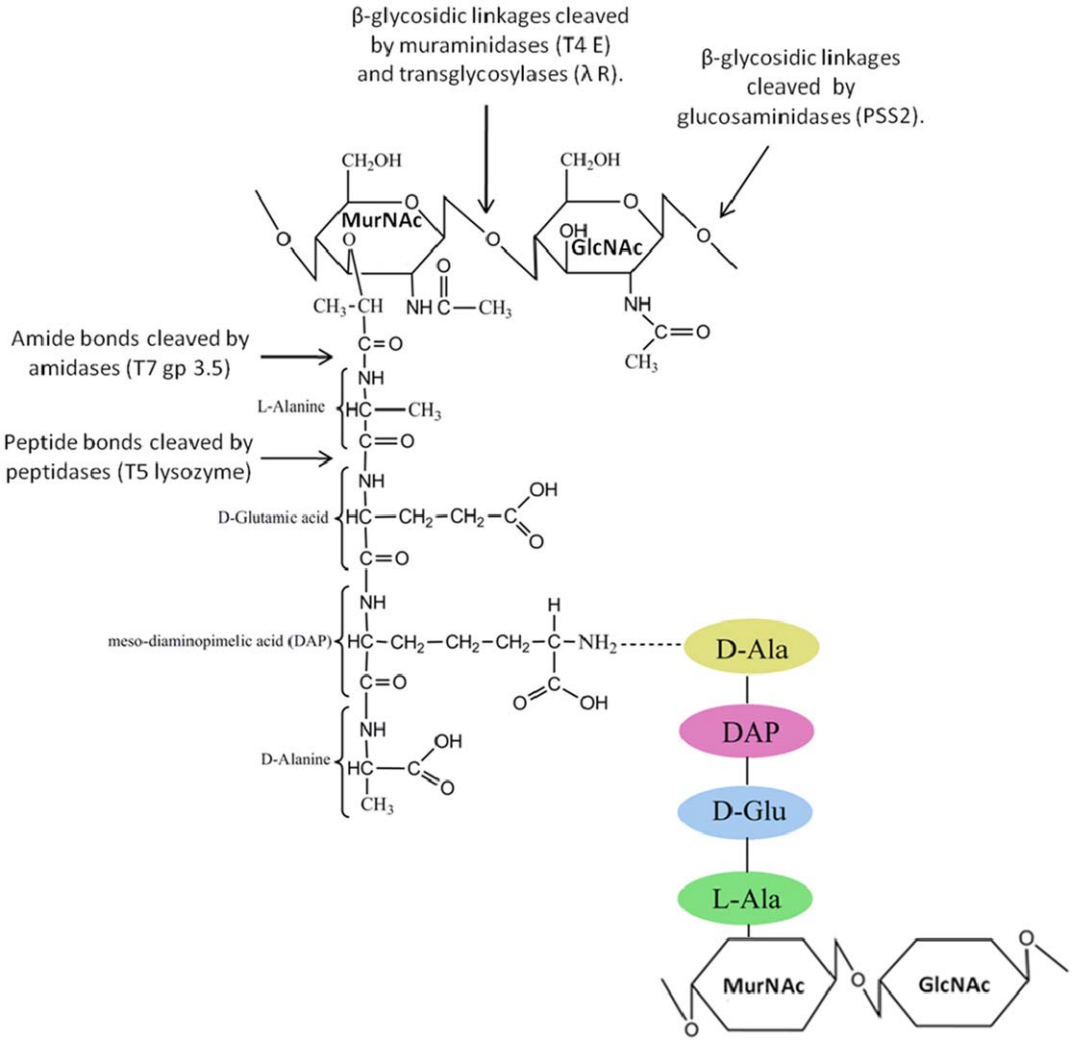

FIG. 1.2. *E. coli* PG and endolysin targets. The bacterial cell wall is composed of PG. PG is made up of repeating units of N-acetyl muramic acid (MurNAc)-N-acetyl glucosamine (GlcNAc) disaccharide. The GlcNAc and MurNAc units are connected by a β-glycosidic linkage. Tetrapeptide crosslinks composed of L-Ala-D-Glu-DAP-D-Ala connect the repeating GlcNAc - MurNAc chains. Endolysins cleave various bonds within the PG.

T4 lysozyme mechanism

The T4 lysozyme, E, is classified as a 1, 4-β-acetyl-muraminidase (47, 73) (EC 3.2.1.17). E uses a hydrolase mechanism and has a catalytic triad composed of a Glu

followed by 5 amino acids, an Asp or Cys followed by 8 amino acids, and a Thr (E-5x-D/C-8x-T) (Fig. 1.3). In the proposed catalytic mechanism, the Glu residue acts as an acid to protonate the glycosidic linkage (58). It was originally thought that the contribution of the Asp was to stabilize the resulting oxocarbonium intermediate; however, the discovery that a D20C mutant retained 80% activity shifted the importance of this residue. At physiological pH, a Cys would remain protonated. Thus in its uncharged state, it would be less likely to stabilize an intermediate electrostatically; however, a sulfhydryl at this position would make an ideal nucleophile (73). Recently it was proposed that a nucleophile in the middle catalytic triad position abstracts a proton from a water molecule, which then attacks C1. The Thr interacts with and stabilizes the water molecule involved in the hydrolysis of the glycosidic bond. What results is a single nucleophilic substitution reaction in which the remaining PG chain is the leaving group along with the glycoside oxygen.

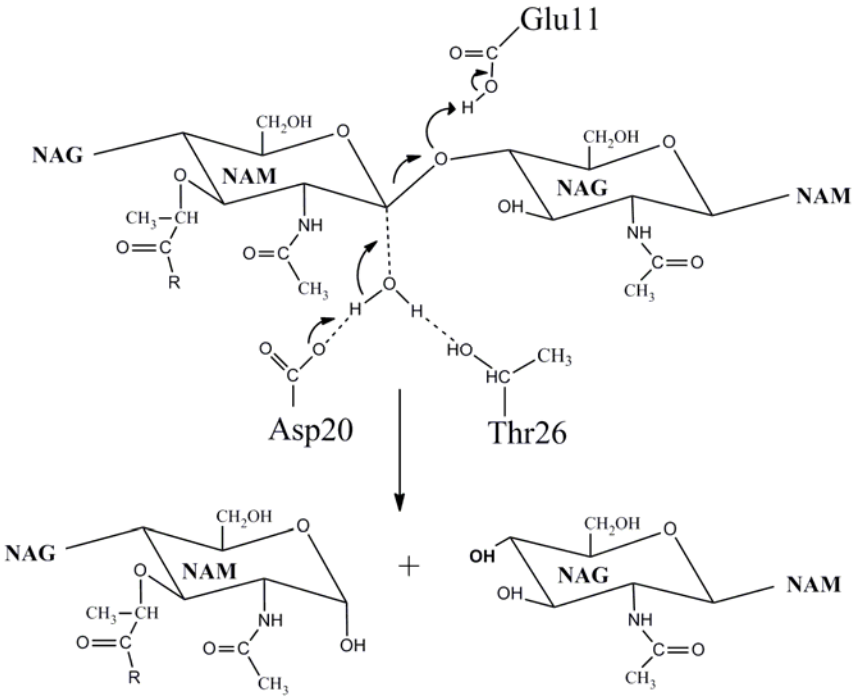

FIG. 1.3. T4 lysozyme mechanism. The T4 lysozyme, E, uses an E-8x-D-5x-T catalytic triad and acts as a hydrolase. Asp 20 removes a proton from a water molecule stabilized by Thr26 (73). The Thr26-stabilized hydroxyl then attacks the C1 position of the MurNAc moiety. A proton is abstracted from Glu11 resulting in a single displacement reaction. The mechanism inverts the position of the C1 hydroxyl from the β conformation to the α conformation.

Transglycosylase mechanism

The catalytic center of the transglycosylase is structurally similar to that of the lysozyme, except for the presence of only a single catalytic residue (120). As with the lysozyme reaction, the proposed transglycosylase mechanism uses a catalytic carboxylic acid to protonate the glycosidic linkage releasing the GlcNAc moiety (Fig. 1.4). A stable muramoyl oxazolinium-ion intermediate is formed involving the N-acetyl group of the MurNAc residue. The catalytic Glu then acts as a base to remove the C6 hydroxyl proton of the intermediate. An intramolecular, nucleophilic attack at C1 by the C6 oxygen

collapses the intermediate resulting in the formation of a non-reducing, 1, 6-anhydro-N-acetyl-muramoyl product.

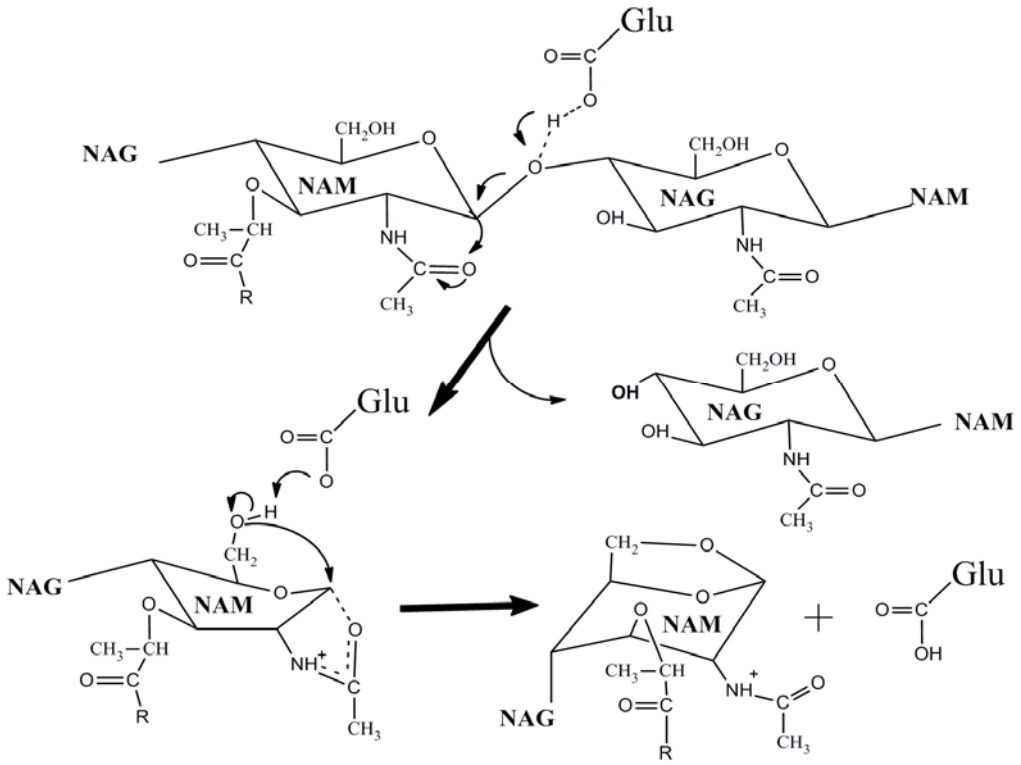

FIG. 1.4. Lytic transglycosylase mechanism. Transglycosylases cleave the β-1,4 glycosidic bond between MurNAc and GlcNAc moieties of PG. A catalytic glutamate acts as an acid to protonate the bond to release the GlcNAc moiety (120). The same Glu then acts as a base to remove a proton from the C6 hydroxyl of the remaining MurNAc residue forming a resonance stabilized intermediate. The intermediate collapses to form a 1, 6-anhydro-N-acetyl-muramoyl product.

Lambda R

Bacteriophage lambda uses a transglycosylase, R, to degrade the PG. R accumulates concomitantly with holin accumulation during late gene expression (148) (Fig. 1.5) . R is an 18KDa soluble protein that accumulates in the cytoplasm in a fully

folded, enzymatically active conformation (13). Since it is sequestered in the cytoplasm, R accumulates with no disruption of cellular structural integrity. After hole formation, R passes through the lesions in the cytoplasmic membrane barrier and degrades the PG. It was originally thought that the resulting disruption of the murein layer was necessary and sufficient for lysis. With no barrier to withstand the turgor pressure of the cell, the cell would burst, releasing progeny phage. This notion has recently been challenged by the results of bioinformatic, biochemical, and genetic study of the Rz/Rz1 spanins (11, 12, 132).

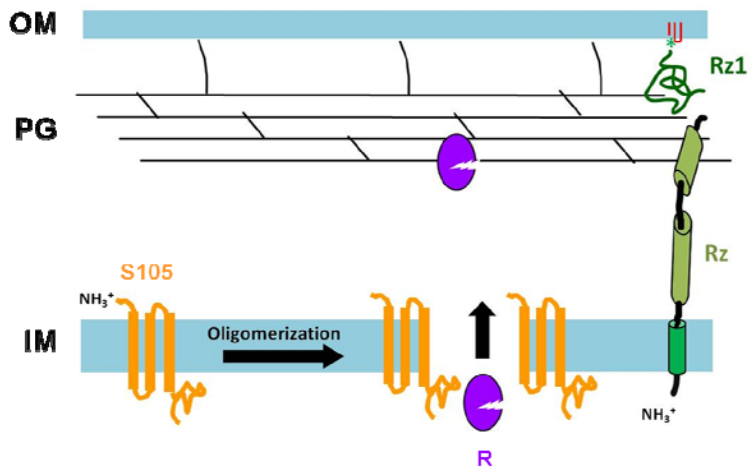

FIG. 1.5. The lambda lysis paradigm. Bacteriophage lambda surpasses the cellular envelope using four proteins. S105, the holin, accumulates in the IM while R, the endolysin, accumulates in the cytoplasm. Rz and Rz1, the spanins, form a laterally immobile complex across the periplasm with Rz and Rz1 in the IM and OM, respectively. Upon holin triggering and lesion formation, R passes through the hole and degrades the PG followed by spanin mediated OM disruption.

Rz and Rz1, the spanins

Rz was initially thought to be a missing endopeptidase whose activity had been detected in lambda phage lysates (138). Under standard laboratory conditions (aerobic liquid culture, vigorous shaking at 37°C), only S105 and R are required for lysis. Under these conditions, defects in Rz/Rz1 prevent lysis only in the presence of >10mM Mg^{2+}. Due to this conditional phenotype, Rz and Rz1 were considered accessory lysis proteins (148). Since S105 and R are sufficient for lysis under standard laboratory conditions, it was thought that the PG degradation by R was so complete that cell lysis followed. Furthermore, since the Rz/Rz1 conditional lysis defect was only apparent after OM stabilization, it was thought that the Rz/Rz1 role in lysis was at the OM. The PG is linked to the OM by > 300,000 oligopeptide links between Braun's lipoprotein, Lpp, and the PG (22, 141). It was hypothesized that Rz/Rz1 has an endopeptidase activity to release the PG from its crosslinks (166).

Recently, work by Berry *et al*. (11, 12, 132) has redefined the previous model of dsDNA phage lysis, and has removed the label of "accessory" from Rz and Rz1. Rz has an N-terminal hydrophobic sequence predicted to be a TMD or a secretory signal. Rz1 contains a putative signal peptidase II cleavage site and a lipoylation signal. The locations of Rz and Rz1 have been confirmed as the IM and the OM, respectively, with both proteins embedded in their corresponding membranes (12). The periplasmic domains of Rz/Rz1 were shown by circular dichroism to form a complex that undergoes a conformational change that adds alpha-helicity. This complex formation between Rz in the IM and Rz1 in the OM, across the periplasm, leads to the term 'spanin' to describe

the Rz/Rz1 proteins. Further evidence for Rz/Rz1 complex formation comes from transmission-electron microscopy studies showing that when mixed, the periplasmic domains of Rz and Rz1 form a single particle with a coil-like pattern (11).

The ubiquity of spanins was shown in a comprehensive bioinformatic survey performed of dsDNA phages of Gram-negative hosts. Rz/Rz1 equivalents were found in most (88%) of the phages analyzed (132). Four gene arrangements were found including: the fully embedded arrangement found in lambda, overlapped genes like those in phage P2 in which the *Rz1* equivalent extends past the end of the *Rz* equivalent, fully separated genes, and even single gene equivalents like T1 *gp11*. Gp11 has a C-terminal TMD and an N-terminal lipoylation signal and was shown to compliment a lambda Rz/Rz1 defect, despite little to no sequence homology. This finding suggests that Rz and Rz1 form a complex linking the IM and OM across the periplasm (132).

The classic lysis paradigm now consists of a three stage process in which each step is required for the next (Fig. 1.5). First, the holin accumulates evenly throughout the membrane while the endolysin accumulates in the cytoplasm. The Rz/Rz1 proteins accumulate in the IM and OM, respectively, and form a complex across the periplasm. Presumably, this spanin complex is relatively immobile due to the constraints of the PG meshwork (11). In the moments before lysis, the holin oligomerizes and forms rafts in the membrane. A local depletion of the PMF leads to hole formation allowing escape of the endolysin resulting in PG degradation. Without the constraint of the PG, the Rz/Rz1 complex would be able to oligomerize laterally. Although the exact mechanism of

Rz/Rz1 function is not known, it has been suggested that the spanin collapses into a coiled coil conformation, fusing the inner and outer membranes, lysing the cell (12).

Phage lysis of Gram-positive cells

Phage-mediated lysis of Gram-positive cells occurs in a similar manner to Gram-negative cells. In phages of Gram-positive hosts, typically the "lysis cassette" contains the *holin* and *R* encoding the endolysin, usually referred to as a lysin. Due to the lack of an OM, Gram-positive phages do not have Rz/Rz1 equivalents. Mycobacteria are unique among Gram positives in that they have an external mycolic acid layer that has proteins mounted in it, effectively making it an OM (169). The mycolic acid OM is attached to an arabinogalactan (AG)-PG layer by ester linkages. To overcome this barrier, mycobacteriophages encode an additional lysis protein, Lysin B, a mycolyl-arabinogalactan esterase (111, 118). Lysin B cleaves the ester bonds linking the OM to the AG-PG, releasing the mycolic acid layer.

The holins of phages of Gram-positive hosts are proposed to be functionally similar to those of Gram-negative phage (135). Gram-positive endolysins share activities with Gram-negative endolysins (muraminidase, peptidase, amidase), although their architecture tends to differ (42). Most Gram-positive endolysins are modular in organization in that they have an N-terminal catalytic domain and a dedicated C-terminal substrate binding domain called the cell wall binding domain (CWBD). Some Gram-positive endolysins have multiple catalytic domains. For example, the *Staphylococcus aureus* phage Φ11 has a D-alanoyl-glycyl-endopeptidase domain and an N-

acetylmuramoyl-L-alanine amidase domain followed by a CWBD (98). The CWBDs of Gram-positive lysins are very specific and have high affinity for their substrate. PG varies among bacterial species (122) and Gram-positive lysins have specificity for the host PG they were produced from (42). *Streptococcus pneumonia* phage lysins, for instance, kill various strains of group A *Streptococcus*, but not other strains of oral commensal *Streptococcus* (82). Surface plasmon resonance experiments by Loessner *et al.* measured association constants of Gram-positive lysins for their substrate and found them to be in the nanomolar range (representing strong binding), thus endolysins may be one-time use molecules that remain attached to the PG (83). This characteristic would be valuable in the context of a phage infection by preventing diffusion. Gram- positive cells are sensitive to the action of exogenous lysins (42). A free lysin not bound to PG could prematurely lyse a neighboring infected cell resulting in a non-productive infection (83).

It was originally thought that phages of Gram-positive hosts use a lysis mechanism that was the same as their gram negative counterparts. As with the classic lysis paradigm of lambda, the holin of Gram-positive phages is responsible for lysis timing. Recently, a new Gram-positive lysis paradigm has begun to emerge. The lysin, Lys44, from the *Oenococcus oeni* phage fOg44 was found to have a hydrophobic N-terminal region such as that found in a signal sequence (118). Experiments showed that Lys44 encoded a signal peptide and was exported in its mature, cleaved form. This result was unexpected since the phage also encoded a holin. Up to this point, it was hypothesized that phages required a holin for endolysin export since none of the characterized endolysins contained a signal sequence (129).

During phage infection studies, mature, processed Lys44 began to accumulate 80 minutes after infection despite a 150 minute latent period (118). The authors concluded that the phage must have a regulation system to keep Lys44 inactive until it is needed. A model was proposed in which the holin regulates Lys44 by mimicking the cellular mechanism that regulates autolysins involved in cell wall turnover. Support for this comes from reports by Jolliffe *et al.* showing that *Bacillus subtilus* autolysins are activated by the addition of energy poisons, suggesting that the energized membrane plays a role in regulating cellular autolysis (67). Although the exact mechanism for this regulation is not understood, it has been proposed that the chemical and structural environment of the energized membrane inhibits autolysins. Changes in membrane characteristics such as permeability (71) and fluorescent probe binding (32, 99) have been observed after depletion of the PMF. This indicates that the membrane environment differs with and without the PMF.

SAR endolysins: A new lysis paradigm

Recently, a new paradigm for dsDNA phage lysis of Gram-negative hosts has emerged. As with phage fOg44, it was observed that some phages of Gram-negative hosts also encoded endolysins that have a hydrophobic N-terminal domain similar to that of a signal sequence. The lysozyme, Lyz, of phage P1 was found to have such an extension and was chosen for detailed study (159). When Lyz was expressed without a holin, it caused lysis of the host culture suggesting host-mediated secretion of the enzyme. Indeed, Lyz mediated lysis could be prevented by the addition of azide, a SecA

inhibitor. Furthermore, a sub-cellular fractionation of cells expressing Lyz showed that the protein was partitioned both into the membrane and soluble fractions of the cell. Surprisingly, however, the apparent molecular masses of the soluble and membrane bound Lyz were the same, indicating that the secretion was not associated with signal peptide cleavage. Taken together, the data indicated that the N-terminal domain was responsible for: a) secretion of the protein through the Sec pathway, b) tethering the protein to the cytoplasmic membrane, and c) releasing the protein into the periplasm (Fig. 1.6). The domain was thus called a Signal Anchor Release, or SAR domain.

P1 Lyz

Since the lysis genes are expressed simultaneously with all of the genes required for morphogenesis, a secreted enzyme must be maintained in an inactive state to prevent premature lysis of the host cell. This regulation was first studied in detail in P1 Lyz. Sequence analysis showed that Lyz has a catalytic triad of E-5x-**C**-8x-T. This is similar to that of the T4 lysozyme, E, which has an E-5x-**D**-8x-T triad (Fig. 1.7). Mutagenesis studies of T4 E show that a catalytic D to C mutation was very well tolerated showing an activity 80% that of wild type E (58). To date, all SAR endolysins studied are homologs of T4 E. It was observed that Lyz had two cysteines (C13 in the SAR domain and C44) N-terminal to its catalytic C51. The *E. coli* protein DsbA forms disulfide bonds with consecutive cysteines as they emerge in the periplasm (8). It was therefore hypothesized that a disulfide bond would form between C44 and the catalytic C51, leaving the membrane sequestered C13 in its sulfhydryl form (Fig. 1.8). This lead to an elegant

model in which an inactivating disulfide was formed between C44 and C51 and upon release from the membrane, the SAR sulfhydryl at C13 would attack C44, forming a new disulfide bond. This disulfide bond isomerization was hypothesized to release C51, activating the protein (159).

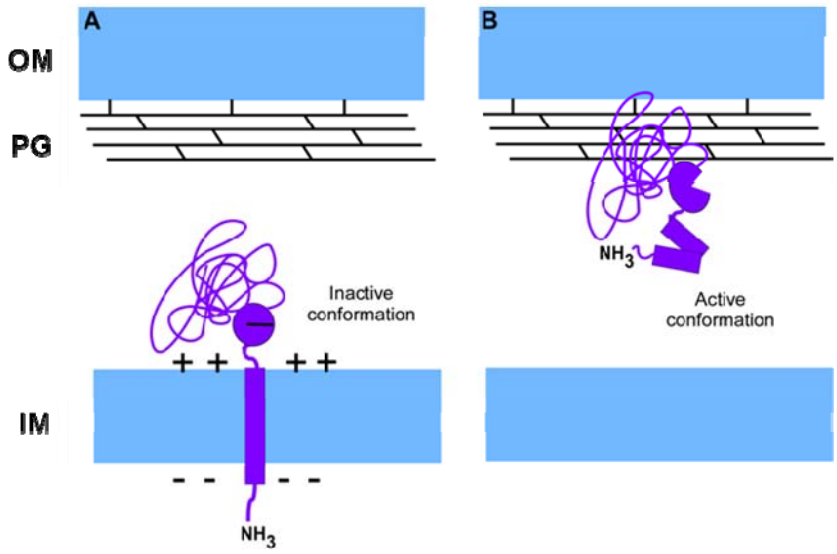

FIG. 1.6. The SAR endolysin paradigm. Unlike the classic lysis paradigm, the SAR endolysin paradigm uses a secreted endolysin. (A) The SAR endolysin is secreted by virtue of its N-terminal Signal Anchor Release domain via the Sec translocon. The SAR domain anchors the endolysin to the cytoplasmic membrane in an inactive conformation. (B) Upon membrane depolarization, the SAR domain is released from the membrane, activating the protein.

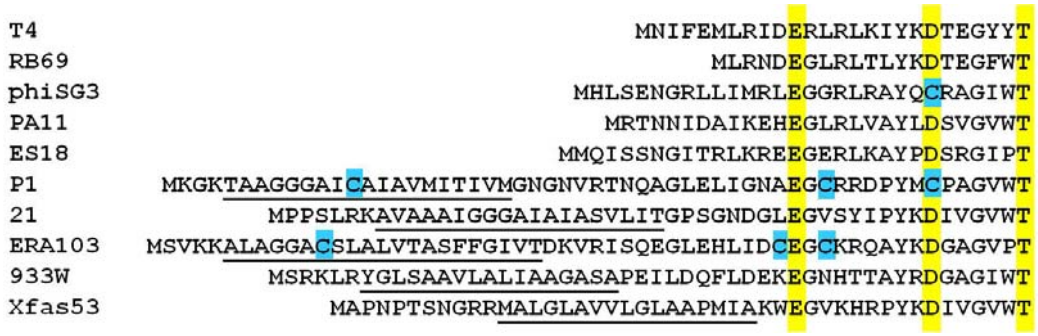

FIG. 1.7. Alignment of phage lysozymes. The lysozyme of phage T4, E, has a catalytic triad composed of E-5x-D-8x-T. Homologs of T4 E use this triad, maintaining the spacing between residues, although in some, a catalytic Cys replaces the catalytic Asp. Here, homologs of T4 E are aligned by their catalytic residues (highlighted in yellow) and Cys residues are highlighted in blue. SAR endolysins are homologs of T4 E (SAR domains underlined).

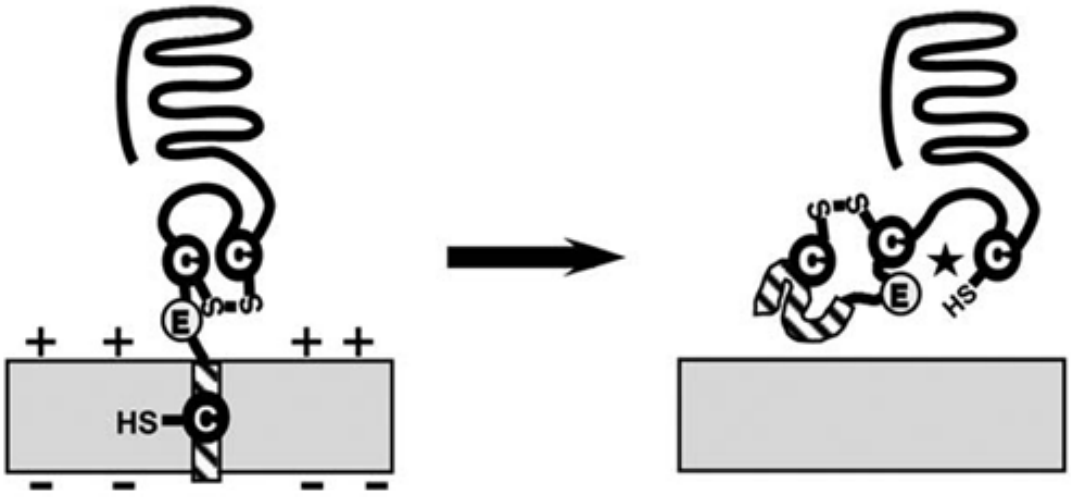

FIG. 1.8. P1 Lyz activation by disulfide bond isomerization. In its membrane bound form, P1 Lyz is inactive due to a disulfide bond formed between a catalytic Cys and a non-catalytic Cys (left panel). Upon release from the membrane, the SAR domain Cys initiates a disulfide bond isomerization to release the catalytic Cys (star, right). (Figure reprinted with permission from AAAS.)

The alternative disulfide bonding pattern was demonstrated *in vivo* using a sulfhydryl cyanylating chemical, 1-cyano-4-dimethylaminopyridinium (CDAP) (156). Treatment of a cyanylated protein with a strong base such as ammonium hydroxide cleaves the protein on the N-terminus of the cyanylated cysteine. When Lyz was cleaved before and after lysis using this method, two distinct cleavage products were observed confirming the disulfide bond isomerization. The crystal structure of Lyz was solved by Arockiasamy Arulandu and Stephanie Swanson in the laboratory of Jim Sacchettini (Texas A&M University) (Fig. 1.9). To obtain the inactive form of Lyz (Lyz^{i}), the periplasmic domain of the protein was crystallized. Full length Lyz was purified and crystallized for the active form (Lyz^{a}). The structures show several interesting features. First, Lyz displays the classic lysozyme fold (150) with an N-terminal catalytic domain and a C-terminal helix bundle connected by an inter-domain helix . The structural rearrangement required to activate the enzyme is also apparent. The catalytic cleft of the periplasmic domain of Lyz was found to be in an inactive conformation with: a) the side chains of the catalytic Glu42 and Thr57 residues pointing outwards and b) a bend in the inter-domain helix (Fig. 1.9 left). The catalytic cysteine of Lyz^{i} was maintained in a disulfide bond with Cys44 in the vicinity of the active site. Upon activation of the enzyme by the disulfide bond isomerization, three helices in the N-terminal domain unwind to form three beta strands that form a sheet (Fig. 1.9 right). This shift in secondary structure repositions the newly freed catalytic Cys51, rotates the catalytic Glu42 and Thr57 inward to form the active site, and relieves a bend in the inter-domain helix. The active structure also shows the newly formed Cys13-Cys44 disulfide.

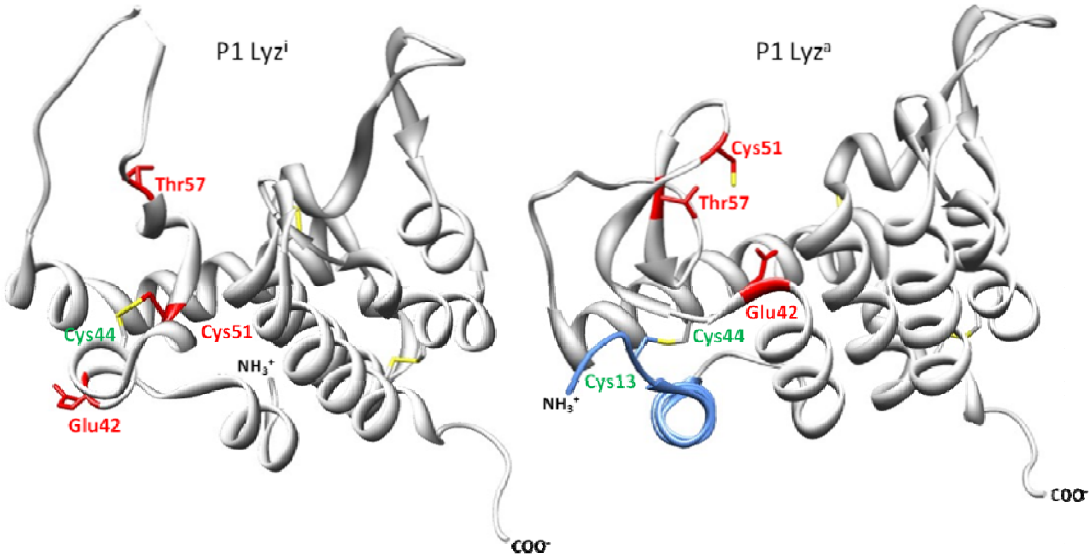

FIG. 1.9. Comparison of inactive and active forms of P1 Lyz. The inactive form of P1 Lyz (Lyz^{i}, left) and active forms of P1 Lyz (Lyz^{a}, right) are shown. Upon release of the SAR domain, a large structural rearrangement occurs that creates the catalytic cleft of activate the enzyme. Catalytic residues (red), disulfide bonds (stick form), Cys residues involved in the disulfide bond isomerization (green), and the SAR domain (blue) are indicated.

In the disulfide bond isomerization regulation mechanism, the role of the SAR domain is two-fold. First, the SAR domain targets the protein for secretion and releases it from the membrane, but this function was found not to be unique to the P1 Lyz SAR domain. When the SAR domain was replaced with the cleavable signal sequence of PhoA, Lyz was still targeted to and released into the periplasm, although activation required addition of exogenous reducing agent. The second function of the Lyz SAR domain is to provide the free sulfhydryl that initiates the disulfide bond isomerization. This function was also found not to be unique to the Lyz SAR domain. When the Lyz SAR domain was replaced with the cysteine-less SAR domain from the SAR endolysin of phage 21, R^{21}, the resulting chimera was not lytic. However, the addition of a Cys on

the proper, protein-interacting face of the SAR domain in this chimera restored lysis. This indicates that the Lyz SAR domain itself is not responsible for the large structural rearrangement that activates the enzyme; rather it is the release of the inactivating disulfide by the SAR sulfhydryl or addition of exogenous reducing agent that triggers the conformational change.

The SAR domain

SAR domains are composed of an over-representation of weakly hydrophobic and polar, non-charged residues such as Ala, Gly, Ser, and Thr (80, 100, 159). The over-representation is compared to TMDs that contain mostly Leu, Ile, and Val residues (41, 155). Presumably, the weakly hydrophobic nature of the SAR domain supports its release from the cytoplasmic membrane. A SAR domain first acts as a TMD so it must span the cytoplasmic membrane. Membrane spanning segments are typically alpha-helices and are 17-35 residues in length on average (68). Preliminary research suggests that SAR domains form a single helix in a lipidic environment (J. Deaton, unpublished data). Addition of energy poisons such as dinitrophenol or cyanide promote premature release of the SAR domain from the membrane suggesting that SAR domains are maintained as meta-stable membrane helices in the context of an energized membrane.

An interesting feature of a SAR domain is its ability to retain solubility in an aqueous environment after release from the membrane thus allowing the SAR domain to fold into the now soluble, active protein. This is a very interesting feature from a thermodynamic standpoint. A SAR domain must be hydrophobic enough to be

recognized by the Sec translocon and partitioned into the membrane, but it must also be somewhat neutral as to not alter the protein's structure upon release. Typically, exposure of a hydrophobic sequence to the aqueous environment would result in aggregation due to the hydrophobic effect (137), yet the SAR domain has evolved a delicate balance as to achieve stability and functionality in both environments. This balance of hydrophobicity can be disturbed by converting Gly or Ala residues to Leu residues, as these mutations block the ability of the SAR domain to exit the bilayer. This feature is protein specific i.e. 3 Gly to Leu substitutions are needed to retain P1 Lyz in the membrane and 2 substitutions are required to retain the SAR domain of R^{21} (74, 133).

ERA103

Phage ERA103 infects the Gram-negative plant pathogen *Erwinia amylovora*, the causative agent of fire blight in fruit producing plants, such as pear and apple trees (104). Current treatment for the crop eradicating disease is streptomycin, but success of treatment is fading due to the increase in resistance among *E. amylovora* populations. Recent studies have suggested the use of phage to control the pathogen. In 2004, a German group looked to the lysozyme of *E. amylovora* phage ΦEa1h as a potential control mechanism. The *lyz* gene from φEa1h was cloned and IPTG induced Lyz^{Ea1h} was found to lyse cells (70).

The sequence of *lyz* from ΦEa1h is identical to that of lyz^{103} from *E. amylovora* phage ERA103. The protein has 51% similarity and 33% identity with P1 Lyz, suggesting that Lyz^{103} is a T4 E homolog. Further examination of the sequence shows

the presence of an N-terminal stretch of 21 weakly hydrophobic amino acids reminiscent of a SAR domain. A Cys is present in the putative SAR domain indicating the potential for a disulfide bond activation mechanism similar to P1 Lyz. Unlike P1 Lyz, Lyz^{103} does not have a catalytic Cys, but rather uses the canonical Asp residue. Thus, if disulfide bond rearrangement is involved in its activation, then the mechanism must be different.

Phage 21

The lambdoid phage 21 is another example of the SAR endolysin-mediated lysis paradigm (133, 159). So far, the study of phage 21 has had the most impact in determining the role of the holin in the SAR lysis system. In the classic lysis paradigm, the holin controls lysis timing by forming lesions in the IM allowing passage of the endolysin into the periplasm where it can degrade the PG. In the SAR lysis system, however, the endolysin begins in the periplasm, so the question was posed: "how does the holin control lysis timing in the SAR lysis paradigm?" Initial study of the SAR endolysin showed that the SAR domain can be released prematurely by simple membrane depolarization (159). Thus, it was hypothesized that the holin regulates lysis timing by causing the collapse of the PMF in a temporally regulated manner. Initial studies of S^{21}, the phage 21 holin, show that it cannot complement an S_{am} mutation in the lambda lysis cassette, despite accumulation of active endolysin (108). This data suggests that S^{21} forms small holes that do not allow passage of lambda R into the periplasm. S^{21} has been termed a "pinholin" to highlight the small size of the holes.

S^{21}, the pinholin

The lysis cassette of phage 21 contains a holin, S^{21}, SAR endolysin, R^{21}, and *Rz*/*Rz1* genes (Fig. 1.10A). Like phage lambda, S^{21} uses a dual start motif to express two proteins: $S^{21}68$, the holin and $S^{21}71$, the antiholin (109). $S^{21}68$ is the prototype class II holin with two TMDs arranged in an N-in, C-in topology (Fig 1.10B).

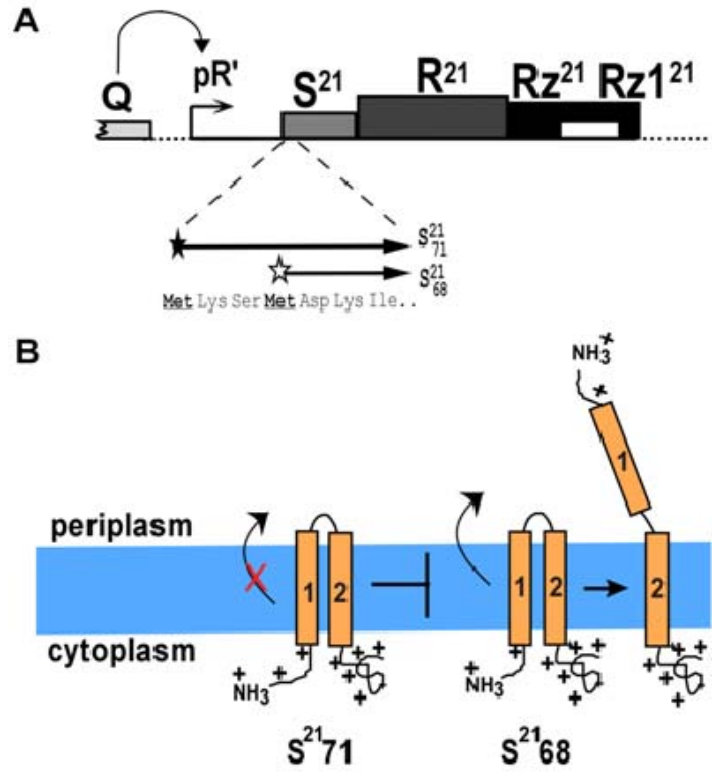

FIG. 1.10. Phage 21 lysis cassette and holin topologies. (A) The lysis cassette of lambdoid phage 21 is composed of a S^{21}, R^{21}, and *Rz*/*Rz1*. S^{21} uses a dual start motif to encode $S^{21}71$, the antiholin and $S^{21}68$, the holin. (B) $S^{21}68$ is a class II holin and has 2 TMDs in an N-in, C-in topology (109). TMD1 is a SAR domain that exists the membrane in the pathway to hole formation (106).

Interestingly, the first TMD is composed largely of weakly hydrophobic residues reminiscent of a SAR domain. When the SAR domain of P1 Lyz was replaced with the $S^{21}68$ TMD1, the full-length chimeric protein $S^{21}68_{TMD1}\Phi Lyz_{\Delta SAR}$ was present in the periplasm. This is consistent with the hypothesis that $S^{21}68$ TMD1 is a SAR domain. To determine if TMD1 actually exits the bilayer within the context of the holin, Park *et al*. replaced Ser16 of TMD1 with a Cys (109). The prediction was that if TMD1 exits the membrane to the oxidizing environment of the periplasm, and if $S^{21}68$ molecules could oligomerize, then disulfide linked dimers would be observed. Membranes of cells expressing $S^{21}68_{S16C}$ indeed contained disulfide linked dimers indicating that TMD1 leaves the membrane, perhaps on the pathway to hole formation. Genetic and biochemical studies by Pang *et al*. indicate that the pinhole is composed of a heptamer of $S^{21}68$ with a lumen of < 2nm (107). Interestingly, not all SAR endolysins are coupled with pinholins (108). The holin of phage P1, LydA, forms large holes as seen by complementation of a lambda S_{am} mutation (157). As a consequence, it cannot be concluded that all SAR endolysins are served by pinholins.

R^{21}: A new class of SAR endolysin

As with any lysis system, the endolysin of 21 needs regulation to prevent premature lysis. With P1 Lyz, this is accomplished by a disulfide bond isomerization initiated by the SAR domain Cys. Inspection of the R^{21} primary structure shows no Cys residue in the SAR domain. Since there is no Cys in the SAR domain, R^{21} cannot be regulated by a disulfide bond isomerization like P1 Lyz. This makes R^{21} the prototype of

a new sub-class of "Cys-less" SAR endolysins. Interestingly, bioinformatic analysis of putative SAR endolysins suggests that R^{21}-like proteins are predominant over P1-Lyz like SAR endolysins, at least among sequenced phages.

933W

The lambdoid phage 933W has gained significant attention recently in the field of human disease, specifically Entero-hemorrhagic *Escherichia coli* (EHEC) serotype O157:H7, the causative agent of hemorrhagic colitis and hemolytic-uremic syndrome (136). EHEC produces the Stx2 toxin in the AB_5 Shiga family of toxins (78). The Stx2 toxin is a complex composed of a monomer of subunit A and a pentamer of subunit B. Subunit A is an N-glycosidase that cleaves an adenine residue from the 28S RNA of the 60S ribosomal subunit, rendering it inactive. The B pentamer is responsible for cell recognition via the Gb3 glycolipid receptor, and for internalization of the A toxin.

The genes encoding the Stx2 toxin are under control the late promoter, pR', of the 933W prophage (Fig. 1.11). While EHEC strain EDL933 has many prophages, only prophage 933W can actively produce phage (61). As with lambda, 933W pR' is induced upon prophage induction as a result of SOS-dependent DNA damage (145). Activation of pR' induces the morphogenesis, lysis, and concomitantly, the Stx2 genes (144). The fact that phage induction results in toxin expression explains why DNA damaging antibiotics such as norfloxacin, a gyrase inhibitor, exacerbate EHEC infections (167).

FIG. 1.11. Phage 933W lysis cassette. As with phage lambda, the lysis genes of prophage 933W of EHEC strain O157:H7 are expressed from the late pR' promotor which is induced by the anti-terminator, Q. The 933W lysis cassette consists of a pinholin (*S*), a SAR endolysin (*R*, star), and spannin genes (*Rz/RZ1*). Also expressed from the late promotor are the *Stx A* and *B* shiga toxin genes.

In EHEC strains induced for the 933W prophage, Stx2 is found extracellularly (124). EHEC contains the LEE pathogenicity island that encodes a type III secretion system (5). The secretion system supports the release of proteins responsible for cell adhesion, but not the Stx2 toxin. It was hypothesized that since the phage lysis genes and toxin genes are co-transcribed, perhaps phage induced lysis provides a mechanism for toxin release. Indeed, 933W ΔSR prophage induction results in complete cell association of Stx2 suggesting that the phage's lysis machinery alone is responsible for toxin release. Inspection of the 933W lysis cassette shows the presence of a putative SAR endolysin, R^{933W}. Other *stx* containing phages such as VT2-Sakai, Stx2-φII, and the prophage of the recent Germany outbreak strain contain endolysins identical to R^{933W} (87, 113, 116, 119). As described above, SAR endolysins are initially secreted to the cytoplasmic membrane in an inactive conformation. In the context of a phage infection or prophage induction, presumably, SAR endolysins are released synchronously due to membrane depolarization by the holin. In the absence of a holin, however, SAR endolysins gradually release from the membrane. Membrane release results in a

refolding event that activates the enzyme (133, 159). Thus, a SAR endolysin expressed alone can cause lysis of the host cell. Keeping in mind the vital role phage mediated lysis plays in shiga toxin release; R^{933W} is a potential drug target in EHEC infections.

Project aims

The goals of this dissertation research are to:

1) Determine if other modes of disulfide bond regulation of SAR endolysins exist and, if so, characterize them.
2) Define the mechanism of regulation in SAR endolysins with no cysteine in the SAR domain.
3) Establish the prevalence of the SAR endolysin within dsDNA phage lysis of Gram-negative hosts.
4) Determine the value of the SAR endolysin as a drug target in EHEC O15:H7 infections.

CHAPTER II

REGULATION OF A PHAGE ENDOLYSIN BY DISULFIDE CAGING*

Introduction

In infections by double-stranded DNA phages, host lysis requires degradation of the peptidoglycan by a phage-encoded endolysin (148). By far the most intensively studied endolysin is the T4 lysozyme, E (EC 3.2.1.17), which attacks the glycosidic bonds between GlcNAc and MurNAc in the murein (7). During the latent period, canonical endolysins are produced as fully active enzymes sequestered in the cytoplasm, thereby preventing premature lysis. Another phage protein, the holin, terminates the infection cycle by forming extremely large, non-specific holes in the membrane, which allow the endolysin to escape and attack the murein layer (163). Recently, studies on the lysis system of bacteriophage P1 have revealed that phage-encoded endolysins are not always dependent upon holins for export (158, 159). Although it is an ortholog of T4 E, the P1 lysozyme, P1 Lyz, is translocated across the cytoplasmic membrane by the host *sec* system, by virtue of an N-terminal transmembrane domain (TMD) absent in E (Fig. 2.1A). Since this transmembrane domain is not removed by signal peptidase, nascent P1 Lyz remains tethered to the membrane with its catalytic residues already present in the periplasm. The P1 Lyz TMD exits the membrane and

---

*Reprinted with permission from "Regulation of a phage endolysin by disulfide caging" by Kuty, G.F., M. Xu, D.K. Struck, E.J. Summer, and R. Young, 2010. Journal of Bacteriology, 192, 5682-5687. Copyright 2010 by ASM Press.

becomes a part of the soluble, periplasmic form of the protein, either at a slow spontaneous rate or, more efficiently, when the holin triggers to depolarize the membrane (109). Because of the unique ability to direct *sec*-mediated export and membrane insertion and also to support release into the periplasm from the bilayer, the TMD of P1 Lyz was designated as a SAR (signal anchor release) domain. More recently, other SAR endolysins have been identified and characterized (130, 133). In fact, bioinformatic analysis suggest that most members of the T4 lysozyme family, recognizable by the Glu-8x-Asp/Cys-5x-Thr catalytic triad (Fig. 2.1A), are SAR endolysins (43 of 58 entries in the Genbank Protein database) (133).

Since the P1 *Lyz* gene is expressed well before progeny phage have been assembled, there must be a mechanism to ensure that the membrane-tethered form of the protein is kept enzymatically inactive so that premature lysis is avoided. A key to the regulation of P1 Lyz is the fact that the P1 enzyme has a catalytic cysteine residue, Cys_{51} (Fig. 2.1A), in the central position of the catalytic triad, in contrast to E and most of its orthologs, which have an Asp in this position (159). Genetic, biochemical, and structural analysis of P1 Lyz demonstrated that the membrane-tethered form is inactive for two reasons: first, the entire catalytic domain is misfolded, so that the active site cleft is completely missing; and second, the catalytic Cys_{51} is covalently occupied in a disulfide bond with another Cys at position 44. This led to a model for activation in which a thiol (Cys_{13}) present in the SAR domain becomes unmasked upon membrane release and triggers a disulfide-bond isomerization, liberating the thiol of the catalytic Cys_{51}. This model was

confirmed by crystal structures showing the alternative disulfide linkages in the inactive and active forms of P1 Lyz (158).

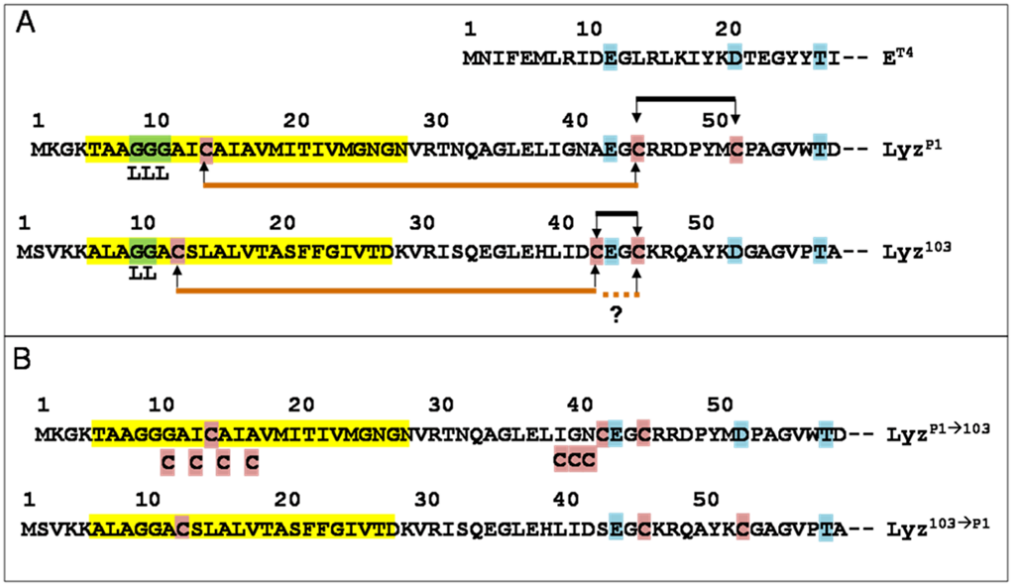

FIG. 2.1. Sequence alignments. (A) Alignment of T4E, P1 Lyz, and Lyz^{103}. Catalytic residues are highlighted in blue, Cys residues in pink, and SAR domains in yellow. Locations of Leu substitutions are indicated below the corresponding Gly (green). Inactivating/caging disulfides for P1 Lyz and Lyz^{103}, respectively, are shown as a black connected arrow above the participating Cys residues. Disulfides resulting from isomerization are shown below the participating Cys residues as brown connected arrows. (B) Alignments of P1 Lyz$^{\rightarrow 103}$ and Lyz$^{103\rightarrow P1}$ conversion mutants. Catalytic residues are highlighted in blue, Cys residues in pink, and SAR domains in yellow. Locations of Cys substitutions are indicated below the corresponding residue.

P1 Lyz became the prototype of a class of SAR endolysins recognizable by the Asp → Cys substitution in the catalytic triad and the presence of activating Cys in the N-terminal SAR domain. However, most SAR endolysins belong to a second major class,

represented by R^{21}, the endolysin of the lambdoid phage 21 (159). These enzymes have the canonical Glu-8x-Asp-5x-Thr catalytic triad and no Cys residue in the SAR domain. Instead, genetic and structural analysis revealed that in the inactive form, the catalytic domain has nearly the correct fold, except for a displacement of the active site Glu, but is subject to steric hindrance by the proximity of the bilayer in which the SAR domain is embedded. In the soluble, active form, the SAR domain of R^{21} has refolded into the main body of the enzyme, providing a floor to the active site and repositioning the catalytic glutamate to its proper place (133). Thus the R^{21} regulatory scheme is markedly different from that of P1 Lyz, where the released SAR domain provides only the free thiol for the disulfide bond rearrangement and makes few contacts with the enzyme itself.

Here we examine the regulation of the endolysin $Lyz^{103}$of the *Erwinia amylovora* phage ERA103 (GenBank: EF160123), which has characteristics of both of these major classes: it has a Cys residue in an N-terminal hydrophobic sequence but retains the canonical Asp residue in the catalytic triad. The results are discussed in terms of a model for SAR-dependent disulfide bond isomerization distinct from that of P1 Lyz and its homologs but which nevertheless confers a covalent constraint on premature activation of the muralytic activity.

Materials and methods

Bacterial strains and growth conditions

The *Escherichia coli* strains XL1-Blue, MC4100, and MG1655 Δ*tonA lacY*⁻ *lacI*^{q} have been described (127, 133). Standard conditions for the growth of cultures and the monitoring of lysis kinetics have been previously described (29, 127). All bacterial cultures were grown in standard LB medium, supplemented with 100 μg/ml ampicillin when appropriate. When indicated, isopropyl β-D-thiogalactopyranoside (IPTG), dinitrophenol (DNP), and dithiothreitol (DTT) were added to final concentrations of 1 mM, 10 mM, and 1 mM respectively.

DNA procedures and plasmid construction

Procedures for the isolation of plasmid DNA, DNA amplification by PCR, PCR product purification, DNA transformation, site-directed mutagenesis, and DNA sequencing have been previously described (52, 126, 128). The construction of the plasmid pP1 Lyz, a derivative of pJF118EH, has been described previously (45, 159). The plasmid pLyz^{103} was constructed by amplifying and inserting *lyz*^{103} between the *EcoR*I and *Hind*III restriction sites of pJF118EH. Similarly, the construct, pETLyz^{103} was constructed by amplifying *lyz*^{103} and inserting it between *Xba*I and *Bam*HI of pET11a (Novagen). For over-expression purposes, the inactive allele *lyz*$^{103}_{D52N}$ was used. Derivative alleles of *lyz*^{103} and P1 *Lyz* were made using site-directed mutagenesis. For detection and purification purposes, the *lyz*^{103} allele was modified to encode an

oligohistidine tag ($Gly_2His_6Gly_2$) appended to Met_{178} by site-directed mutagenesis. All purified protein cited in this work refers to the oligohistidine-tagged version.

SDS-PAGE and Western blotting

SDS-PAGE, Western blotting, and immunodetection experiments were performed as previously described (52). Antiserum against P1 Lyz was prepared in chickens by Aves Labs (Tigard, OR) and was used at a dilution of 1:1,000. A mouse monoclonal antibody against the oligo-histidine epitope tag was purchased from Amersham and was used at a dilution of 1:3,000. Horseradish peroxidase-conjugated secondary antibodies against chicken IgY were purchased from Aves Labs, and were used at a 1:2000 dilution for colorimetric detection, and a 1:300,000 dilution for chemiluminescent detection. The anti-mouse IgG horseradish peroxidase-conjugated secondary antibody was supplied with the SuperSignal chemiluminescence kit (Pierce) and was used at a 1:5000 dilution. Blots were developed using the chromogenic substrate 4-chloro-1-naphthol (Sigma) or with the West Femto SuperSignal chemiluminescence kit (Pierce). Chemiluminescent signal was detected using a Bio-Rad ChemiDoc XRS.

Subcellular fractionation

Soluble or membrane localization was determined as described previously (133, 159). Briefly, 25 mls of an induced culture were collected by centrifugation at 5,000 x g in a Sorvall Superspeed RC2-B centrifuge, and resuspended in 2 mls of French press

buffer (0.1 M sodium phosphate, 0.1 M KCl, 5 mM EDTA, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, pH 7.0). Cells were disrupted by passage through a French pressure cell (Spectronic Instruments, Rochester, N.Y.) at 16,000 lb/in^2. The membrane and soluble fractions were separated by centrifugation at 100,000 x g in a Beckman TL-100 Ultracentrifuge for 60 min. Equivalent amounts of each fraction were examined by SDS-PAGE and Western blotting as described above.

Sulfhydryl modification using PEG-OPSS

To detect the presence of free cysteines in SAR endolysins, a 5 ml sample of a culture induced in logarithmic phase for 25 min was precipitated by TCA. The pellet was resuspended in 1 ml of PEGylation buffer (500 mM Tris pH 7.0, 1% SDS, 1 mM EDTA) and mPEG-OPSS (Nektar Transforming Therapeutics, Huntsville, Alabama) (84) was added to a final concentration of 3 μM. The mixture was incubated for 30 min at room temperature and then precipitated by the addition of 1ml of ice-cold acetone. The samples were held at –20°C for 10 min, after which the precipitate was collected by centrifugation at 18,000 x g at 4°C for 15 min. The pellets were air-dried, resuspended in non-reducing SDS sample buffer and examined by SDS-PAGE and western blotting. All samples were run with controls that had not been exposed to PEG-OPSS.

Lyz^{103} expression and purification

Since wild type Lyz^{103} lyses cells rapidly when over-expressed, the enzyme activity was abolished by replacing the catalytic Asp with Asn. Hence, all purified

Lyz^{103} cited in this work refers to Lyz^{103}_{D52N}. pET *lyz^{103}_{D52N} cHis* was transformed into BL21 (DE3) cells (Invitrogen) harboring pLysS, and fresh transformants were cultured and induced for 1 hr at 30°C. Cells were collected at 4K RPM for 30 min at 4°C in a Sorvall RC-3B centrifuge and resuspended in Lyz^{103} buffer (20 mM Tris-HCl (pH 8), 100 mM NaCl). Protease Inhibitor Cocktail for His-tagged protein (Sigma)) was added as per manufacturer's instructions and cells were lysed by passage through a French pressure cell (Spectronic Instruments, Rochester, N.Y.) at 20,000 lb/in^2. After removing unlysed cells and debris, the lysate was filtered through a 0.2 μm syringe filter. The cleared lysate was then applied to Talon Metal Affinity Resin (Clontech). Protein was eluted in elution buffer (20 mM Tris-HCl, 100 mM NaCl, 500 mM imidazole, pH 8), and was used with no further purification.

CDAP cleavage

Purified Lyz^{103}_{D52N} was precipitated and resuspended in 100 μl CDAP buffer (4M Guanidine-HCl, 0.1M Citrate, pH 3). 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP, Sigma) prepared fresh in CDAP buffer, was added in 1,000-fold molar excess and was incubated at room temperature for 15 min. NH_4OH (EMD Chemicals) was added to 1M final concentration, and the mixture was incubated at room temperature for 3 hrs (97, 156). Total protein was precipitated by addition of 1 ml of ice cold ethanol and incubation overnight at -20°C. Samples were pelleted, dried, and resuspended in sample loading buffer with or without 5% β-mercaptoethanol, as indicated. Equivalent amounts were analyzed by SDS-PAGE and western blotting.

Results and discussion

The lysozyme from bacteriophage ERA103 has an N-terminal SAR domain

We first wished to confirm that the N-terminal hydrophobic domain of Lyz^{103} is a SAR domain. Inductions of lyz^{103} resulted in lysis of *E. coli*, even without a holin gene present, indicating a spontaneous release from the membrane and consequent lysis (Fig. 2.2A). Moreover, like P1 Lyz, Lyz^{103} was found in both the soluble and membrane fractions of cells (Fig. 2.3A). For both proteins, the membrane-associated and soluble forms migrate identically in SDS-PAGE, indicating the latter is not derived by proteolytic cleavage of the former. Additionally, energy poisons such as DNP accelerated the lysis of cultures expressing lyz^{103} (Fig. 2.2A), indicating that collapse of the proton motive force facilitates the membrane release and activation of Lyz^{103}, as it does for P1 Lyz (159). The SAR domains characterized in P1 Lyz and R^{21}differ from conventional TMDs in that they have a high content of weakly hydrophobic and uncharged polar residues such as Ala, Gly, and Ser . In both cases, substitution of Leu residues for Gly residues (3 in P1 Lyz (Figs. 2.2, 2.3C); 2 in R^{21} (133)) in the SAR domain blocked release from the membrane and host lysis. The same molecular and cellular phenotype was observed for Lyz^{103} when Gly residues at positions 9 and 10 in the SAR domain were converted to Leu (Figs. 2.2C, 2.3D). These results are consistent with the requirement for the SAR domains to exit the membrane, liberating the thiol in the SAR domain to attack the inhibitory disulfide. Thus Lyz^{103} is the third SAR endolysin to be characterized in terms of its physiological and topological

characteristics. The predicted amino acid sequence of Lyz^{103} is identical to that of the lysozyme from bacteriophage ϕEa1h (69). Thus, the puzzling lethality observed when *lyz*$^{\phi Ea1h}$ was expressed in the absence of its holin (70) is due to the fact that, as a SAR endolysin, it reaches the periplasm by a holin-independent mechanism.

The activity of Lyz^{103} is regulated by a disulfide "cage"

Lyz^{103}, like T4 E but unlike P1 Lyz, has an Asp as the middle residue of its catalytic triad, and thus is not regulated by covalent blocking of an active site Cys (Fig. 2.1A). Nonetheless, the N-terminal catalytic domain of Lyz^{103} contains three Cys residues (Cys_{12}, Cys_{42}, Cys_{45}), including one in the SAR domain, suggesting that it might, like P1 Lyz, exist in two isomeric forms differing in the arrangement of intramolecular disulfide bonds. To test this possibility, a C12S mutant of Lyz^{103} was tested for function. This mutant was found to be lytically inactive (Fig. 2.2B), although it was released from the inner membrane as efficiently as the wild type protein (Fig. 2.3B). Moreover, addition of the reducing agent, dithiothreitol, to cells expressing the C12S mutant resulted in lysis (Fig. 2.2B).

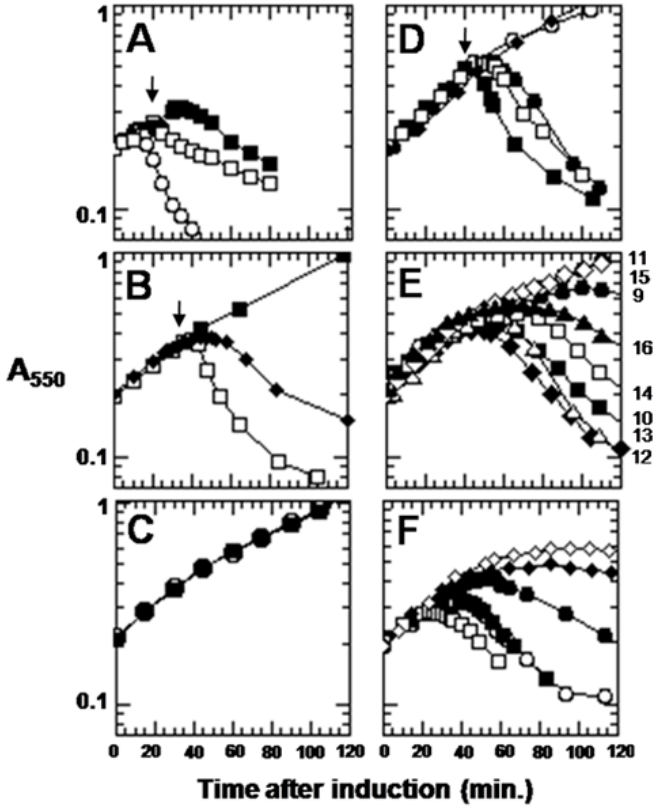

FIG. 2.2. Lysis profiles of SAR endolysin derivatives. In each experiment, cultures of XL1 Blue containing P1 Lyz or Lyz^{103} in pJF118EH were induced at time 0. Optical density of the culture was followed as a function of time in mins. (*A*) Expression of Lyz^{103} lyses the cell independently of a holin. (○) P1 Lyz, (■) Lyz^{103}, (□) Lyz^{103} + 10mM DNP (arrow) 20 mins after induction. . (*B*) Lysis by Lyz^{103} is dependent on the SAR domain Cys. (■) Lyz^{103} C12S, (□) Lyz^{103} C12S + 1mM DTT at 36 mins (arrow) after induction, (♦) Lyz^{103} C12,42,45S. (*C*) P1 Lyz and Lyz^{103} can be locked into the membrane by Leu titration into the SAR domain. (○) Lyz^{103} G9, 10L, (■) P1 Lyz G8,9,10L. (*D*) P1 Lyz and Lyz^{103} can be interconverted. $Lyz^{P1\rightarrow 103}$ was achieved by two amino acid mutations: A41C and C51D (●). (○) $Lyz^{P1\rightarrow 103}$ C13S, (■) $Lyz^{P1\rightarrow 103}$ C13S + 1mM DTT at 45 min., $Lyz^{103\rightarrow P1}$ was achieved by two amino acid mutations: C42S and D52C (□), (♦) $Lyz^{103\rightarrow P1}$ C12S. (*E*) There is optimal positioning of the SAR domain Cys, Cys substitutions are labeled adjacent to their respective curves. (●) $Lyz^{P1\rightarrow 103}$ G9C, (■) $Lyz^{P1\rightarrow 103}$ G10C, (○)$Lyz^{P1\rightarrow 103}$ A11C, (♦) $Lyz^{P1\rightarrow 103}$I12C, (Δ) $Lyz^{P1\rightarrow 103}$, (□) $Lyz^{P1\rightarrow 103}$A14C, (◊) $Lyz^{P1\rightarrow 103}$I15C, (▲) $Lyz^{P1\rightarrow 103}$ A16C. (*F*) The placement of the caging disulfide is stringent. (●) P1 Lyz I38C C51D, (○) P1 Lyz G39C C51D, (■) P1 Lyz N40C C51D, (□) P1 Lyz C13S I38C C51D, (♦) P1 Lyz C13S G39C C51D, (◊) P1 Lyz C13S N40C C51D.

Finally, the triple mutant, $Lyz^{103}_{C12, 42,45S}$, was found to be lytically active (Fig. 2.2B). The behavior of these two mutants is consistent with the model that nascent Lyz^{103} is inactive because of an inhibitory Cys_{42}-Cys_{45} disulfide that is disrupted by Cys_{12} after the release of the SAR domain from the membrane. Since the inhibitory disulfide predicted to exist in nascent Lyz^{103} involves cysteines that flank an essential catalytic residue (Fig. 2.1A), we refer to it as a "caging" disulfide, to distinguish it from the inactivating disulfide present in P1 Lyz. It is of interest to note that a similar inactive, disulfide-caged form of T4 gp*e* was constructed 21 years ago (90). In this form of T4 gp*e*, the active site residues are unaltered, but an engineered disulfide occludes the active site. As shown here with Lyz^{103} and earlier with P1 Lyz the disulfide-caged form of T4 gp*e* could be activated by reducing agents.

Disulfide bond isomerization

The results presented above indicate that the Cys_{42}-Cys_{45} disulfide cages the catalytic Glu43 in inactive Lyz^{103} and that the Cys_{12} thiol from the SAR domain is required for activation, but do not reveal whether one of the two Cys residues flanking Glu_{43} participates in a disulfide exchange with Cys_{12} and, if so, which one (Fig. 2.1A). To address this question, we used the sulfhydryl cyanylation reagent 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP).

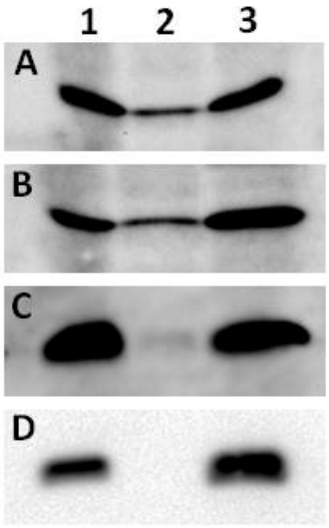

FIG. 2.3. Cellular locations of bacteriophages P1 and ERA103 SAR endolysins. For all panels, the total protein is represented by lane 1, the soluble protein fraction is represented by lane 2, and the membrane fraction is represented by lane 3. (*A*) Lyz^{103} exists in the soluble and membrane fractions of the cell. (*B*) Lyz^{103} C12S displays a non-lytic phenotype, but is released from the cytoplasmic membrane as efficiently as the wild type protein. (*C*) P1 Lyz can be confined to the inner membrane by replacing three Glycines (residues 8-10) with Leucine residues. (*D*) Lyz^{103} can be confined to the inner membrane by replacing two Glycines (residues 9 and 10) with Leucine residues.

Treatment of a CDAP-cyanylated protein with a strong base cleaves the protein on the N-terminal side of the cyanylated Cys, leaving an N-terminal 2-iminothiazolidine-4-carboxylyl (itz)-modification on the C-terminal product (97, 156). If activated Lyz^{103} has a disulfide between Cys_{12} and Cys_{42}, Cys_{45} will be open for CDAP modification, and subsequent cleavage would create two polypeptide fragments which could be resolved from uncleaved protein without the presence of reducing agent (Fig. 2.4A). In contrast, a linkage between Cys_{12} and Cys_{45} will leave Cys_{42} open for modification. In this case, cleavage with strong base would result in polypeptide fragments that were still

covalently linked and which would require treatment with reducing agent to be resolved from uncleaved material (Fig. 2.4B). When purified Lyz^{103} was subjected to CDAP treatment and alkaline cleavage and analyzed by Western blot with anti-oligohistidine tag antibody, a polypeptide product of appropriate size (16 KDa) could be resolved from uncleaved protein (20 KDa) by SDS-PAGE without reducing agent (Fig. 2.4C). The yield of cleavage product was low, approximately 10-15% of total protein. However, treatment of the CDAP-modified product with PEG-OPSS revealed that >60% of the protein was cyanylated by CDAP and therefore unPEGylated (not shown), indicating that that the cleavage reaction, rather than sulfhydryl accessibility, was limiting, presumably due to the high incidence of the competing β-elimination reaction, as noted elsewhere (156). Taken together, these results indicate that the Lyz^{103} activating linkage resulting from the disulfide bond isomerization after SAR extraction is between Cys_{12} and Cys_{42}, rather than between Cys_{12} and Cys_{45}, as predicted from the P1 Lyz and Lyz^{103} sequence alignments (Fig. 2.1A). Thus, the position of the disulfide-bond exchange is not conserved.

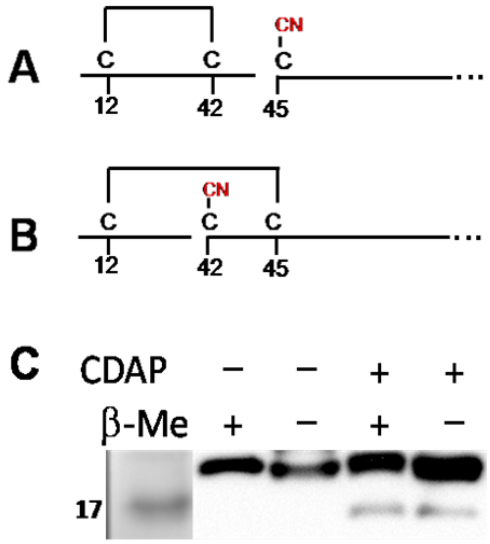

FIG. 2.4. Analysis of the 'cage' relieving disulfide. Purified Lyz^{103} $_{D52N}$ cHis was subjected to CDAP cyanylation and alkaline cleavage. Proteins were resolved by SDS-PAGE with and without reducing agent, as indicated. (A) A cage relieving disulfide between Cys_{12} and Cys_{42} will allow Cys_{45} to be modified by CDAP. Cleavage of this protein results in a C-terminal cleavage fragment that is resolvable without addition of reducing agent (β-Me) (B) A cage relieving disulfide between Cys_{12} and Cys_{45} results in a cleavage product that is resolved only upon addition of β-Me. (C) Treatment of Lyz^{103} D52N with 1000 fold molar excess of CDAP reagent resulted in a resolvable C-terminal cleavage product (16KDa) indicating the formation of a Cys_{12}-Cys_{42} disulfide.

The P1 Lyz and Lyz^{103} regulatory schemes are inter-convertible

Comparing the crystal structures of the active and an inactive form of P1 Lyz demonstrates that much of the N-terminus of the protein is capable of adopting markedly different conformations depending upon the placement of a single intramolecular disulfide (158). To explore the structural malleability of the N-terminal catalytic domain of P1 Lyz (Glu_{42}, Cys_{51}, and Thr_{57}), we attempted to convert P1 Lyz, in which the inactivating disulfide covalently blocks a catalytic Cys, into a Lyz^{103}-type endolysin,

with a caging disulfide sequestering a catalytic Glu, by introducing the A41C and C51D mutations (Fig. 2.1B, Table 2.1). This mutant, $Lyz^{P1\rightarrow103}$, was found to be lytically active, although lysis was delayed and more gradual when compared with the wild type enzyme (Fig. 2.2D, Table 2.1). As seen with P1 Lyz, removal of the SAR domain Cys in $Lyz^{P1\rightarrow103}{}_{C13S}$ rendered the enzyme dependent upon the addition of an exogenous reductant (Fig. 2.2D). Moreover, introducing the C42S and D52C mutations into Lyz^{103} (Fig. 2.1B) converted this enzyme into a P1 Lyz-type endolysin that still required the presence of the Cys residue in its SAR domain for lytic function (Fig. 2.2D). Thus, with regard to the nature of inhibitory disulfide and the use of a catalytic Cys or Asp, P1 Lyz and Lyz^{103} are fully inter-convertible. Since we had polyclonal antibodies at hand for P1 Lyz which were much more efficient than the antibodies against the oligo-histidine tag for Lyz^{103}, the $Lyz^{P1\rightarrow103}$ construct was selected for further analysis.

TABLE 2.1. Lyz^{103} mutant alleles. The catalytic triads, cysteines participating in enzyme regulation, mutations required for the $Lyz^{P1\rightarrow103}$ and $Lyz^{103\rightarrow P1}$ conversions, and approximate lysis times of P1 Lyz, Lyz^{103}, $Lyz^{P1\rightarrow103}$, and $Lyz^{103\rightarrow P1}$ are listed.

	P1 Lyz	**Lyz^{103}**	**$Lyz^{P1\rightarrow103}$**	**$Lyz^{103\rightarrow P1}$**
Catalytic triad	E42, C51, T57	E43, D52, T58	E42, D51, T57	E43, C52, T58
Regulating cysteines	C13 (SAR), C44, C51 (catalytic)	C12 (SAR), C42, C44	C13 (SAR), C41, C44	C12 (SAR), C44, C52 (catalytic)
Conversion mutations	n/a	n/a	A41C, C51D	C42S, D52C
Lysis time (min.)	15	35	50	45

Cysteines located in the SAR domain and catalytic cysteines are indicated.

Optimal positioning of the SAR Cys residue

Since we had previously demonstrated that the *in vivo* activity of P1 Lyz depended upon the placement of the Cys_{13} residue in its SAR domain (158), we reasoned that the poor lytic profile obtained with the $Lyz^{P1\rightarrow103}$ protein might be due to the suboptimal positioning of this critical residue. To test this notion, constructs were generated in which the Cys_{13} residue in the SAR domain was moved to eight new positions, four in each direction, from the parental site in the context of $Lyz^{P1\rightarrow103}$. The results validated the hypothesis, in that earlier and sharper lysis profiles were obtained with a Cys occupying positions 12, 10, 14 and 16, in order of apparent lytic activity (Fig. 2.2E). This pattern is different from that observed using the equivalent substitutions in P1 Lyz, where the Cys_{13} form is the most active and moving the thiol to position 12 eliminated the ability to activate (158). These differences could be due to a combination of factors, since the activation of P1 Lyz involves major structural changes within the N-terminal catalytic domain and all of the mutations necessary to convert P1 Lyz into $Lyz^{P1\rightarrow103}$ and its derivatives occurred within this domain.

Strict positional requirement for the inhibitory disulfide cage

We next determined the effect of moving the position of the first Cys (Cys_{41}) of the inhibitory disulfide cage in $Lyz^{P1\rightarrow103}$ towards the N-terminus of the protein (Fig. 2.1B). All three variants (N40C, G39C, and I38C) tested were lytic (Fig. 2.2F). Surprisingly, the lethality of these proteins was not completely dependent upon the presence of a Cys in the SAR domain, as lysis still occurs, although much later and more

gradually (Fig. 2.2F). The simplest explanation for this finding is that disulfides between Cys_{38}, Cys_{39}, or Cys_{40} and Cys_{44} do not effectively cage the active site glutamate. We directly assessed the status of the cysteines in $Lyz^{P1\rightarrow103}$, $Lyz^{P1\rightarrow103}{}_{C13S}$, and P1 $Lyz_{C13S, I38C, C51D}$, the most lytically active cage mutant, by testing their susceptibility to modification by PEG-OPSS. PEG-OPSS has a PEG 5000 moiety affixed to an orthopyridyl−disulfide. When reacted with a free sulfhydryl, PEG-OPSS will covalently bind, and will shift the apparent molecular weight of the protein (84). As expected, the active endolysin, $Lyz^{P1\rightarrow103}$, was found to have a single free cysteine (Cys_{13}) while the inactive enzyme, $Lyz^{P1\rightarrow103}{}_{C13S}$, had none (Fig. 2.5). The lytically active P1 $Lyz_{C13S, I38C, C51D}$ protein was not modified with PEG-OPSS indicating that all six of its cysteine residues were present in disulfide linkages. Thus, while the Cys_{38}-Cys_{44} disulfide forms, it is not inhibitory.

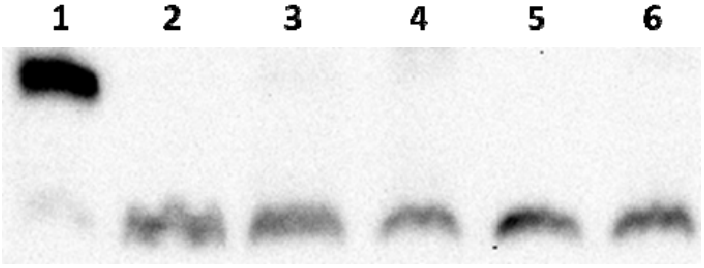

FIG. 2.5. Assessment of the status of cysteine residues**.** Lane 1 contains molecular weight standards. P1 $Lyz^{\rightarrow103}$ is an active endolysin and has a single free sulfhydryl. It is modified by PEG-OPSS (Lane 2). $Lyz^{P1\rightarrow103}$ C13S has no free cysteine residues and is not modified by PEG-OPSS (Lane 3). $Lyz^{P1\rightarrow103}$ C13S, I38C, C51D is an active endolysin despite the presence of the caging disulfide. As confirmation of the presence of this disulfide, the protein is not modified by PEG-OPSS indicating that all of its cysteines are in disulfide bonds. Lanes 5-7 contain UnPEGylated $Lyz^{P1\rightarrow103}$, $Lyz^{P1\rightarrow103}$ C13S, and $Lyz^{P1\rightarrow103}$ C13S, I38C, respectively.

Conclusion

Two modes of SAR endolysin negative regulation have previously been experimentally established. The first is that of an inactivating disulfide involving a catalytic Cys, as is present in P1 Lyz(158); the presence of the disulfide not only precludes participation of the catalytic Cys but also stabilizes a drastically misfolded N-terminal domain. The second mode was found in R^{21}, in which, although the inactive protein has more subtle folding defects, the active site is compromised by the proximity of the membrane to which it is tethered by the embedded SAR domain. The results presented here demonstrate that Lyz^{103} represents a third distinct mode of SAR endolysin regulation: disulfide-bond caging of the active site in the inactive form of the enzyme (Fig. 2.6). Despite these differences in the regulatory mode, however, the two proteins appear to be completely interconvertible by simply repositioning the regulatory Cys residues and adjusting the position of the activating Cys residue in the SAR domain. These findings serve to emphasize both the plasticity of the SAR regulatory system and the importance of regulating the timing of lysis in the infection cycle.

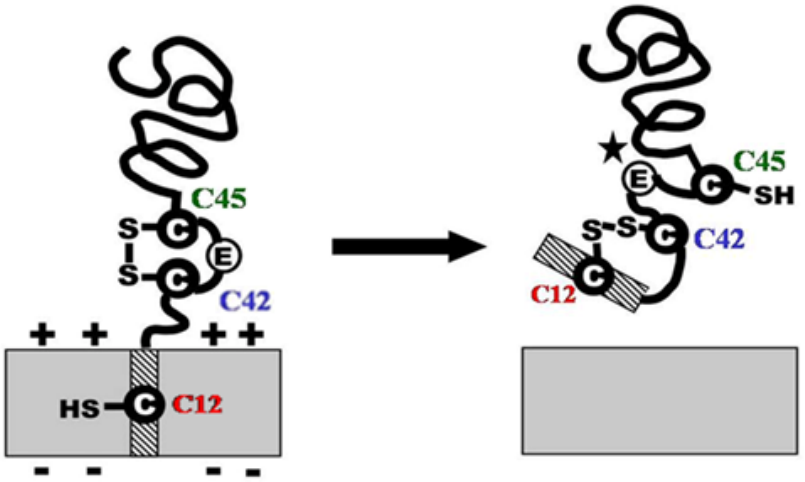

FIG. 2.6. Model for Lyz^{103} activation. In its inactive form, Lyz^{103} is tethered to the charged cytoplasmic membrane. A caging disulfide exists between Cys42 (blue) and Cys45 (green) that sequesters a catalytic Glu from the remaining members of the catalytic triad. Upon SAR domain (hatched bar) release from the membrane, the SAR thiol, Cys_{12} (red) attacks the caging disulfide at Cys_{42}, forming a cage relieving disulfide. The catalytic Glu is released from the cage and can form the catalytic cleft.

Finally, it should be noted that another instance of disulfide-bonding affecting the steric blockage of an enzyme active site has been reported in NADP-malate dehydrogenase from chloroplasts (26). In this case, intramolecular steric regulation is achieved when a C-terminal helix, positioned and stabilized by a disulfide bond, occupies the active site and forms hydrogen bonds with catalytic residues. The negative regulation is released when the conserved disulfide is reduced by thioredoxin in response to light stimulation. A key difference is that the Lyz^{103} activation, driven by the release of the SAR domain from the bilayer, is almost certainly irreversible and thus the stabilities of the inactive and active states are unlikely to be comparable.

CHAPTER III

REGULATION OF A MURALYTIC ENZYME BY DYNAMIC MEMBRANE TOPOLOGY

Introduction

It has been estimated that phage lysis events occur $>10^{29}$ times a day and thus represent one of the major factors in the homeostasis of biomass (134). In general, lysis requires the enzymatic degradation of the host cell wall by a phage-encoded endolysin. For a successful infection, endolysin activity must be blocked throughout the latent period, typically 30 - 60 min, until, at lysis, there is a short period of acute enzymatic catalysis on the time scale of seconds (53). The canonical phage endolysins, like λ R, T4 E, and T7 gp3.5, meet these requirements passively, in that they are synthesized as fully active enzymes but are restricted to the cytoplasm, where they have no substrate, until the instant when the phage-encoded holins trigger and open holes in the membrane (64, 162). This robust regulation was thought to be universal until the discovery of endolysins with secretory signals (118, 159, 162). Since secretory endolysins are not restricted from their substrate by the membrane, they require not only strict post-secretory negative regulation but also the means to become activated in a timely manner. Phage-encoded holins control the timing of lysis by causing a temporally-regulated, saltatory disruption of the membrane but show no specificity for endolysins (53, 148, 163). Thus the regulatory system of a secretory endolysin must depend on the integrity of the membrane, such that the loss of integrity causes quantitative activation of the muralytic function. This requirement is met in SAR endolysins.

A SAR (Signal Anchor-Release) domain is an N-terminal transmembrane domain (TMD) that has the unique ability to exit the lipid bilayer completely (159). The first SAR domain was identified in P1 Lyz, the lysozyme of bacteriophage P1 and a homolog of the well-studied T4 lysozyme, protein E. During the latent period, P1 Lyz is expressed and secreted to the periplasm, where it accumulates tethered to the membrane by the SAR domain. Premature destruction of the cell wall is avoided because in the membrane-tethered form, P1 Lyz is catalytically inactive, both conformationally and covalently. In the inactive form (^{i}P1 Lyz) , the N-terminal domain is in a radically different fold from the active enzyme (^{a}P1 Lyz) and completely lacks its catalytic cleft. In addition, a catalytic Cys residue is occupied in a disulfide bond. Activation occurs via two dramatic transitions concomitant with the escape of the SAR domain from the bilayer (158, 159). Covalent activation is achieved when a free thiol in the SAR domain causes a disulfide bond isomerization, releasing the catalytic cysteine. In addition, the entire catalytic domain undergoes a large conformational reorganization, unwinding 3 α-helices to form 3 β-strands in the active site. The extracted SAR domain itself contributes only the liberating thiol, remains largely helical, and makes no intimate contacts with the body of the enzyme.

This elegant regulation, with its conformational and covalent levels, ensures that the phage morphogenesis period is not shortened by premature lysis. However, many SAR endolysins do not have cysteine residues either in the transmembrane domain or the active site and thus must be regulated in a different way. Here, we characterize one such enzyme, R^{21} (Fig. 3.1), the lysozyme of the lambdoid phage 21 (9, 20, 159), by genetic,

biochemical, and structural means. R^{21} has the canonical Glu35-Asp44-Thr50 catalytic triad and a cysteine-free SAR domain which is not replaceable by that of P1 Lyz. The results are discussed in terms of unique phase-specific roles for the SAR domain in the inactive and active forms of the endolysin.

Materials and methods

Bacterial strains and growth conditions

The *Escherichia coli* strains XL1-Blue and MG1655 *ΔtonA lacI*^{q} have been previously described (127). To create MG1655 *ΔtonA lacI*^{q} *ΔlacY*::*Kan* (RY16323), MG1655 *ΔtonA lacI*^{q} (RY17307) was transduced to Kan^R with a P1 lysate grown on BW25113 *ΔlacY::Kan* (RY17459) using standard procedures (93). Standard conditions for the growth of cultures and the monitoring of lysis kinetics have been previously described (29, 127). When indicated, isopropyl-β-D-thiogalactopyranoside (IPTG) and chloroform ($CHCl_3$) were added to final concentrations of 1 mM and 1%, respectively. Phase-contrast images were taken 100 minutes after induction using a Nikon Digital Sight DS-5M camera on a Nikon OPTIPHOT microscope.

Standard DNA procedures

Procedures for the isolation of plasmid DNA, DNA amplification by PCR, PCR product purification, DNA transformation and DNA sequencing have been previously described (52, 126, 128). Oligonucleotides were obtained from Integrated DNA

Technologies, Coralville, IA, and used without further purification. All enzymes were purchased from New England Biolabs, with the exception of Pfu polymerase which was from Stratagene. Site-directed mutagenesis was performed using the QuikChange kit (Stratagene). DNA sequencing was performed at the Laboratory for Plant Genome Technology of Texas A&M Agrilife.

Plasmids

To construct the plasmid pZE-R^{21}, a derivative of the IPTG-inducible, medium-copy vector pZE12 (86), the *R^{21}* gene was amplified from pJFR^{21} (159) and cloned into pZE12. The control plasmid pZE-luc (86) is pZE12 with a luciferase insert. An allele encoding R^{21} carrying the c-MYC tag (EQKLISEEDL) at the C-terminus was constructed using ExSite PCR (Stratagene). To construct pZE12 with the chimeric allele encoding P1 $Lyz_{1\text{-}26}\Phi R^{21}{}_{27\text{-}165}$, the DNA encoding the SAR domain was amplified from pJFLyz (159) and the PCR product was then used as a mega-primer to replace the DNA encoding the R^{21} SAR domain. Similarly, pZE12 derivatives carrying chimeric alleles (Fig. 3.1) encoding R^{21} with an artificial TMD ($MK_3L_2I_2V_8I_2L_2$) added at the N-terminus or with the PelB signal sequence replacing the residues 1-26 were constructed from pZER^{21} using ExSite PCR. The plasmids were designated pZE-$TMD_{art}\Phi R^{21}$ and pZE-$pelB_{ss}\Phi R^{21}{}_{27\text{-}165}$, respectively. The T7 over-expression plasmid pETR^{21} was constructed by amplifying the *R^{21}* gene from pBR121 (9) with primers that added DNA encoding the sequence GGHHHHHHGG at the C-terminus and inserting the product between the NdeI and BamHI sites of pET11a. Similarly, to construct pETpelBΦ $R^{21}{}_{27\text{-}165}$, the DNA

encoding the truncated $R^{21}_{27\text{-}165}$ periplasmic domain was amplified and cloned into the NcoI and XhoI sites of pET22b (Novagen, San Diego, CA).

Expression and protein purification

For production of $^{a}R^{21}$, the plasmid $pETR^{21}$ was transformed into Rosetta (DE3) pLysS (Novagen, Madison, WI). Transformants were grown in Terrific Broth (BD, Franklin Lake, NJ) to A600 ~ 1 - 1.5 and induced with 0.5 mM IPTG. The culture was arrested between 60 to 90 minutes after induction, depending on the extent of lysis. The partially lysed culture was spun at 4000 rpm for 30 minutes at 4°C. Harvested cell pellets were stored at -20°C. The cells were resuspended in buffer (25 mM Tris-HCl (pH 8.0), 500 mM NaCl, 3 mM β-Mercaptoethanol, 1% Triton X-100, and EDTA-free protease inhibitor cocktail (Roche Molecular)), and were disrupted by 2 passages through a French pressure cell at ~16,000 psi. After the lysate was centrifuged at 40,000 rpm (Beckman Ti50.2) for 60 - 90 min, the supernatant was collected and applied to a 5ml HiTrap-chelating Sepharose column (Amersham Pharmacia) connected in series, charged with $NiCl_2$ and equilibrated with buffer A (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 2 mM β-mercaptoethanol and 10 mM imidazole). The column was washed with buffer A followed by increased concentration of imidazole (25 mM). Bound protein was eluted with a 25-500 mM imidazole gradient and the $^{a}R^{21}$ enriched fractions were pooled, concentrated, and dialyzed against a buffer containing 25 mM of sodium acetate pH 5.0, 25 mM NaCl and 1 mM DTT. All other contaminating proteins were precipitated upon dialysis with the minor fraction of $^{a}R^{21}$. Supernatant containing $^{a}R^{21}$

was concentrated and 50 µl aliquots were flash frozen in liquid nitrogen and stored at -80 °C for further use.

For production of $^{i}R^{21}$, BL21(DE3) cells were transformed with pETpelBΦ $R^{21}_{27\text{-}165}$ and grown at 37°C in LB medium containing 100 µg/ml of carbenicillin until A_{600} ~ 0.8. The culture was induced with 0.5 mM IPTG and incubated for 18 h at 25 °C. $^{i}R^{21}$ protein was purified through a Ni-affinity column and size-exclusion chromatography as described above.

Crystallization and structure determination

Purified $^{a}R^{21}$ at 20~30 mg/ml concentration was used to screen a large number of commercially available crystallization conditions by vapor diffusion method. Diffraction quality crystals of $^{a}R^{21}$ were obtained in a condition containing 150 mM sodium acetate pH 4.6 and 2 M sodium formate. $^{i}R^{21}$ was crystallized in the condition of 100 mM Bis-Tris propane pH 7, 2.5M sodium nitrate. The best crystals grew in 2 weeks and were flash frozen in mother liquor with 20% glycerol then stored in liquid nitrogen. Data sets were collected in the lab (Rigaku MM007HF generator with Bruker Smart 6000 CCD detector) and synchrotron beamlines at Advanced Photon Source (APS 23-ID and 19-ID), Argonne National Laboratory, Chicago.

The initial structure was solved by single-wavelength dispersion using an iodine derivative obtained by quick soaking of the crystal in the cryo with 0.5M NaI for 30~90s. Then the model was refined against the native dataset iteratively until the R/R_{free} reached 0.212/0.249 at 1.95 Å for $^{a}R^{21}$ and 0.199/0.231 at 1.7 Å for $^{i}R^{21}$. The phasing

and refinement were carried out with the Phenix package (2). The models were built in Coot (39). PyMol (Delano Scientific; San Carlos CA) was used for structure analysis and rendering.

$^{a}R^{21}$ crystallizes in space group $P2_12_12_1$ with unit cell dimensions of a = 78.2 Å, b = 94.8 Å, c = 97.7 Å, α = β = γ = 90° (Table 3.1) and four $^{a}R^{21}$ molecules in the asymmetric unit (ASU). All the residues can be traced in the electron density map. The first three molecules are nearly identical (backbone RMS = 0.4~0.5 Å). The fourth molecule in the ASU is the same as the other three except two loops, G27-S38 and T137-G146. $^{i}R^{21}$ crystallizes in space group $P2_12_12$ with unit cell dimensions of a = 64.2 Å, b = 109.7 Å, c = 45.0 Å, α = β = γ = 90° and two $^{i}R^{21}$ molecules in the ASU (backbone RMS =0.2 Å). The first seven and last four residues are invisible in the structure. For both $^{a}R^{21}$ and $^{i}R^{21}$, the first molecule (chain A) in the asymmetric unit is used as the representative structure.

TABLE 3.1. Data collection and refinement statistics

	full-length $^{a}R^{21}$	**truncated $^{i}R^{21}_{27\text{-}165}$**
Data collection[1]		
Space group	$P2_12_12_1$	$P2_12_12$
Cell dimensions a, b, c (Å)	78.2, 94.8, 97.7	64.2, 109.7, 45.0
Wavelength (Å)	0.9795	0.9796
Resolution (Å)	1.95 (2.05-1.95)	1.70 (1.79-1.70)
I/sigI	19.3 (3.8)	23.9 (4.9)
Completeness (%)	99.6 (99.2)	100 (99.8)
R_{sym}	0.031 (0.268)	0.023 (0.213)
Redundancy	7.2 (5.7)	6.1 (6.1)
Refinement[2]		
# reflections	50721	35715
R_{work}/R_{free}	0.2116/0.2485	0.1991/0.2311
# atoms	5282	2483
RMSD bonds	0.005	0.004
RMSD angles	0.700	0.647

[1]The R^{21} datasets were collected at AP 23ID and processed with HKL2000.
[2]Refinement was done with Phenix.

Subcellular fractionation

To determine whether a particular protein was present in the soluble or membrane fractions, 25 mls of an induced culture were collected by centrifugation at 5000 x g and resuspended in 2 ml of buffer (0.1M sodium phosphate, 0.1M KCl, 5 mM EDTA, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, pH 7.0) and then disrupted by passage through a French pressure cell (Spectronic Instruments, Rochester, N.Y.) at 16,000 lb/in^2. The membrane and soluble fractions were separated by

centrifugation at 100,000 x g for 60 min. Equivalent amounts of the fractions were examined by SDS-PAGE and Western blotting as described below. To separate the periplasmic fraction, cells expressing pZER[21] derivatives were collected after 30 minutes of induction by centrifugation at 5000 x g. Cells were resuspended in 0.5 ml of spheroplasting buffer (25% sucrose, 30 mM TrisHCl, pH 8). Ten µl of 0.25 M EDTA, ten µl of lysozyme (20 mg/ml in water), and 500 µl of distilled water were added sequentially (24). The mixture was inverted several times and incubated for 5 minutes at room temperature. Examination under the microscope showed that ~ 95% of the cells were spheroplasts. Spheroplasts were pelleted by centrifugation at 8,000 x g for 30 minutes at 4 °C. Equivalent amounts of the supernatant (periplasm) and pellet (cytoplasm plus membrane) fractions were analyzed by SDS-PAGE and Western blotting as described below.

SDS-PAGE and Western blotting

SDS-PAGE, Western blotting, and immunodetection experiments were performed as previously described (52). A mouse monoclonal antibody against the c-MYC tag was purchased from Covance and was used at a 1:1000 dilution. The anti-mouse IgG horseradish peroxidase-conjugated secondary antibody was supplied with the SuperSignal chemiluminescence kit (Pierce) and was used at a 1:5000 dilution. Blots were developed using the West Femto SuperSignal chemiluminescence kit (Pierce). Chemiluminescent images were acquired using a Biorad ChemiDoc XRS.

Accession codes

The atomic coordinates and structure factors have been deposited in the PDB with the accession codes 3HDE and 3HDF for $^{a}R^{21}$ and $^{i}R^{21}$, respectively.

Results

The R^{21} class of SAR endolysins predominates in sequenced phage genomes

Prior to the discovery of SAR endolysins, all phage lysozymes that had been characterized, like T4 E, lambda R, and T7 gp*3.5*, were found to be soluble enzymes that accumulate fully folded and enzymatically active in the cytoplasm (64, 162). Thus degradation of the cell wall, and subsequent lysis, could be controlled by maintaining the integrity of the membrane until it was permeabilized at the appropriate time by the phage-encoded holin. However, when we surveyed 196 sequenced phage genomes, we found 58 endolysin genes with the Glu-8aa-Asp/Cys-5aa-Thr catalytic triad (Fig. 3.1) that characterizes the canonical T4 E, a true lysozyme. Unexpectedly, most (43/58) have N-terminal hydrophobic domains with the characteristics of SAR domains (Table 3.2); i.e., predicted N-terminal transmembrane α-helices which have a higher representation of Gly, Ala, and Ser residues compared to typical transmembrane domains (TMDs). Thus SAR-type endolysin regulation is the rule, rather than the exception, at least among the T4 E homologs of sequenced phage genomes. Moreover, among these 43 SAR endolysins, only nine have a Cys residue within the SAR domain and a catalytic cysteine as observed in P1 Lyz and Lyz^{103} (74, 158, 159). This suggests that, at least among the

sequenced genomes of phages known to be viable, the disulfide bond-dependent regulation of P1 Lyz is the exception rather than the rule, and that the remaining 34 SAR enzymes must be regulated differently. One of these enzymes, R^{21}, has already been confirmed as a SAR endolysin (159), and was chosen for detailed study.

TABLE 3.2. SAR endolysins dominate the true lysozyme proteins.

endolysin subtype	# phage genomes	characterized enzymes[3]
cytoplasmic	15	T4 E, P22 gp*19*
SAR with Cys	9	P1 Lyz
SAR without Cys	34	R^{21}
total	58	

Endolysin subtype was defined by existence of predicted TMDs or SAR domains at N-terminus. Absence of both resulted in scoring as cytoplasmic. 196 phages of Gram-negative hosts were analyzed from the Phage category of the GenBank Genome database. Characterized enzymes are those which have been purified and characterized by biochemical or structural methods.

The SAR domain of R^{21} is essential for its lysozyme activity

Previously, we demonstrated that the SAR domain of P1 Lyz is not an essential component of the structure or the enzymatic activity of the protein. In other words, the SAR domain is used only to regulate by tethering the protein to the membrane and, after release from the bilayer, to supply a cysteine for activation of the enzyme (158). To determine whether this was also true for R^{21}, we replaced the N-terminal 26 residues of R^{21} with the cleavable secretory signal sequence $PelB_{ss}$ (76) (Fig. 3.1). When the *$pelB_{ss}$-R^{21}_{27-165}* chimera was expressed in *Escherichia coli*, no significant cell lysis was observed (Fig. 3.2A) even though large amounts of the processed R^{21}_{27-165} protein

accumulated in the periplasm (Fig. 3.2B). Moreover, an *in vitro* lysozyme assay of the purified protein did not detect lysozyme activity (data not shown). The lack of *in vivo* and *in vitro* activity of the truncated R^{21} indicated that the SAR domain of R^{21} is necessary for its enzymatic function.

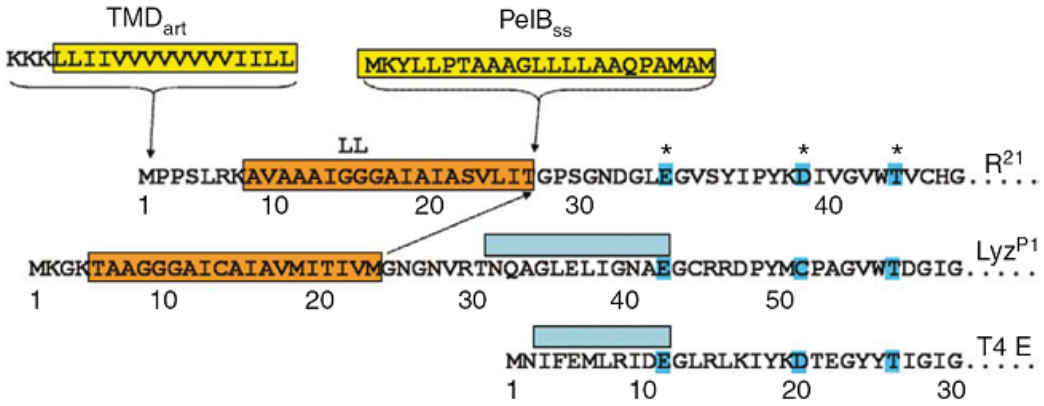

FIG. 3.1. N-terminal sequences of R^{21}, P1 Lyz, and T4 E. The N-terminal domains of R^{21}, P1 Lyz and T4 E are shown aligned by the E-8aa-(D/C)-5aa-T catalytic triad. SAR domains are boxed in orange. Leu substitutions made in R^{21} are shown above Gly14 and Gly15. The PelB signal sequence ($PelB_{ss}$; yellow) and the artificial TMD (TMD_{art}; yellow) are shown with arrows indicating points of fusion in the chimeric constructs. In the N-terminal domain, catalytic triad residues are in blue. Residues involved in positioning the catalytic glutamate are highlighted by a light blue box in P1 Lyz and T4 E.

Further, when Gly14 and Gly15 of the SAR domain were replaced with Leu residues (Fig. 3.1), the R^{21} SAR domain lost its ability to escape the membrane (Fig. 3.2E) and lysis was blocked unless $CHCl_3$ was added to disrupt the bilayer (Fig. 3.2A, F). On the other hand, appending an artificial TMD to the N-terminus of full-length R^{21} resulted in a chimera that was membrane-tethered (Fig. 3.2C) and, although not

explicitly lytic (Fig. 3.2A), had sufficient enzymatic activity to convert all the induced cells to spherical morphology (Fig. 3.2G). Apparently, in this chimera, the location of the R^{21} SAR sequence distal to the TMD prevented its recognition by the *sec* system and caused it to be secreted into the periplasm with the rest of the polypeptide. Moreover, replacing the SAR domain of R^{21} with that of P1 Lyz resulted in a chimera that, although retaining the ability to be released from the membrane (Fig. 3.2D), was inactive (Fig. 3.2A). In contrast, P1 Lyz is still functional when its SAR domain is replaced by that of R^{21} (158). Taken together, these results demonstrate that, besides controlling the topology of protein, the SAR domain in R^{21} plays a specific and more integral role in the catalytic activity of enzyme, compared to that in P1 Lyz.

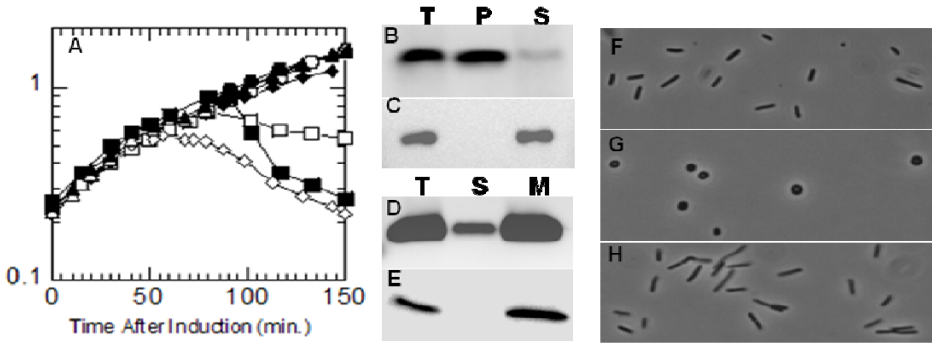

FIG. 3.2. Induction of *R*[21] and derivative alleles. (*A*) Lysis profiles. (○) pZE-luc, (●) pZE-$R^{21}_{G14,15L}$, (□) pZE-luc + 1mM $CHCl_3$ at 80 minutes after induction, (■) pZE-$R^{21}_{G14,15L}$ + 1mM $CHCl_3$ at 80 minutes after induction, (◊) pZE-R^{21}, (♦)P1 $Lyz_{1-26}\Phi R^{21}_{27-165}$, (Δ) pZE-$pelB_{ss}\Phi R^{21}_{27-165}$, (▲) pZE-$TMD_{art}\Phi R^{21}$. (*B-E*) Localization and processing of R^{21} and derivatives after induction of the indicated plasmids. For panels **b** and **c**, lane 1 = total, lane 2 = periplasm, and lane 3 = spheroplast fraction. For panels **d** and **e**, lanes 1, 2, and 3 represent the total, soluble, and membrane fractions, respectively. (**b**) pZEpelBΦR^{21}_{27-165}, (**c**) pZE-$TMD_{art}\Phi R^{21}$, (**d**) P1 $Lyz_{1-26}\Phi R^{21}_{27-165}$, (**e**) pZE-$R^{21}_{G14,15L}$. (*F-H*): Morphologies of cells carrying indicated R^{21} plasmids at 100 minutes after induction. (**f**) pZE-$R^{21}_{G14,15L}$, (**g**) pZE-$TMD_{art}\Phi R^{21}$ (**h**) pZE12-luc.

The structural basis of R^{21} regulation

To explore the structural basis for this novel regulation, the crystal structures of the active, full-length enzyme, $^{a}R^{21}$, and the inactive enzyme, $^{i}R^{21}$, which is missing the entire SAR domain, were solved to high resolution by Quingan Sun in the lab of Dr. Jim Sacchettini at Texas A&M University (Table 3.1). $^{a}R^{21}$ shares the characteristic dumbbell structure of the canonical T4 lysozyme, with the 42 residues distal to the SAR sequence forming a relatively independent catalytic domain, containing a catalytic triad (Glu35, Asp44 and Thr50) connected by a long, domain spanning α-helix (Lys68 to Tyr89) to the bundle of C-terminal α-helices (helix α4 to α8) (Fig. 3.3A, C). Despite

marginal sequence identities (33.9% of R^{21} with P1 Lyz, 30.3% with P22 gp*19*, 18.8% with T4 E), these 3D structure elements are conserved in all of the known structures of T4 E-like lysozymes (R^{21}, P1 Lyz, P22 gp*19*, and T4 E) and the geometry of the catalytic triad in $^{a}R^{21}$ (Glu35, Asp44, and Thr50) is nearly identical to those of the other three (Fig. 3.3B) (10, 96, 158).

In the structures of T4 E (Fig. 3.3C) and P1 Lyz (Fig. 3.3D), an α-helix from the N-terminal of the protein (Ile3 - Glu11 in E; Asn31 – Gly43 in P1 Lyz) interacts extensively with the C-terminal lobe and serves to position the essential glutamate in the active site as part of the catalytic triad. In P1 Lyz, this stabilizing helix is 12 residues in length and separated from the SAR domain by a turn (NVRT) (158). Residues 14-26 of the SAR helix in ^{a}P1 Lyz assume an α-helical conformation that packs lateral to the stabilizing α-helix. However, in $^{a}R^{21}$, the SAR domain is folded into two anti-parallel α-helices, α1 (residues Pro3 to Gly15) and α2 (residues Ala17 to Thr26), connected by a sharp turn consisting of triple Glycines 14-16. These two helices pack against the C-terminal helical bundle (α4 – α8) at an angle of about 45º (Fig. 3.3A). Importantly, inspection of the predicted electrostatic surfaces reveals that the interface between the SAR domain and the body of the enzyme is dominated by hydrophobic contacts. The critical Glu of $^{a}R^{21}$ (Glu35) is located not on the helix itself, but on a loop directly downstream of the SAR helix. This loop is stabilized by a network of hydrogen bonds which includes residues of the SAR domain. The catalytic Glu35 appears to be further stabilized by the salt bridge between its Oε and the side chain of Arg152 from helix α8 of the C-terminal helical bundle, a feature well-known for T4 E (131) and also shared by

the P22 and P1 enzymes. Sequence alignment of R^{21}-like endolysins indicates that this Glu-Arg salt bridge is fully conserved in this family (Fig. 3.4).

Additional differences were observed in the structure of the R^{21} inactive form ($^{i}R^{21}$) compared to the active enzyme (Fig. 3.5). Although overall the $^{i}R^{21}$ crystal diffracted to higher resolution than that of $^{a}R^{21}$ (Table 3.1), the five N-terminal residues preceding Asn31 were not visible in the electron density map. The catalytic Glu35 can be only built to Cβ, presumably due to the flexibility of this region in the inactive form. Whereas in P1 Lyz, the active and inactive forms have radically different structures, the overall fold of $^{i}R^{21}$ is nearly identical to that of $^{a}R^{21}$, with every major secondary structural element preserved, except the changes to the active site, described below. Like P1 Lyz, R^{21} has two disulfide bonds that provide structural stability, but unlike P1 Lyz, there is no difference in the disulfide bonding pattern between the active and inactive forms. In the catalytic domains of $^{a}R^{21}$ and $^{i}R^{21}$, the loop regions between Ser38 and Thr67, which includes two of the residues of the catalytic triad, Asp44 and Thr50, are superimposable (rmsd ~0.361 Å). However, the absence of the two α-helices of the SAR domain and the adjacent H-bond network has a dramatic effect on the position of the Glu35, which is displaced about 10Å. Its position in the active enzyme is occupied by Lys147 in $^{i}R^{21}$, in the loop connecting two of the C-terminal helices, α7 and α8. Lys147 is likely to further ensure that the active site in $^{i}R^{21}$ would not be capable of binding substrate and concomitantly fold into active geometry. Given the indispensable role of the catalytic Glu demonstrated in T4 E (117), the displacement and disorder of Glu35 account for the lack of enzymatic activity in $^{i}R^{21}$.

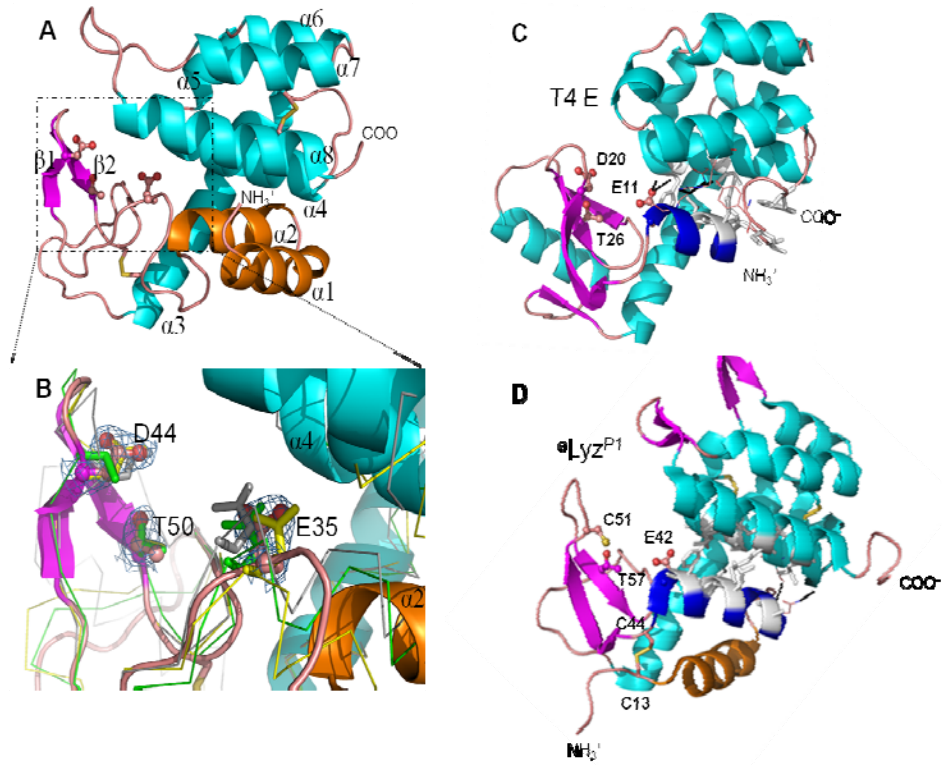

FIG. 3.3. Structural comparison of $^{a}R^{21}$ and T4 E. (*A*) Overall structure of $^{a}R^{21}$. The cartoon is colored according to secondary structure (helix in cyan, sheet in magenta, coil in brown), while the SAR helices (α1 and α2) are in orange. The catalytic triad and the two pairs of disulfide bonds are shown as stick structures (Q. Sun, Texas A&M University). (*B*) The catalytic triad (Glu35, Asp44 and Thr50). The catalytic loop regions of T4 E (Glu11-Thr59, grey C-α trace), P1 Lyz (Glu42-Lys72, green C-α trace) and P22 gp*19* (Glu16-Ile50, yellow C-α trace) were superimposed on $^{a}R^{21}$ (Glu35 - Thr67); the RMSs are 0.66Å, 1.03Å and 0.40 Å, respectively. The catalytic triads of T4E, P1 Lyz, and P22 gp*19* are shown in stick. The 2Fo-Fc electron density map of $^{a}R^{21}$ around the triad is shown at the 1.0 σ level. (*C*) The structure of T4 E. The cartoon is colored as described for Fig 3a. The helix responsible for placement of the catalytic Glu is shown in blue. Catalytic residues are shown in stick. Residues involved in the hydrophobic interaction between the Glu placing helix and the C-terminal bundle are shown in grey stick. (*D*) The structure of aP1 Lyz. The cartoon is colored as described for Fig 3A. As in 3c, the helix responsible for placement of the catalytic Glu is shown in blue with hydrophobic interaction residues in grey stick. Disulfide bonds are represented in yellow stick, and Cys residues participating in the activating disulfide bond isomerization (Cys13 and Cys44) are labeled.

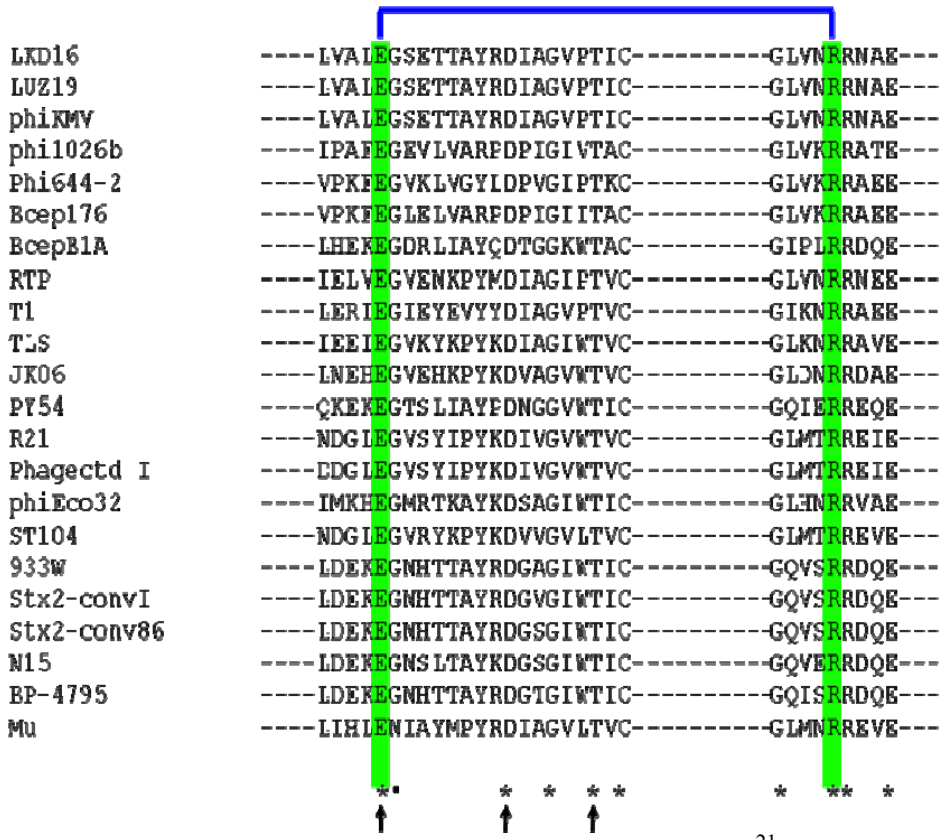

FIG. 3.4. The Glu-Arg salt bridge is fully conserved in the R^{21}-like endolysin family. Salt bridge residues are highlighted in green. Only the salt bridge region is shown. The catalytic glutamate residue forms a salt bridge (blue bracket, top) with an Arg in helix 8. Catalytic (arrow), fully conserved (*), and nearly completely conserved residues (dot) are indicated. The labels of the aligned sequences correspond to the following Genbank records: LKD16: gi|158345072, LUZ19: gi|167600491, phiKMV: gi|33300856, phi1026b: gi|38707914, phi644-2: gi|134288680, Bcep176: gi|77864683, BcepB1A: gi|48697551, RTP: gi|81343992, T1: gi|45686348, TLS: gi|148734541, JK06: gi|71834140, PY54: gi|33770570, R^{21}: gi|215468, Phagectd I: gi|148609440, phiEco32: gi|167583572, ST104: gi|46358689, 933W: gi|9632511, Stx2-convI: gi|20065952, Stx2-conv86: gi|116221999, N15: gi|3192716, BP-4795: gi|157166033, Mu: gi|6010396.

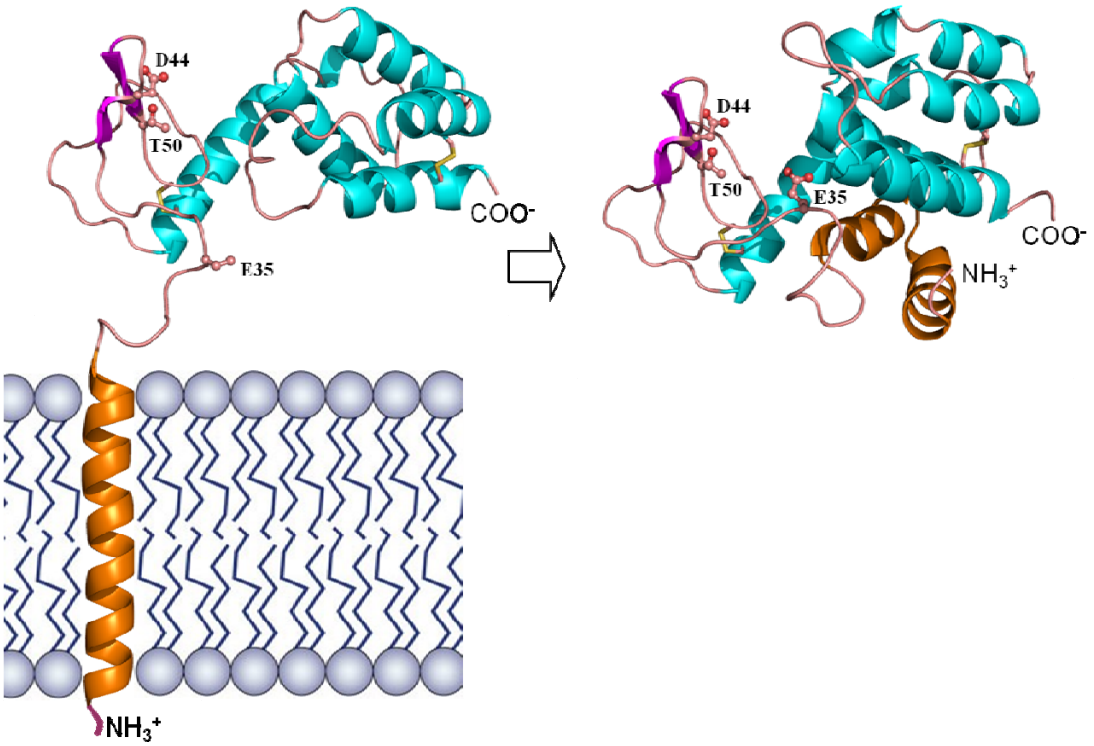

FIG. 3.5. Topological and conformational dynamics of R^{21} activation. Color scheme as in Fig. 3.3. On the left, the SAR domain is shown as a membrane-spanning helix in the inactive, membrane-tethered form.

Discussion

A new mechanism for the regulation of SAR endolysins

Mechanistic insight for the regulation of secretory endolysins became available only recently, for the SAR endolysin of coliphage P1 (158, 159). After *sec*-mediated export, P1 Lyz is tethered to the membrane by its N-terminal SAR domain and maintained in an inactive state by virtue of a disulfide bond between the catalytic Cys51 and a second Cys residue (Cys44) in the catalytic domain, as well as by the complete

disorganization of its N-terminal catalytic domain. The SAR domain escapes from the membrane, at a low constitutive rate or quantitatively and synchronously when the holin triggers to disrupt the membrane. In the active, soluble form of P1 Lyz, a Cys residue in the SAR domain replaces the catalytic Cys in the disulfide bond, and the N-terminal catalytic domain is completely refolded, yielding a properly-arranged catalytic triad straddling the substrate-binding cleft (Fig. 3.3D) (158). In this active structure, the SAR domain makes no significant contacts with the body of the enzyme other than the disulfide linkage. In fact, except for the disulfide linkage, the SAR domain of P1 Lyz is not essential for the function of the endolysin. Even the position of the Cys residue within the SAR domain is very flexible, suggesting that the SAR domain can be arranged in a variety of orientations with respect to the catalytic domain (158). The studies presented here describe an entirely different mode of regulation for the SAR endolysin of phage 21, the prototype of a class of SAR endolysins much more frequently represented in annotated phage genomes than the P1 Lyz class (Table 3.2). Although $^{a}R^{21}$ shares both its overall structure and the geometry of its catalytic triad and substrate-binding cleft with P1 Lyz and with T4 E (Fig. 3.3), there is a fundamental difference. In $^{a}R^{21}$, the SAR domain, when removed from the membrane where it is presumably entirely helical, "jack-knifes" into a helix-turn-helix structure that makes many intimate contacts with the N-terminal catalytic domain and the C-terminal helical bundle and is essential to the activity of the enzyme. Moreover, the overall structure of the N-terminal catalytic domain is much more preserved between the $^{i}R^{21}$and $^{a}R^{21}$ structures, in comparison to the dramatically different folds of ^{i}P1 Lyz and ^{a}P1 Lyz. Only a loop

segment of the catalytic domain adjacent to the SAR sequence is arranged differently, with the Glu residue of the catalytic triad averted from the enzymatic cleft in $^{i}R^{21}$. Another striking contrast is that activation of R^{21}involves significant reorientation and restructuring of the C-terminal helical bundle, resulting in the formation of a hydrophobic binding site for the helix-turn-helix structure of the released SAR domain. In ^{i}P1 Lyz and ^{a}P1 Lyz, the C-terminal domains are very similar (158). Overall, the structural transitions are much less dramatic in R^{21} activation than for P1 Lyz. Thus, paradoxically, after extraction from the membrane the SAR domain of R^{21} enters the structure of the enzyme intimately without causing fundamental restructuring. In contrast, the extracted SAR domain of P1 Lyz provides only its thiol group for covalent activation of the catalytic triad and can do so in a number of orientations, yet activation of P1 Lyz is associated with a complete refolding of the entire N-terminal catalytic domain and *de novo* creation of the substrate-binding cleft. The simplest notion is that the periplasmic domain of ^{i}P1 Lyz is much less stable. Upon insertion into the membrane, P1 Lyz is remodeled by the host DsbA (159), and may be conformationally remodeled by host chaperones after release. For R^{21}, the small but distinct differences between the conformations of the bundled C-terminal helices, especially helices α4 and α8, suggest a possible induced fit mechanism for its activation. Once the SAR domain has escaped from the membrane, the simplest way to minimize the exposure of its most hydrophobic residues is to pack against the non-polar face of helix α4, which could lead to dislodgement of helix α8 and further conformational adjustments. In this perspective, the newly released SAR domain activates R^{21} through the spatially adjacent C-terminal

helical bundle, rather than the sequence-close N-terminal domain.

Comparison of the two strategies

The regulatory strategies adopted by P1 Lyz and R^{21} have significant differences. ^{i}P1 Lyz has two levels of negative control: covalent inactivation of its active site Cys and a N-terminal catalytic domain with radical conformational disability. In contrast, $^{i}R^{21}$ appears to lack only the correct placement of its catalytic Glu, which is sterically blocked from attaining its position by a hydrogen-bonded lysine side-chain. These striking differences suggest evolutionary trade-offs exacted to meet the conflicting dual responsibilities of the SAR endolysin: maintenance of an inactive state throughout the infection cycle and also the ability to effect degradation of the murein as rapidly as possible, once the holin has triggered and depolarized the membrane. Moreover, in contrast to canonical endolysins, which are completely blocked from premature muralytic function by sequestration in the cytoplasmic compartment, SAR endolysins are inherently leaky, in the sense that even without holin function, enzymes are constitutively released from the membrane, thus becoming activated during the latent period (109). P1 Lyz appears to be better suited for negative control in view of the Cys substitution for Asp in the catalytic triad, which has been shown to reduce enzyme activity in the T4 E context (58). R^{21} seems to be more poised for muralytic function, with its canonical catalytic triad and the minimally dysfunctional conformation of the inactive enzyme. The particular evolutionary path taken by these two SAR endolysins would reflect the selective pressures exerted on the lysis process. In particular, analyses

based on the Minimum Value Theorem for predator-prey relationships have indicated that conditions of sparse host availability and/or poor growth conditions favor extended infection cycles and thus favoring an endolysin less prone to premature activation (147). In contrast, target-rich and/or rapid growth environments favor shorter infection cycles and more rapid release of the progeny virions, which would favor the more efficient and muralytic endolysin.

Dynamic membrane topology of the SAR domain

Compared to canonical transmembrane domains, SAR domains are enriched in Gly and Ala residues, a feature that is critical for dynamic membrane topology. In R^{21}, replacement of two Gly residues with Leu eliminates the ability of the SAR endolysin to escape from the membrane (Fig. 3.2E). Nevertheless, SAR domains still have numerous hydrophobic residues which, after extraction from the lipidic milieu, are a liability in the aqueous environment of the periplasm. In R^{21}, the cluster of Gly residues in the SAR domain provides a second structural function, serving as the hinge allowing the extracted SAR domain to fold into two shorter helices. Besides creating a new helix-helix interface to accommodate hydrophobic residues, this also compacts the SAR domain sufficiently that it can fit into a hydrophobic pocket on the surface of the main body of the enzyme. Consequently, nearly all the hydrophobic residues of the SAR domain are buried in the active structure, and the most solvent-exposed surface is populated by hydrophilic residues.

Is R^{21} representative of the Cys-less SAR endolysins in terms of the structure and role of the SAR domain in the active enzyme? A compilation of the N-terminal sequences including the SAR domains and the catalytic triads of 22 Cys-less SAR endolysins suggests that there may be several different modes (Fig. 3.6), since the SAR domain varies significantly in its proximity to the active site Glu. R^{21} shares similar SAR domain arrangements with endolysins from phages encoding potent human toxins, including the cytolethal distending toxin (6), CDT, and the Shiga toxin (113); moreover, in the latter case, the acute dispersal of the toxin in the mammalian gut is dependent on the muralytic action of the SAR endolysin (101, 103). In no case does the hydrophobic core of the SAR domain extend farther than 3 - 4 helical turns without encountering a Gly or Pro residue. These residues are tolerated in helices in the membrane but are helix-breakers in aqueous solvent (77), suggesting that in most or all of these proteins, the SAR domain uses some version of the R^{21} strategy of converting the SAR sequence into a helix-turn-helix packed against the main body of the enzyme to solve the problem of what to do with a transmembrane domain out of its native lipidic environment. In some homologs, the SAR domain immediately abuts the catalytic Glu residue; in these enzymes, most of the SAR domain must constitute the stabilizing helix. Taken together with the structures described here, it seems clear that SAR domains represent structural elements of unparalleled evolutionary flexibility, both in terms of membrane topology and the capacity for integration as a regulatory component.

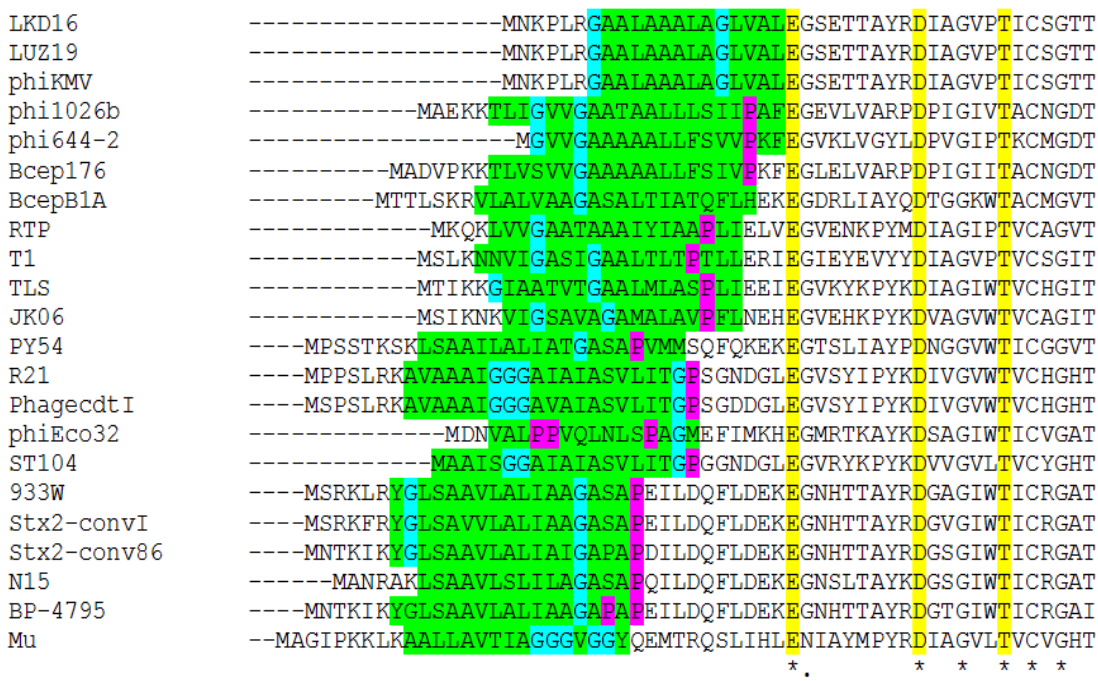

FIG. 3.6. Alignments of unique SAR endolysins lacking Cys residues in their SAR domains. The N-terminal SAR and catalytic domains of R^{21}-like SAR endolysins are shown aligned by the catalytic triad (yellow). These sequences were selected from the 34 sequences identified in Table 3.2 by the criterion of having at least one amino acid residue difference in the complete sequence. Completely conserved (asterisk) and nearly completely conserved (dot) residues are indicated. SAR domains are indicated in green, along with helix-breaking residues Gly (blue) and Pro (purple). The Genbank records correspond to those used in figure 3.4.

CHAPTER IV

THE PREVALENCE OF SAR ENDOLYSINS AMONG dsDNA PHAGES OF GRAM-NEGATIVE HOSTS

Introduction

The classic lysis paradigm for the lysis of Gram-negative hosts by dsDNA phages is typified by phage lambda. It consists of three components: a cytoplasmic endolysin; an inner membrane (IM) channel-forming protein, the holin; and the spanin complex, consisting of an integral IM protein and an outer membrane (OM) lipoprotein (163). A three-step model has been proposed for the coordinated lytic function of these proteins. At the end of the vegetative cycle, the holin forms large, non-specific lesions in the IM that allows passage of the muralytic endolysin. Consequently, the peptidoglycan (PG) is compromised, triggering the action of the spanin complex. Endolysins in this paradigm include homologs of phage T7 gp*3.5* (an amidase), lambda R (a transglycosylase), T4 E (a lysozyme), and T5 Lys (a peptidase). The lysozymes and transglycosylases belong to the lysozyme-like superfamily and share a common fold and the amidases belong to the peptidoglycan recognition protein (PGRP) superfamily. Amidases cleave the amide bond between MurNAc and L-alanine, transglycosylases and lysozymes cleave the β-1,4 glycosidic linkages between MurNAc and GlcNAc moieties, and peptidases cleave the peptide bonds between amino acids of the tetrapeptide crosslinks of the PG (Fig. 1.2).

Recent work has uncovered a new lysis paradigm. Rather than a cytoplasmic endolysin, this new paradigm depends on a secreted SAR (Signal Anchor Release)

endolysin (159). A SAR endolysin contains an N-terminal transmembrane domain (TMD), the SAR domain, which, compared to normal TMDs, is richer in weakly hydrophobic or polar residues such as Ala, Gly, and Ser. SAR endolysins are secreted by the SecYEG translocon and accumulate in the periplasm tethered to the energized IM . The SAR domain is so named because it first acts as a membrane anchor but has the capacity to escape from the membrane. The release of the SAR domain occurs spontaneously at a low rate but occurs quantitatively when the holin triggers to depolarize the membrane. Some SAR endolysins are coupled with pinholins, a class of holins that form small lesions (109). The pinholin depletes the proton motive force (PMF) of the IM, causing the immediate release of the tethered SAR endolysin. Similar to the lambda paradigm, this system also contains a spanin complex for OM disruption. To date, several SAR endolysins including P1 Lyz, R^{21}, Lyz^{103}, and N4 gp*61* have been experimentally characterized (74, 130, 133, 158).

SAR endolysins are expressed from the beginning of the late gene expression period, along with all the morphogenesis proteins. To prevent premature lysis of the host cell, it is imperative to block the PG degrading activity (133). The first SAR endolysin to be genetically and structurally characterized was P1 Lyz from phage P1 (158). It is a homolog of T4 E and contains an E-8x-C-5x-T catalytic triad. In its membrane-tethered form, the protein is maintained in a covalently inactivated state, in that the catalytic Cys in position 51 is held in a disulfide bond. This inhibition is relieved when the SAR domain is released from the membrane. A sulfhydryl in the SAR domain attacks the inactivating disulfide, releasing Cys_{51}, freeing the catalytic thiol and allowing refolding

of the catalytic domain (158). Lyz^{103}, the SAR endolysin of *Erwinia amylovora* phage ERA103, is also regulated by a disulfide bond isomerization. Rather than using an inactivating disulfide, Lyz^{103} is regulated by means of disulfide caging of a catalytic residue (74). Most SAR endolysins, however, do not contain Cys residues in their SAR domains, and thus cannot be regulated by disulfide bond isomerization (133). These proteins are also T4 homologs and have the canonical E-8x-D-5x-T catalytic triad. The prototype of this class of SAR endolysins is R^{21} from lambdoid phage 21. The inactive form of R^{21} was shown to have a misfolded catalytic cleft, mainly in that the catalytic Glu residue is displaced by 10 angstroms. Upon release from the membrane, the SAR domain associates with the C-terminal domain to position the catalytic Glu, thus forming the active site. Bioinformatic analysis suggests that this latter regulatory mechanism is most predominant among SAR endolysins, based on the sequenced phage genomes.

Here we undertake a comprehensive bioinformatic search for endolysins of phages of Gram-negative hosts. The goals of this study are to determine the extent of endolysin variety in regards to classes/mechanisms, to discern any new lysis paradigms, and to determine the prevalence of the SAR endolysin paradigm among phages of Gram-negative hosts. The results are discussed in terms of evolution of phage lysis and of the topologically dynamic SAR endolysin.

Materials and methods

Bioinformatic endolysin identification and phyolgenetics

Our search for endolysin genes encompassed the annotated dsDNA phages of Gram-negative bacteria listed in the NCBI Entrez Genome and Nucleotide databases to date. Sequence similarity searches were carried out against the non-redundant protein sequences database using NCBI BLASTp (3). Initial rounds of identification were performed using a single query sequence: T7 gp*3.5*, T4 E, lambda R, and T5 Lys representing the amidase, glycosidases, transglycosylase, and endopeptidase classes, respectively. Subsequent rounds of identification were performed using the lowest scoring (*E*-value), non-redundant sequences from these searches. This method was repeated until no additional endolysin hits could be identified. Endolysins not identified using this method were found by hand either by verifying proteins annotated by the submitters as endolysins or by manual BLASTp and PSI-BLAST searches. Manual searches utilized the NCBI Conserved Domain Database (CDD). Conserved domain searches were performed for protein coding sequences between 100 and 200 amino acids proximal to the phage structural genes. Any remaining phage genomes were submitted to EMBL-EBI InterProScan (114) as FASTA files for functional domain prediction. Phylogenetic tree analysis was performed using Phylogeny.fr, a free online suite that performs multiple sequence alignments using MUSCLE, phylogenetic tree construction with PhyML, and tree rendering with TreeDyn (35). Unrooted trees are presented due to

the difficulty in selecting an outgroup among organisms that are known to transfer genes among each other frequently (36).

Bacterial strains and growth conditions

The *Escherichia coli* strains BL21 (DE3) pLysS, XL1-Blue, MC4100, and MG1655 Δ*tonA lacY*⁻ *lacI*^{q} have been described (127, 133). Standard conditions for the growth of cultures and the monitoring of lysis kinetics have been previously described (29, 127). All bacterial cultures were grown in standard LB medium, supplemented with 100 μg/ml ampicillin when appropriate. When indicated, isopropyl β-D-thiogalactopyranoside (IPTG) and dinitrophenol (DNP) were added to final concentrations of 1 mM and 10 mM, respectively.

DNA procedures and plasmid construction

Procedures for the isolation of plasmid DNA, DNA amplification by PCR, PCR product purification, DNA transformation, site-directed mutagenesis, and DNA sequencing have been previously described (52, 126, 128). The plasmid pZEBcep22gp*79*, a derivative of the medium copy, IPTG-inducible vector pZE12 (86), was constructed by PCR amplifying Bcep22 *gp79* from the genome and inserting it between the *Kpn*I and *Xba*I sites of pZE12 digested with the same enzymes. Derivative alleles were constructed using site-directed mutagenesis. The construct pETΦV10gp*29* was constructed by PCR amplifying ΦV10 *gp29* from the genome (kindly provided by Bruce Applegate) and inserting it between the *Nde*I and *Bam*HI sites of pET11a. For

detection and purification purposes, Bcep22gp*79* and ΦV10gp*29* were modified by ExSite PCR to encode an oligohistidine tag ($Gly_2His_6Gly_2$) appended to Gln_{174} and Val_{209}, respectively. All purified protein cited in this work refers to the oligohistidine-tagged versions.

Subcellular fractionation

Soluble or membrane localization was determined as described previously (133). Briefly, 25 mls of an induced culture were collected by centrifugation at 5,000 x g in a Sorvall Superspeed RC2-B centrifuge, and resuspended in 2mls of French press buffer (0.1M sodium phosphate, 0.1M KCl, 5 mM EDTA, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, pH 7.0). Cells were disrupted by passage through a French pressure cell (Spectronic Instruments, Rochester, N.Y.) at 16,000 lb/in^2. The membrane and soluble fractions were separated by centrifugation at 100,000 x g in a Beckman TL-100 Ultracentrifuge for 60 min. To separate the periplasmic fraction, cells expressing pZER[21] derivatives were collected after 30 min of induction by centrifugation at 5000 x g. Cells were resuspended in 0.5 ml of spheroplasting buffer (25% sucrose, 30 mM TrisHCl, pH 8). Ten µl of 0.25 M EDTA, 10 µl of lysozyme (20 mg/ml in water), and 500 µl of distilled water were added sequentially (24). The mixture was inverted several times and incubated for 5 minutes at room temperature. Examination under the microscope showed that ~ 95% of the cells were spheroplasts. Spheroplasts were pelleted by centrifugation at 8,000 x g for 30 min at 4 °C. Equivalent amounts of each fraction were examined by SDS-PAGE and Western blotting as described below.

SDS-PAGE and Western blotting

SDS-PAGE, Western blotting, and immunodetection experiments were performed as previously described (52). A mouse monoclonal antibody against the oligo-histidine epitope tag was purchased from Amersham and was used at a dilution of 1:3000. The anti-mouse IgG horseradish peroxidase-conjugated secondary antibody was supplied with the SuperSignal chemiluminescence kit (Pierce) and was used at a 1:5000 dilution. Blots were developed using the West Femto SuperSignal chemiluminescence kit (Pierce). Chemiluminescent signal was detected using a Bio-Rad ChemiDoc XRS.

Protein expression and purification

pETphiV10gp29cHis was transformed into BL21 (DE3) cells (Invitrogen) harboring pLysS, and fresh transformants were cultured and induced for 3 hrs at 30°C. Cells were collected at 4K RPM for 30 min at 4°C in a Sorvall RC-3B centrifuge and resuspended in purification buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8). Protease Inhibitor Cocktail for His-tagged protein (Sigma) was added as per manufacturer's instructions and cells were lysed by passage through a French pressure cell (Spectronic Instruments, Rochester, N.Y.) at 20,000 lb/in^2. After removing unlysed cells and debris, the lysate was filtered through a 0.2 µm syringe filter. The cleared lysate was then applied to Talon Metal Affinity Resin (Clontech). Protein was eluted in elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 5).

Lysozyme activity assay

Peptidoglycan degradation activity was determined by turbidometry using resuspended *Micrococcus lysodeikticus* cells. Briefly, various quantities of the specified purified protein were added to micro-plate wells. Lyophilized *M. lysodeikticus* cells were resuspended in reaction buffer (50 mM sodium acetate, pH 6.5) at a concentration of 1mg cells/mL. 100 μL cells were added to each well and the optical density was read at A_{550} in a Tecan Infinite M200 Pro in 2 min time points for 1 hr.

Results

Identification of phage endolysin genes

A bioinformatic search for endolysins from dsDNA phage infecting Gram-negative hosts was undertaken. 279 Genbank dsDNA phage entries were searched, and their respective endolysins were identified (Table 4.1). Phages of archaea and cyanobacteria were excluded from the search as these phages have lysis systems that are beyond the scope of this review. Our treatment also excluded Gram-positive phage endolysins, since they tend to be modular in architecture and can have more than one catalytic domain (163). Our search was limited to sequenced phage genomes as these were acquired by DNA isolated from phage particles released presumably by functional endolysins, hence prophage were excluded.

The identified endolysins fall into four superfamilies (Table 4.1). The superfamilies group based on structural homology, i.e. members of the lysozyme-like

superfamily share a similar "classic lysozyme" fold. The superfamilies differ in PG bond specificity: lysozyme-like superfamily members cleave the β-1,4-glycosidic bond, amidase superfamily members cleave the amide bond between MurNAc and L-Ala, and members of the Hedgehog/DD-peptidase superfamily cleave the peptide bonds between peptide crosslinks. The lysozyme-like superfamily is further broken down into sub-classes based on catalytic mechanism, i.e. lysozymes, chitinases, transglycosylases, etc. The muraminidase mechanism has been inferred by catalytic residues as described below. One class of endolysins could only be identified by its PG binding domain. These proteins compose their own PG_binding_1 superfamily. Few phages contained no identifiable endolysin candidate.

Lysozymes

The members of the most dominant endolysin class (32% of total) are homologs of the muraminidase T4 E, a true lysozyme (73). In the past, the term lysozyme has been used loosely to describe PG degrading enzymes (47). For example, T7 gp*3.5* (30) , lambda R (14), goose egg white lysozyme (GEWL) (140), and the endolysin of phage T5 (34) have been described as lysozymes despite their more accurate classifications of gp*3.5* as an amidase, R and GEWL as a transglycosylases, and the T5 endolysin as a peptidase. Here, we will use the term as it applies to T4 E, the original lysozyme and its homologs.

TABLE 4.1. Endolysin candidates from dsDNA phages of Gram-negative hosts.

Superfamily[A]	Endolysin Sub-class[B]	Percent of Total[C]	Members[D]
Lysozyme-like Superfamily	Lysozyme	32	**21**, 25, 31, 65, **2851**,44RR2.8t, **933W**, Acj61, Aeh1, APSE-1, AR1, **Bcep176**, **BcepB1A**, CC31, Cd, **cdtI**, **D108**, **ERA103**, ES18, **F108**, Felix01, **Fels-2**, **HP1**, **HP2,** IME08, IME09, **JK06**, JS10, JS98, **K139**, **kappa**, **KP34**, **LKA1**, **LKD16**, **LUZ19, Min27**, **Mu**, **N15**, **P1**, P7, P22, PA11, PAJU2, PaP3, **phi1026b, phi-2, phi4795**, **phi644-2**, phiAS4, phiAS5, **phiEA100**, phiEA104, **phiEA1H**, phiEa21-4, phiEco32, phiHISC, **phikF77**, **phiKMV**, **phiMHaA1**, **phi-O153**, **phiPLPE**, phiSG1, phiSG3, **PT2**, **PT5**, PX29, **PY54**, RB14, RB32, RB51, RB69, RE-2010, **RTP**, **SE1**, SP18, **ST104**, **Stx-conv I**, **Stx2-conv 171**, **Stx2-conv 86**, **Stx2-conv I**, **Stx2-conv II**, **T1**, T4, T4T, **TLS**, vB_EcoM-VR7, **VT2-Sakai**, WV8, **Xfas53**, **YYZ-2008**
	Transglycosylase[E]	13	186, *B3*, BcepC6B, ***BcepMu***, BcepNazgul, c341, ***CP220, Cpt10***, D3, DE3, E1, Eco1230-10, *EL*, epsilon34, g341c**,** HK022, HK620, HK97, ***KP15***, ***KS10***, L-413C, lambda, *M6*, P2, ***phiE255***, *phiKZ*, phiO18P, phiSboM-AG3, PsP3, RM378, rv5, *SC1*, *SC2*, Sf6, SfV, WPhi, ***YuA***
	Glycoside hydrolase family 19 chitinase	8	11b, 14-1, 16-3, Aaphi23, AB1, BcepF1, epsilon15, F8, F10, Fels-1, JG024, LBL3, LMA2, OP2, PB1, phiAB1, phiV10, phiW-14, RSL1,S1249, SN, ST64B
	Glycoside hydrolase family 24	8	133, **Bcep22**, **BcepIL02**, **F116**, Gifsy-1, Gifsy-2 K1G, K1H, K1ind1, K1ind2, K1ind3, **KPP10**, KS7, **LIMElight**, LUZ24, **OP1**, **phiL7**, **RSB1**, SETP3, So-1, SS3e, SSL-2009a, **Xop411**, **Xp10**
	DUF 847 superfamily	3	BIP-1, BMP-1, BPP-1, **N4** (enterobacteria), phiHAP-1, VHML, VP58.5, VP882

TABLE 4.1. Continued

Superfamily[A]	Endolysin Sub-class[B]	Percent of Total[C]	Members[D]
N-acetyl muramoyl-L-alanine amidase-like superfamily	Amidase	12	13a, 285P, Ac42, Acj9, BA3, BA14, Berlin, DSS3phi2, EcoDS1, EE36phi1, gh-1, ICP3, ICP3_2007_A, ICP3_2008_A, ICP3_2009_B, K11, K1F, KP32, Kvp1, MmP1, VchoN4, PBC5, phi15, phiA1122, phiIBB-PF7A, phiSG-JL2, phiYeO3-12, SIO1, T3, T7, Vi06, VP4, Yepe2, Yep-phi
Hedgehog/DD-peptidase superfamily	Carboxypeptidase	11	119X, B40-8, EcP1, EPS7, H105/1, IN93, JSE, K1-5, K1E, KS9, KVP40, Mx8, PA73, PaP2, Phi1, phiEcom-GJ, phiJL001, phiKO2, phiP27, RB16, RB43, RB49, SP6, SPC35, ST160, ST64T, T5, VP2, VP5, VP93, Xp15
PG_binding_1 superfamily	Pfam01471	8	203phi2-1, Bcep1, Bcep43, Bcep781, BcepNY3, KL3, KS14, KS5, L17, PAK_P1, phi52237, phiCTX, phiE12-2, phiE202, phiRSA1, PR4, PR5, PR772, PRD1, S1, Vi01
n/a	No candidate	5	BcepGomr, D3112, DMS3, LIT1, LUZ7, MP22, MP29, MP38, P4, PA1/KOR/2010, PM2, VpV262

SAR endolysins are in **bold**.

[A] Endolysin superfamily determined by BLAST sequence homology.

[B] Endolysin type determined by conserved domains and homologs found during BLAST searches using experimentally verified proteins of each type.

[C] Percent of 279 total genomes analyzed

[D] Phage names are as listed in the Viral Genome Resource at NCBI [http://www.ncbi.nlm.nih.gov/genomes/VIRUSES/35237.html]. All phages are Caudovirales except for phages L17, PR3, PR4, PR5, PR772, and PRD1, which are members of Tectiviridae.

[E] Phages in *italics* are members of the GEWL family of transglycosylases (see text).

In general, muraminidases cleave the β-1,4 linkages between MurNAc and GlcNAc residues. Members of the lysozyme class were characterized by their E-5x-D/C-8x-T catalytic triad, where the center position can be either an Asp or a Cys with little penalty to activity (58). Lysozymes are bi-lobal with a catalytic N-terminal domain and an α-helical C-terminal domain connected by a long inter-domain helix. Upon further inspection of the T4 homologs, it was found that a majority of them (52 of 90 or 57% of lysozymes in the survey) contain an N-terminal hydrophobic extension that is not present in E (Table 4.2). This extension is characterized as being 16-22 amino acids in length and has a high proportion of weakly hydrophobic and non-charged polar residues such as Ala, Gly, Ser, and Thr. The sequences were analyzed for transmembrane propensity using TMHMM. If the above criteria were met, the corresponding stretch of amino acids was designated as a SAR domain and the endolysins were classified as SAR endolysins. The putative SAR endolysins were further classified based on the presence or absence of a Cys in the SAR domain. Those that contain a SAR domain Cys were presumed to be disulfide bond regulated (DSB) and those without were designated as R^{21}-like.

Glycoside hydrolase family 24

Twenty-three out of 279 phage endolysins were identified as belonging to the endolysin/autolysin super-family. The endolysin/autolysin enzymes are predicted to have a general muraminidase function. Members of this class include a variety of PG degrading enzymes such as T4 E, lambda R, chitinases, and bacterial autolysins involved

in cell wall turn-over. To gather more information regarding possible mechanism, members of our endolysin/autolysin group were queried in CAZY, the Carbohydrate Active enZYmes database. All of the members of our endolysin/autolysin group belong to the glycoside hydrolase family 24 (GH24). GH24 proteins are predicted by UniProt to use a catalytic Glu to protonate the glycosidic bond in a hydrolase mechanism similar to lysozymes. Inspection of the GH24 sequences showed no canonical lysozyme-like catalytic triad, so these proteins were not classified as true T4 E homologs. In our search, almost half (43%) of these enzymes contained putative SAR domains and were classified as "GH24" SAR endolysins.

Transglycosylases

The second largest category of endolysins (13%) identified were the lytic transglycosylases (LTs). Like T4 E, LTs are muraminidases as they cleave β - 1,4 linkages between MurNAc and GlcNAc residues, however, they use a different mechanism. The LTs identified fell into two sub-categories: those related to R, the LT of phage lambda (55% of LTs) and those that are related to goose egg white lysozyme (GEWL) and bacterial soluble lytic transglycosylases (SLT) (45% of LTs). Goose type and phage LTs represent families 1 and 4, respectively, according to a transglycosylase compilation which identified four families based on sequence similarity and conserved domains (15).

TABLE 4.2. Putative SAR endolysins.

Phage	Phage accession[A]	Host	Endolysin entry	Type[B]
21	215466	*Enterobacteria*	67436	R^{21} like
2851	209407371	*Enterobacteria*	209407416	R^{21} like
933W	9632466	*Enterobacteria*	9632511	R^{21} like
Bcep176	77864625	*Burkholderia*	77864683	R^{21} like
Bcep22	158997720	*Burkholderia*	158997736	GH24
BcepB1A	68165913	*Burkholderia*	48697551	R^{21} like
BcepIL02	238801616	*Burkholderia*	238801690	GH24
BcepMu	48696910	*Burkholderia*	48696932	LT
cdtI	148609382	*Escherichia coli*	148609440	R^{21} like
CP220	294337971	*Campylobacter*	294338091	LT
CPt10	294338166	*Campylobacter*	294338289	LT
D108	281199644	*Escherichia coli*	281199665	R^{21} like
Era103	125999960	*Erwinia*	126000009	DSB
F108	73918060	*Pasteurella*	109302925	DSB
F116	48527494	*Pseudomonas*	56692926	GH24
Fels-2	169936017	*Enterobacteria*	169936041	DSB
HP1	9628600	*Haemophilus*	9628630	DSB
HP2	13752188	*Haemophilus*	17981840	DSB
JK06	71834082	*Enterobacteria*	71834140	R^{21} like
K139	17975105	*Vibrio*	17975138	DSB
kappa	165292211	*Vibrio*	165970268	DSB
KP15	294661423	*Klebsiella*	294661543	LT
KP34	294661411	*Klebsiella*	282554633	R^{21} like
KPP10	297591648	*Pseudomonas*	297591686	GH24
KS10	197725423	*Burkholderia*	198449271	LT
LIMElight	308071837	*Pantoea*	308071892	GH24
LKA1	114796430	*Pseudomonas*	158345187	R^{21} like
LKD16	114796377	*Pseudomonas*	158345072	R^{21} like
LUZ19	161168305	*Pseudomonas*	167600491	R^{21} like
M6	89155211	*Pseudomonas*	149408252	LT
Min27	163955703	*Enterobacteria*	170783656	R^{21} like
Mu	9633494	*Enterobacteria*	6010396	R^{21} like
N15	9630464	*Enterobacteria*	3192716	R^{21} like
N4	117650898	*Enterobacteria*	119952230	DUF847
OP1	84662593	*Xanthomonas*	84570669	GH24
P1	46401626	*Enterobacteria*	33338666	DSB
P7	33323506	*Enterobacteria*	33323523	DSB
phi1026b	38707890	*Burkholderia*	38707914	R^{21} like
phi-2	281306659	*Pseudomonas*	281306699	R^{21} like
phi4795	31044225	*enterobacteria*	31044270	R^{21} like

TABLE 4.2. Continued

Phage	Phage accession[A]	Host	Endolysin entry	Type[B]
phi644-2	148734199	*Burkholderia*	134288680	R^{21} like
phiE255	134288788	*Burkholderia*	134288836	LT
phiEA100	311875322	*Erwinia*	311875369	DSB
phiEa1H	311875202	*Erwinia*	311875248	DSB
phikF77	225626323	*Pseudomonas*	225626372	R^{21} like
phiKMV	33300811	*Pseudomonas*	33300856	R^{21} like
phiL7	238695587	*Xanthomonas*	238695615	GH24
phiMHaA1	90110542	*Mannheimia*	109289945	DSB
phi-O153	56682761	*Escherichia coli*	32170989	R^{21} like
phiPLPE	195964712	*Iodobacteriophage*	197085629	R^{21} like
PT2	165880710	*Pseudomonas*	195546752	R^{21} like
PT5	158187602	*Pseudomonas*	195546690	R^{21} like
PY54	33770508	*Yersinia*	33770570	R^{21} like
RE-2010	312601701	*Salmonella*	312601724	DSB
RSB1	197935896	*Ralstonia*	197935896	GH24
RTP	81343928	*Enterobacteria*	81343992	R^{21} like
SE1	219681191	*Salmonella*	219681236	R^{21} like
ST104	46358648	*Enterobacteria*	46358689	R^{21} like
Stx1-conv I	32170834	*Escherichia coli*	32170989	R^{21} like
Stx2-conv 86	115500802	*Escherichia coli*	116221999	R^{21} like
Stx2-conv 1717	209447126	*Escherichia coli*	209447172	R^{21} like
Stx2-conv I	20065797	*Escherichia coli*	20065952	R^{21} like
Stx2-conv II	32171002	*Escherichia coli*	32171161	R^{21} like
T1	45686283	*Enterobacteria*	45686348	R^{21} like
TLS	38046731	*Enterobacteria*	148734541	R^{21} like
VT2-Sakai	9633396	*Enterobacteria*	9633441	R^{21} like
Xfas53	273810419	*Xylella*	273810445	R^{21} like
Xop411	148727181	*Xanthomonas*	148727205	GH24
Xp10	32128412	*Xanthomonas*	32128440	GH24
YuA	162135082	*Pseudomonas*	162135144	LT
YYZ-2008	209427726	*Enterobacteria*	209427769	R^{21} like

SAR endolysins were identified in the genomes of 71 phages of Gram-negative hosts

[A] The GenBank identifier numbers for the phage and endolysin are given.

[B] The type of SAR endolysin was determined by manual inspection. T4 E homologs with no Cys in the SAR domain were designated as R21 like. The E homologs with a Cys in the SAR domain were presumed to be regulated by disulfide bond isomerization (DSB). Members of the glycoside hydrolase family 24 were deemed 'GH24' SAR endolysins. Candidates that are GEWL homologs are designated as SAR containing lytic transglycosylases (LT). The SAR endolysin of phage N4 is categorized as a DUF847 protein (130).

Three of the families identified in that study were bacterial and the fourth contained transglycosylases similar to lambda R. Interestingly, the bacterial SLT/GEWL proteins have a lysozyme fold (described above) (139). In our study, members of these categories were classified based on conserved domains as described by Blackburn and Clarke (15). Of the two families identified here, only the family 1 LTs (SLT/GEWL homologs) were found to have putative SAR domains and were classified as "LT" SAR endolysins (Table 4.2).

Chitinases

Twenty-two phages were found to encode an endolysin containing a COG3179 domain that has the predicted function of a chitinase. More specifically, members of this class are members of the glycoside hydrolase family 19. Chitin is composed of repeating units of GlcNAc, which is the main component in fungal cell walls and is also found in the exoskeletons of insects, crustaceans, and other invertebrates. Chitinases hydrolyze the β-1,4-glycosidic linkages between GlcNAc residues using a single-replacement reaction that results in inversion of the PG sugar configuration at C_1 (55). Chitinases belonging to the family 19 have two catalytic Glu residues positioned eight residues apart. These two catalytic residues and their spacing are conserved among the putative members of this family found in our search. Divergence from a common ancestor has been suggested between chitinases and lysozymes based on structural similarities (95). Despite the lack of sequence similarity, both enzymes contain N and C-terminal lobes, are highly alpha helical, and have a large catalytic cleft.

DUF847 superfamily

Eight phages were found to have endolysins belonging to the DUF847 superfamily. A notable entry of this type is gp*61* from the enterobacteria phage N4. Gp*61* has been characterized as an N-acetylmuramidase by mass spectrometric detection of reduced muropeptide fragments after hydrolysis (130). Although it is not a true T4 E homolog, it is presumed to use a hydrolase mechanism (described in detail below). Interestingly, when gp*61* was cloned and expressed in *Escherichia coli* cells, rapid lysis occurred independently of a holin. Fractionation showed that the protein was found in the IM and the periplasm at the same apparent molecular mass. These characteristics suggest that N4 gp*61* is a SAR endolysin. Other members of this family contain a catalytic Glu (E_{26} in gp*61*), which is shared with other N-acetylmuramidases like T4 E, but do not have a predicted SAR domain.

Amidases

Thirty-four out of 279 phages encoded putative amidases. Mutational analysis of the T7 amidase, gp*3.5*, has shown that residues Tyr_{46} and Lys_{128} are responsible for amidase activity and residues His_{17}, His_{122}, and Cys_{130} are important for Zn^{2+} binding, required for amidase function (30). All of the putative amidases identified in this search contained the Tyr and Lys catalytic residue equivalents and a predicted metal binding site. Twenty-eight of the 34 amidase containing phages are T7-like podophages. *Silicibacter* phage DSS3phi2 and *Sulfitobacter* phage EE36phi1 are N4-like podophages. Interestingly, the *Acinetobacter* phages Ac42 and Acj9 are T4-like myophages.

Peptidases

Eleven percent of the endolysins were identified as peptidases. These are usually metallopeptidases that use zinc or some other divalent cation as a cofactor. Peptidases cleave the peptide bonds between the amino acids of the PG tetrapeptide crosslinks. The lytic peptidases similar to the phage T5 peptidase are specific for the L-alanoyl-D-glutamate peptide bond. Members of the peptidase class were identified by the canonical catalytic motif of H-6x-D-Xx-D-2x-H where the H-6x-D motif represents the metal binding site and the final D-2x-H (separated from the metal binding site by a variable amino acid region) is the active site (19). The peptidase containing phages are a mix of sipho, myo, and podophages.

PG binding domains

Eight percent of the endolysins were identified by their peptidoglycan binding domain (pfam01471). This domain contains a helix bundle that is found in several PG degrading enzymes including the endolysins from *Pseudomonas aeruginosa* bacteriophages ΦKZ and EL (23). The proteins in this category have the DGX-Pho-GXXT consensus sequence, where Pho represents a hydrophobic amino acid (49). The tectiviruses (L17, PR4, PR5, PRD1, and PR772) and *pseudomonas* phage PAK_P1 contain DUF3380 domains that are usually associated with PG binding domains. Other than that association, no additional information is available for these domains. Sequence analysis by PSI-iBLAST provides no additional information regarding possible mechanism or catalytic residues of the proteins in this category.

No candidate

Five percent of the phages searched contained no bioinformatically identifiable endolysin candidate. Phages in this category were manually inspected for gene annotation and ORFs were checked for homology to known lysis proteins. Furthermore, phage genomes were submitted to InterPro Scan analysis (164). No identifiable endolysin hits were recovered. Of the 12 phages with no identifiable candidate, 8 are *Pseudomonas* phages. The genomes of phage with no endolysin candidate were acquired by sequencing DNA from phage preps indicated that phage particles had been released in some way (4, 27, 56, 57, 60, 88, 165). Interestingly, some of the phages (BcepGomr, D3112, and DMS3) have identifiable holin and spanin candidates, but no muralytic enzyme. Perhaps these phages contain a novel endolysin class that has yet to be characterized.

Distribution of SAR endolysins

Overall, the SAR endolysins were found to compose 25% of the endolysins identified. Within the SAR endolysins, 75% are homologous to T4 gp*e* and are true lysozymes. Fourteen percent are GH24 proteins and 11 percent are transglycosylases. The SAR endolysin of *enterobacteria* phage N4 belongs to the DUF847 family (130). Interestingly, these four protein types are all members of the Lysozyme-like superfamily. They share motifs of the classic lysozyme fold including a catalytic Glu and cleave the β-1,4-glycosylidic linkage between MurNAc and GlcNAc moieties. For the most part,

their amino acid sequences are unrelated, but they are hypothesized to share a common ancestor (62, 75, 95).

Bcep22 gp79 is a SAR endolysin, experimental confirmation

Of the 23 proteins in the GH24 category, 10 contained putative SAR domains. The putative SAR endolysin from *Burkholderia* phage Bcep22 was selected as a representative of this group for further analysis. Since a SAR endolysin can spontaneously release from the membrane, it can lyse cells independently of holin expression. Accordingly, when pZEgp*79* was expressed, overt culture lysis occurred (Fig. 4.1A). In the context of a phage infection, however, SAR endolysins are released synchronously by the depolarizing action of the holin. Energy poisons such as dinitrophenol (DNP) mimic the effect of the holin and cause premature release from the membrane and thus early lysis, as is the case here (Fig. 4.1A). A subcellular fractionation revealed that the protein was present in both the soluble and membrane fractions of the cell. The apparent molecular mass of the released and bound protein was the same indicating that the latter was not a proteolysed version of the former (Fig 4.1B). As with previously studied SAR endolysins (74, 133), the protein could be restricted to the membrane by replacement of adjacent Gly or Ala residues with Leu residues, with either two or three substitutions required, depending on the SAR endolysin. In this case, the replacement of two Ala residues converted gp*79* to a membrane-bound protein and also eliminated its lytic function (Fig 4.1A, C). SAR endolysins are secreted through the Sec translocon (159) and their secretion through the translocon can be prevented by the

SecA inhibitor, azide (105). Accordingly, azide blocks spontaneous lysis by gp*79* **(**Fig. 4.1A**)**. Taken together, these results unambiguously categorize Bcep22 gp*79* as a SAR endolysin.

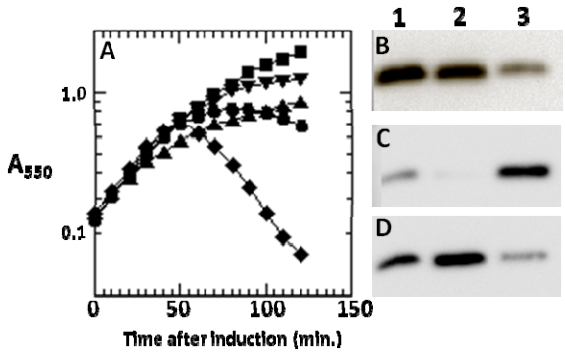

FIG. 4.1. Bcep22 gp*79* is a SAR endolysin. In each experiment, cultures of MG1655 containing pBcep22gp*79* were induced at time 0. Optical density of the culture was followed as a function of time in mins. (*A*) Expression of Bcep22 gp*79* lyses the cell independently of a holin. (●) wt, (■) Bcep22gp*79* 2A2L, (♦) wt + 10mM DNP added at 50 min. after induction, (▲) wt + NaN_3 at time 0, (▼) Bcep22gp*79* E26Q. For (B-D), the total protein is represented by lane 1, the soluble protein fraction is represented by lane 2, and the membrane fraction is represented by lane 3. (B) Bcep22 gp*79* is in the soluble and membrane fractions of the cell. (C) Bcep22 gp*79* can be confined to the inner membrane by replacing two Glycines (residues 16 and 17) with Leucine residues. (D) Bcep22 gp*79* E26Q is non-lytic, but is still released from the membrane as efficiently as wt.

Bcep22 gp*79* represents a new class of SAR endolysins

As mentioned above, members of the GH24 family are thought to hydrolyze the MurNAc-GlcNAc glycosidic linkage like true lysozymes. These proteins, however, do not contain the T4 E-like E-5x-D/C-8x-T catalytic triad and are not true T4 E homologs. As a member of the GH24 family, Bcep22 gp*79* may constitute a new class of SAR

endolysins. Since all previously identified SAR endolysins are homologs of T4 gp*e*, and GH24 proteins are putative hydrolases, we wanted to know how the GH24 proteins were related to lysozymes. We were also interested in determining the relationship between SAR lysozymes and SAR GH24 proteins. Phylogenetic tree analysis was performed on the lysozyme and GH24 protein sequences (113 total sequences). Overall, the GH24 proteins (Fig. 4.2, purple (soluble GH24) and green (SAR GH24)) do not group together in a single clade. Instead, they share clades, mostly with SAR lysozymes. The base of the tree is unclear making it difficult to determine early evolutionary relationships; however, some conclusions can be drawn from the branches. Feature 1 highlights a clade containing a GH24 SAR protein from the *Panteoa* phage LIMElight and an R^{21}-like SAR lysozyme from *Klebsiella* phage KP34. Since both of these proteins contain a SAR domain, we could hypothesize that perhaps these proteins diverged from a common ancestor that was a SAR endolysin. It may even be that in this case, the SAR domain predates the functional split (GH24 vs. lysozyme) between the two proteins. Another interesting relationship is shown in feature 2 between P1 Lyz-like SAR lysozymes from *Vibrio* phages K139 and kappa, and the R^{21}-like SAR lysozyme of *Xylella* phage Xfas53. A similar relationship is seen in feature 5 between the SAR lysozymes of *Enterobacteria* phages T1, Fels-2, and RE-2010. In both features, the P1 Lyz-like and R^{21}-like SAR lysozymes share a branch point indicating that they may share a common SAR domain containing ancestor, perhaps indicating that the SAR domain was acquired before the development of the regulation style. Feature 3 points out a clade containing soluble GH24 proteins from *Salmonella* phages Gifsy-1 and Gifsy-2 among soluble lysozymes

from *Soldalis* phages phiSG1 and phiSG3 and a *Acyrthosiphon pisum* (Pea aphid) secondary endosymbiont phage APSE-1 (43). What is unique about these lysozymes is that rather than using an E-8x-**D**-5x-T catalytic triad like the other soluble lysozymes, these proteins use an E-8x-**C**-5x-T triad similar to P1 Lyz. These proteins could represent the missing link between SAR endolysins and soluble lysozymes (discussed in more detail below). A more defined relationship between the Gifsy-1 and Gifsy-2 proteins and the lysozymes is difficult to pinpoint due to a polytomy (a node with more than two branches) in the tree. A polytomy can indicate a speciation event but most often indicates a lack of genetic information and an inability to define phylogeny (48). Feature 5 shows another relationship between disulfide bond regulated SAR endolysins and R^{21}-like SAR endolysins. As stated previously, all R^{21}-like proteins identified to date use the canonical E-8x-**D**-5x-T catalytic triad. Of the disulfide bond regulated SAR endolysins, the SAR endolysins of *Erwinia* phages Era103 (74), phiEa100, and phiEa1H (70) are the only proteins to use the canonical triad as opposed to the P1 Lyz-like E-8x-**C**-5x-T triad. Since the *Erwinia* SAR endolysins are in a large clade of R^{21}-like proteins, we can hypothesize that in this instance, the SAR domain predates the disulfide bond regulation (discussed further below).

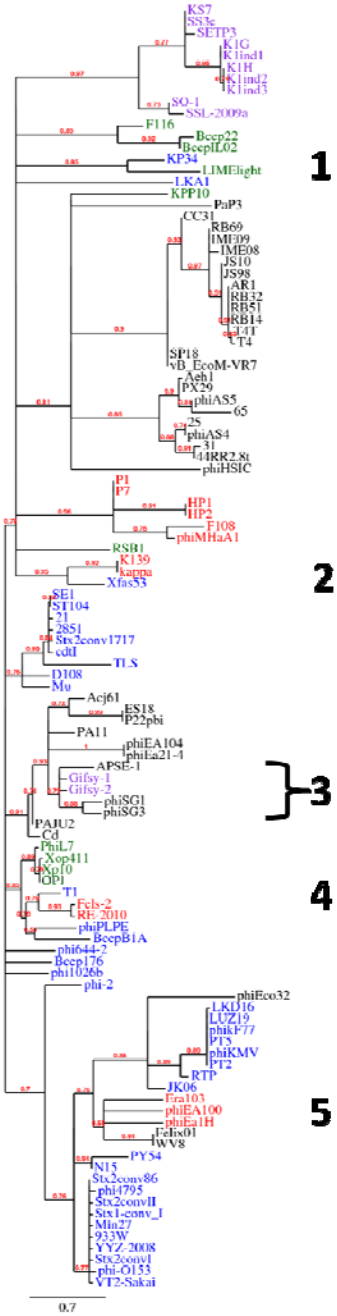

FIG. 4.2. Phylogenetic analysis of lysozymes and GH24 proteins. An unrooted maximum likelihood tree is presented where R^{21}-like SAR lysozymes are in blue, P1Lyz-like SAR lysozymes are in red, and soluble lysozymes are in black. GH24 SAR endolysins are represented in green and soluble GH24 proteins are in purple. For explanation of numbered features, see text. Branch support values are shown before each node in red. All branches with a support value less than 0.5 were collapsed. Branch lengths represent relative divergence although the unit of length is arbitrary.

BcepMu gp*22* is a putative SAR transglycosylase

Previously, it was thought was that all SAR endolysins were homologs of T4 E, a true lysozyme (74, 133, 158, 159). E uses a E-8x-**D**-5x-T catalytic triad in a hydrolase mechanism to cleave the β−1, 4 linkages between MurNAc and GlcNAc units (Fig. 1.3). The Glu residue acts as an acid to protonate the glycosidic linkage. The nucleophile at position 20 abstracts a proton from a water molecule, which then attacks C1 of the MurNAc moiety. The Thr at position 26 interacts with and stabilizes the water molecule involved in the hydrolysis of the glycosidic bond. What results is a nucleophilic substitution reaction in which the remaining peptidoglycan chain is the leaving group along with the glycoside oxygen (73). The cleavage mechanism results in inversion (β to α conformation) of the anomeric carbon.

Like lysozymes, transglycosylases like R from phage lambda cleave the β−1, 4 linkages between MurNAc and GlcNAc units (Fig. 1.4). The catalytic center of a transglycosylase is structurally similar to that of the lysozyme, except for the absence of the catalytic Asp in the transglycosylase (120). As with the muramidase reaction, the transglycosylase mechanism uses a catalytic carboxylic acid to protonate the glycosidic linkage (121). After the formation of a stable intermediate involving the N-acetyl group of the MurNAc moiety, the glycosidic linkage oxygen leaves with the remainder of the PG chain. The catalytic Glu then acts as a base to remove the C6 hydroxyl proton of the intermediate. The final step occurs when an intramolecular, nucleophilic attack at C1 by the C6 oxygen collapses the intermediate. This final step transfers the muramoyl bond

back onto the C6 hydroxyl group of the same muramic acid moiety. This results in the formation of a non-reducing, 1, 6-anhydro-N-acetylmuramyl product.

Burkholderia phage BcepMu contains a putative GEWL-type LT that may be a SAR endolysin. BcepMu thus represents yet another new class of SAR endolysins. We were interested in determining the relationship between the SAR LTs and the soluble LTs. More specifically, we wanted to know if the SAR LTs evolved independently of the other LTs. The importance of this relationship is that if the two groups evolved independently (i.e. clade separately), we may be able to infer that the SAR domain was acquired early in LT evolution. To determine these evolutionary relationships, phylogenetic tree analysis of lambda R and GEWL LTs was performed (Fig. 4.3). Unlike the GH24 SAR endolysins, the putative LT SAR endolysins (Fig. 4.3, red) are mostly grouped together in a clade highlighted in feature 1. An outlier of feature 1 is the *Pseudomonas* phage B3 (21). When analyzed by TMHMM, the B3 LT is predicted to have two N-terminal TMDs and an N-out, C-out topology making it a putative membrane bound LT.

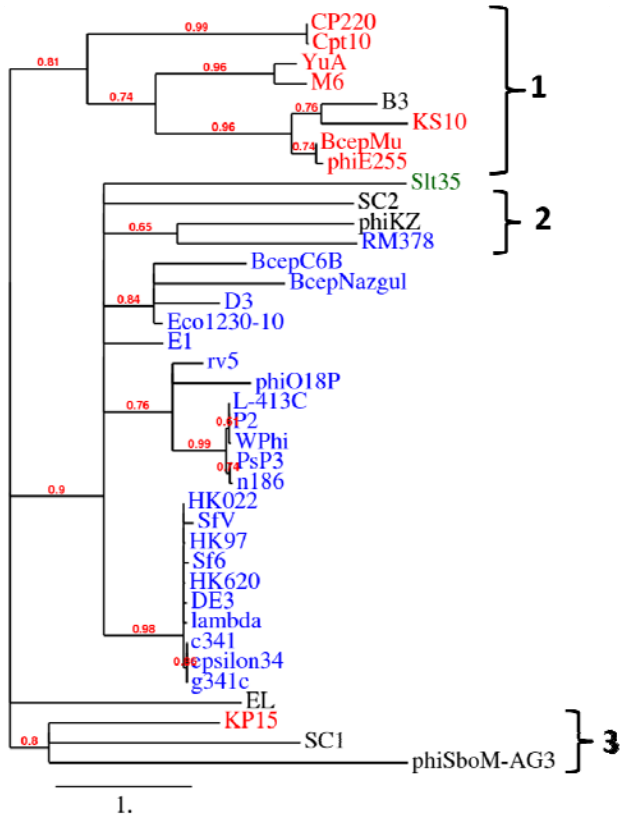

FIG. 4.3. Phylogenetic tree of lytic transglycosylases. An unrooted maximum likelihood tree is presented where lambda R homologs are in blue, GEWL homologs are in black, and GEWL SAR endolysins are in red. For explanation of numbered features, see text. For explanation of numbered features, see text. Branch support values are shown before each node in red. All branches with a support value less than 0.5 were collapsed. Branch lengths represent relative divergence although the unit of length is arbitrary.

The evolutionary origin of the soluble GEWL LT from *Liberibacter* phage SC2 (feature 2) is difficult to resolve due to a polytomy. The soluble GEWL LT in this feature from *Pseudomonas* phage phiKZ appears to be related to the lambda R-like protein from *Rhodothermus* phage RM378. This suggests that GEWL LTs and lambda like LTs may share a common ancestor. Feature 3 shows a predicted relationship between the SAR GEWL protein from *Klebsiella* phage KP15 and soluble GEWL LTs

from phages SC1 and phiSboM-AG3, although the exact origin is difficult to resolve due to a polytomy. In this case, perhaps KP15 picked up its SAR domain in a recent transfer event. The LT tree was derived with the *E. coli* Stl35 (40) (Fig. 4.3, green). Slt35 groups in the major clade containing the lambda-like LTs possibly suggesting horizontal gene transfer between bacteria and phage (discussed further below).

ΦV10 gp*29* has PG degradation activity

Many phage endolysins (8 %) are classified as general function chitinases by conserved domain analysis. We were interested in determining the enzymatic substrate specificity of the phage chitinases, more specifically, if they could degrade bacterial PG. The chitinase from phage ΦV10, gp*29*, was chosen for further experimental analysis. Phage ΦV10 is a temperate phage specific for EHEC strain O157:H7 (112). As expected, cells expressing the presumably cytoplasmic gp*29* did not lyse. Unexpectedly, the addition of $CHCl_3$ did not achieve lysis. However, when cells expressing gp*29* were subjected to a round of freeze-thawing, culture clearing was observed, suggesting the gp*29* is $CHCl_3$-sensitive. Despite classification as a chitinase, purified ΦV10 was shown to have activity against bacterial PG by turbidometry analysis (Fig. 4.4). Several *Pseudomonas* chitinases have been shown to have dual chitin and PG degrading activities (149). However, the chitinase activity of ΦV10 gp*29* was not detectable using a colorimetric colloidal chitin azure assay. This assay measures chitinase activity using chitin from crab shells covalently linked to Remazol

Brilliant Violet 5R dye as a substrate. Chitin degradation is measured as a function of dye release.

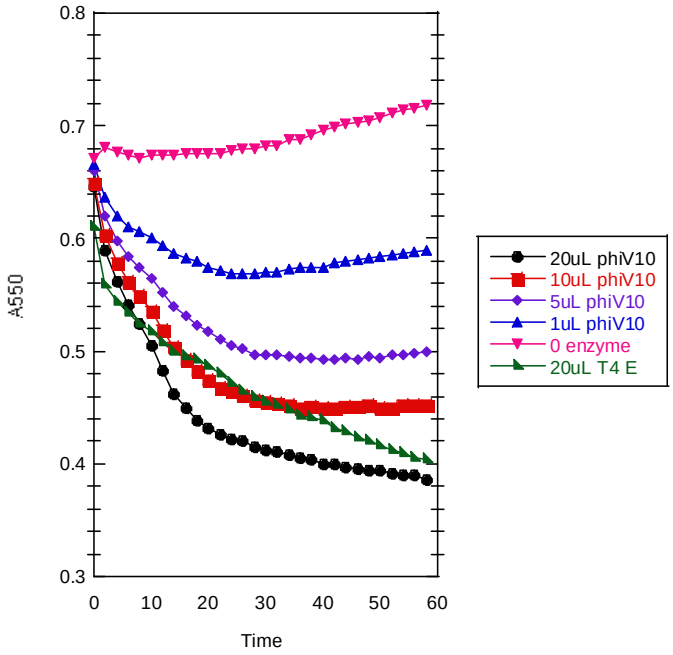

FIG 4.4. ΦV10 gp*29* has PG degradation activity. The endolysin from phage ΦV10 is a member of the glycoside hydrolase family 19 (GH19) with a predicted chitinase function but shows PG degradation activity. The indicated amount of pure protein (approx. 1mg/mL) was added to 100µg *M. lysodeikticus* cells. Optical density was monitored over time. Pure T4 E acted as an assay control.

Amidase phylogenetic analysis

Overall, the amidases show very little sequence divergence. Since a majority of the amidase containing phages are similar T7-like phages, we were interested in knowing if the amidase proteins were also similar to each other. To address this, phylogenetic analysis was performed on the amidase sequences. We were also interested in the possible evolutionary relationships between phage and bacterial enzymes, the *E.*

coli amidases involved in cell wall turnover, AmiA, AmiB, and AmiC were included in the amidase analysis (Fig. 4.5). The amidases form two major clades (Fig. 4.5, features 1 and 2). *Roseophage* SIO1 effectively acts as an outgroup, but is a distant relative of phage T7 (123). The *E. coli* amidases group in clade 1 with the bulk of the phage amidases suggesting possible host-to-phage horizontal transfer (discussed further below). The host amidases are most related to the amidases from the T4-like *Acinetobacter* myophages Ac42 and Acj9. These phages are unusual because most (28 of 34) of the amidase containing phages are T7-like podophages. BLAST analysis of AmiA, AmiB, and AmiC show homology with amidases from other bacterial species including *Shigella*, *Salmonella*, and *Klebsiella*.

Peptidase phylogenetic analysis

In contrast to the amidases, the peptidases display low sequence similarity. This may suggest isolated evolution of members of this class. To test this, phylogenetic analysis was performed on the peptidases sequences (Fig. 4.6). The tree is rich in polytomies, possibly due to the high sequence diversity, making analysis difficult. Again, a polytomy may represent a speciation event, possibly through species isolation, but can also mean that there is insufficient information to form a bifurcated branch. As with the amidase tree, *E. coli* metallopeptidase MepA, which is involved in PG turnover, was included in the analysis (89). Rather than act as an outlier, MepA clades with the phage peptidases suggesting a possible gene transfer event. BLAST analysis shows

MepA homologs in other Gram-negative species such as *Shigella*, *Salmonella*, and *Erwinia*.

An interesting find among the peptidases are phages SP6 and K1-5, which infect *Salmonella* and *E. coli*, respectively (38, 123). SP6 and K1-5 are members of the T7 supergroup. T7 and its homologs encode gp*3.5* which acts as an amidase for host cell lysis. T7 gp*3.5* also binds to and inhibits T7 RNA polymerase (94) to aid in the transition from transcription to replication and packaging of phage DNA (168). As members of the T7 supergroup, SP6 and K1-5 were expected to encode gp*3.5*-like amidases rather than peptidases. Genome analysis reveals that SP6 and K1-5 are more closely related to each other than they are other members of the T7 supergroup, thus they have been placed in their own 'SP6 group' within the T7 supergroup (123). The divergence of the SP6 group from other members of the T7 group may explain why these phages do not have T7-like amidases.

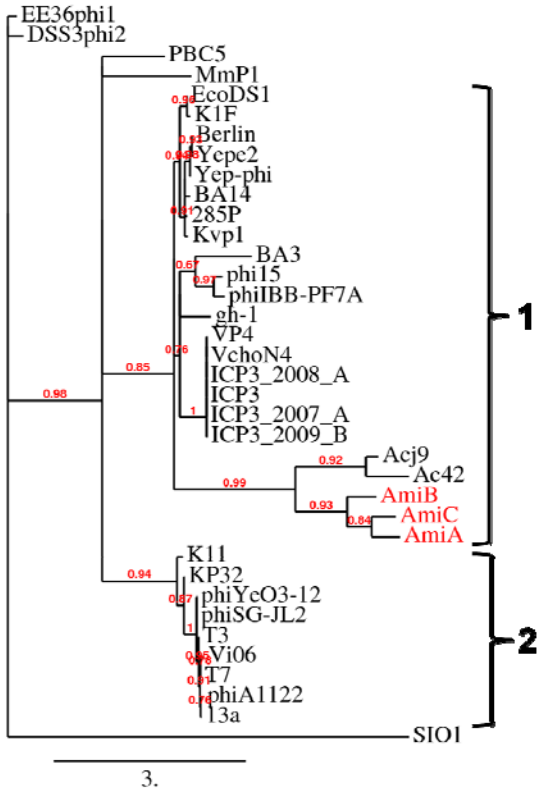

FIG. 4.5. Phylogenetic tree of amidases. An unrooted maximum likelihood tree is presented where phage amidases are in black and *E. coli* amidases are in red. For explanation of numbered features, see text. Branch support values are shown before each node in red. All branches with a support value less than 0.5 were collapsed. Branch lengths represent relative divergence although the unit of length is arbitrary.

The transfer of genes from phage to bacteria occurs frequently (25, 102) and the transfer of genes within phage communities has been well established (36). The amidase and peptidase phylogenetic analyses presented here may represent a case of a phage acquiring and using a host gene. A host autolysin could likely serve a phage well as an endolysin. In the context of a phage infection, endolysins are exported by a holin. Holins form non-specific lesions in the IM and function irrespective of the endolysin. It is

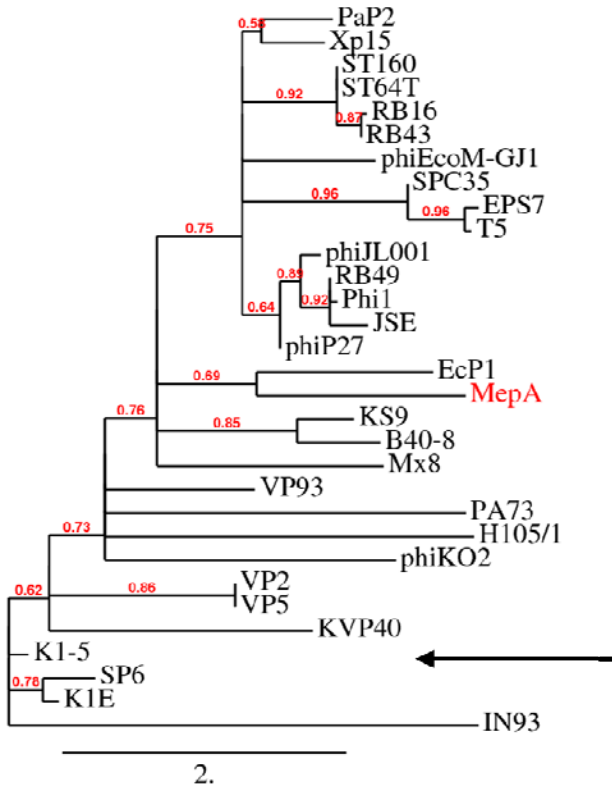

FIG. 4.6. Phylogenetic tree of peptidases. An unrooted maximum likelihood tree is presented where phage peptidases are in black and the *E. coli* peptidase MepA is in red. Branch support values are shown before each node in red. All branches with a support value less than 0.5 were collapsed. Branch lengths represent relative divergence although the unit of length is arbitrary.

reasonable that soluble endolysins of phage using the classic lysis paradigm are interchangeable and could possibly be replaced with a host protein of the same function.

Discussion

SAR endolysin sub-classes

So far, among T4 E homolog SAR endolysins, two subclasses have been established: those that contain Cys residues in their SAR domains and those that do not. Those that contain Cys residues like P1 Lyz are presumably regulated by a disulfide bond isomerization after release of the SAR domain from the bilayer. Those that do not, like R^{21}, are hypothesized to be activated by refolding of the active site upon SAR domain release. The disulfide bond regulated SAR endolysins can be further broken down into two sub-categories based on catalytic triad. Those that contain an ECT triad are most likely regulated by an inactivating disulfide as seen with P1 Lyz. A few disulfide bond regulated SAR endolysins (Lyz^{103} and others) use the canonical EDT triad. All of the members of the R^{21}-like SAR endolysins found in this search use an EDT triad. Of noted interest were the phages that contain a soluble T4 E homolog with an ECT triad. These proteins could be the missing link between P1 Lyz like SAR endolysins and soluble T4 E homologs (discussed in more detail below).

Generally, SAR endolysins can also be broken into three classes distinguishable by the distance of the catalytic residues from the SAR domain (Fig 4.7). P1 Lyz and Lyz^{103} represent a class whose catalytic residues 16 residues or farther from the SAR domain. Since the catalytic triad of these proteins lies far into the periplasm, it is reasonable that in order to prevent premature lysis, these SAR endolysins must have a more substantial, dedicated negative regulation mechanism in the form of DSB

inactivation. In this class, the SAR domain contributes only the free sulfhydryl and is not necessary for activity if exogenous reductant like DTT is added to the culture (74, 158). R^{21} represents a second, middle class whose catalytic residues lay 6 to 10 residues from the SAR domain. These proteins are dependent on the release of the SAR domain for proper folding of the catalytic cleft. It is thought that this refolding is SAR domain specific, in that the P1 Lyz SAR domain cannot activate R^{21} (133). Bcep22 gp*79* represents the third class of SAR endolysins in which the catalytic residue(s) is positioned directly adjacent to the end of the SAR domain. These SAR endolysins may be regulated by simple membrane proximity in which the catalytic residues are held close to the membrane and 'out of reach' from the PG. This question would be answered by testing a Bcepp22 gp*79* construct in which the SAR domain was replaced by a cleavable secretion signal. If the protein is regulated by membrane proximity, the expectation would be that secretion of the periplasmic domain would result in rapid lysis. However, when such a construct was induced, no lytic activity was observed, nor was there any detectable protein accumulation, suggesting that this protein is unstable (data not shown).

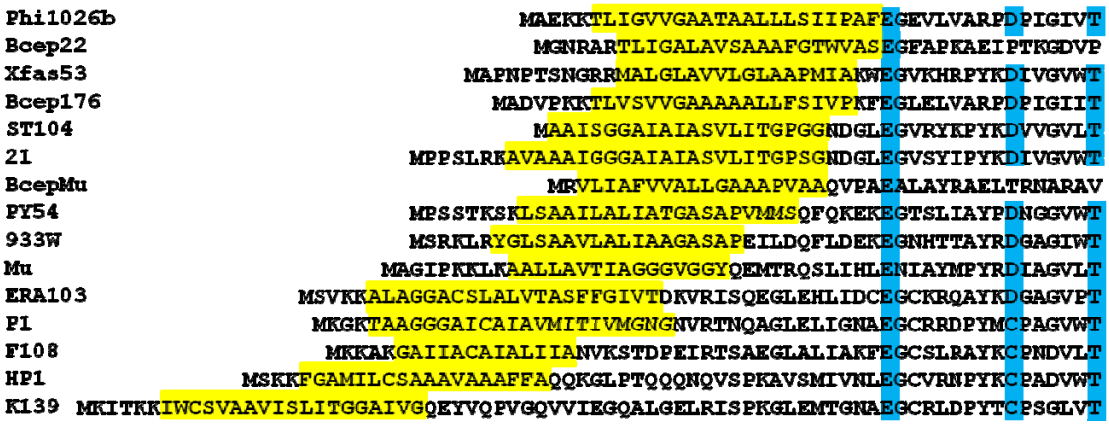

FIG. 4.7. Alignment of SAR endolysins by catalytic residues. SAR endolysins fall into three sub-classes defined by the distance of the catalytic triad from the SAR domain. Here, the N-terminal domains of SAR endolysins are aligned by their catalytic residues. The SAR domain is highlighted in yellow and catalytic residues in blue.

The argument above addresses the question of why we do not see SAR endolysins in the other classes of endolysins such as the amidases and peptidases. The catalytic residues in these two classes of proteins lie much further from the N-terminus of the protein making negative regulation of a membrane tethered amidase or peptidase problematic. The catalytic cleft of amidases lies in the center of the protein (30) as opposed to in an N-terminal lobe as it is with lysozymes and LTs. It is conceivable that this arrangement of the catalytic cleft is not conducive to negative regulation and thus the endolysin must be sequestered in the cytoplasm during the morphogenesis period. Among the SAR endolysins, the most extreme case in terms of catalytic residue distance from the SAR domain is seen in *Vibrio* phage K139. In this instance, the catalytic Glu lies 33 residues from the end of the SAR domain (Fig. 4.6) however, K139 uses a catalytic Cys analogous to Cys51 of P1 Lyz and has an upstream Cys analogous to C44

of P1 Lyz. So, despite the large distance from the SAR domain to the catalytic triad, tethered K139 would remain covalently inactivated, presumably in a manner similar to P1 Lyz.

Evolution of phage lysis

In general, there are three possible models for the evolution of dsDNA phage lysis. The first model places the SAR paradigm before the canonical paradigm. Since a SAR endolysin can lyse a cell without a holin, it is the simplest, self-contained lysis protein. Thus, the earliest lysis systems could have contained only a SAR endolysin. A primordial, SAR endolysin-only system would be somewhat inflexible in regards to lysis timing. This simple paradigm would have developed a more fine tuned timing mechanism by adding a holin to allow for synchronous release and activation of the protein. The holin would only have to form lesions large enough to deplete the PMF (i.e. a pinholin). The next step towards the classic lysis paradigm is the development of the holin to one that forms holes large enough to allow a soluble, cytoplasmic endolysin passage. Interestingly, the holin of phage P1 complements lambda ΔS (157), despite the fact that the P1 endolysin is a SAR protein. Thus phage P1 could represent this step in lysis evolution. If P1 Lyz were to lose its SAR domain and fold in the cytoplasm, it could hypothetically still be released into the periplasm by holin triggering. In fact, several phages contained soluble T4 E homologs that have the P1 Lyz like ECT catalytic triad. Perhaps these proteins represent this step in lysis evolution. From here, the protein need only gain the beneficial Cys to Asp mutation in the catalytic site, as it has been

shown that the Cys to Glu mutation in T4 E results in a 20% decrease in catalytic activity (58).

A second model is that the two component system consisting of a canonical, cytoplasmic endolysin and a large-hole holin came first in the evolution of phage lysis. This would require the acquisition of a SAR domain affixed to a cytoplasmic endolysin. Since holin genes and endolysin genes are adjacent in the lysis cassette, a possible mechanism for this would be a partial gene duplication event that placed a TMD at the N-terminus of the endolysin. If this duplication leads to a successful gene fusion, the endolysin could become membrane bound. Previous experiments have shown that SAR endolysins are capable of lysis in the membrane by proximity (74), thus a membrane bound protein would presumably not allow for lysis time modulation. The TMD could evolve to a SAR domain by mutation of hydrophobic residues to weakly hydrophobic residues such as Ala. At this point, the pseudo-SAR endolysin could evolve a control mechanism for lysis timing by developing a mode of regulation based on its dynamic membrane topology.

An alternative to the above described models is that the canonical lysis paradigm and the SAR lysis paradigm originated and evolved independently of one another. This isolated evolution does not mean, however, that both holins and endolysins evolved separately. For example, a single holin could have evolved and then lent itself to other phages by horizontal gene transfer. Due to the variation seen in endolysin mechanisms (i.e. lysozyme, amidase, etc.) it is difficult to conclude if any one endolysin mechanism

predates the others. The phylogenetic trees presented do not have clear roots complicating any attempt to predict a common ancestral endolysin.

Overall, dsDNA phages of Gram-negative hosts have developed a rapid and malleable method by which to lyse cells. The holin-endolysin coupling allows for adaptability and optimization since single amino acid changes in the holin result in drastic changes in lysis time (153). Despite its possible position as the most primitive lysis system in dsDNA phages, the SAR endolysin system remains strongly represented composing a large portion (71 of 279 or 25%) of endolysins found in this search. The utility of the SAR endolysin as a well controlled, yet rapidly dispersed lysis protein has maintained it in the gene pool and has possibly allowed for its emergence among other muraminidase mechanisms.

CHAPTER V

SAR ENDOLYSINS AS POTENTIAL DRUG TARGETS

Introduction

Enterohemorrhagic *Escherichia coli* (EHEC) is the causative agent of many cases of food poisoning and can cause the deadly hemolytic uremia syndrome (HUS) in children and the elderly. The disease state associated with EHEC is caused by the Shiga toxin (Stx). Stx is secreted by several human pathogens including *Shigella dysentariae*, *E. coli* O157:H7, and *E. coli* O104:H4. *E. coli* pathogens that carry the toxin are generally referred to as Shiga toxin-producing *E. coli* (STEC). According to recent estimates by the Centers for Disease Control, 265,000 cases of STEC are reported each year worldwide. Cases of EHEC are on the rise worldwide and over 2,000 cases were reported due to the most recent outbreak in Germany.

The shiga toxin, Stx2, encoded by STEC is in the AB5 family of toxins. The Stx2 toxin proteins form a complex composed of a monomer of subunit A and a pentamer of subunit B. Subunit A is an N-glycosidase that cleaves an adenine residue from the 28S RNA of the 60S ribosomal subunit while the B pentamer is responsible for endothelial cell invasion (78). The Stx2 toxin genes of STEC strains are transcribed from the late promoter of a lambdoid prophage along with the phage lysis genes (Fig. 1.11). STEC strains typically have >20 prophages, but in general only one can give rise to viable phage particles, while the others are cryptic. For example, the active, Stx encoding prophage of O157:H7 strain EDL933 is 933W (61).

As with lambda, the late promoter of 933W, pR', is induced after prophage induction of the vegetative cycle. Upon commitment to the vegetative cycle, the prophage must excise from the host chromosome, replicate DNA, and turn on the late genes, which include the morphogenesis and lysis genes. To induce the late pR' promoter, Q, the phage encoded anti-terminator binds to Q utilization sites (*qut*) and then to the RNA polymerase, allowing read-through of downstream terminators (160).

During STEC infections, spontaneous prophage induction occurs in a fraction of the cell population in the gut (81). In addition, treatment with any DNA-damaging antibiotics causes population-wide induction by inducing the SOS response. The phage protein responsible for repressing the lytic promoters during lysogeny, CI, is a homolog of the SOS response repressor protein LexA. DNA damage activates RecA, the universally conserved bacterial recombinase/co-protease, leading to cleavage of both LexA and the LexA-like CI repressor (79). Thus not only are all the LexA-repressed promoters derepressed (the SOS response), the prophage is also induced. Consequently, DNA damaging antibiotics such as norfloxacin, a gyrase inhibitor, exacerbate EHEC infections and are not recommended for therapy.

EHEC contains the LEE pathogenicity island, which encodes a type III secretion system (5). The secretion system supports the release of proteins responsible for cell adhesion, but not the Stx2 toxin. A recent study shows that the Stx2 toxin relies on phage induced lysis to escape the cell (124). When O157:H7 strain EDL933 was induced for the 933W prophage, Stx2 was found extracellularly. However, when a mutant 933W ΔSR prophage, lacking the holin and endolysin genes, was induced, Stx2 was completely

restricted to the bacterial cells. Thus, it is the phage's lysis machinery alone that is responsible for toxin release.

Inspection of R^{933W}, the muralytic enzyme of 933W, suggests that it is a SAR (Signal Anchor Release) endolysin. SAR endolysins are signaled for secretion to the IM by the Sec translocon, and are anchored in an inactive conformation with their catalytic residues in the periplasm (159). The SAR domain is maintained in the energized bilayer. PMF depletion by the holin causes synchronous release of the SAR endolysin resulting in rapid lysis (Fig. 5.1). In the absence of the holin, the SAR domain will spontaneously release from the membrane resulting in gradual, holin-independent lysis. This characteristic of SAR endolysins makes the holin a poor target for drug development. If the holin function is disrupted, toxin accumulation would continue until SAR endolysin mediated lysis. The holin functions independently of the endolysin. If the SAR endolysin function is disrupted, then the holin will still trigger immediately depleting the PMF, killing the cells. In the context of an Stx phage infection, toxin would be trapped in the dead cells. This makes the SAR endolysin a potential drug target in EHEC infections.

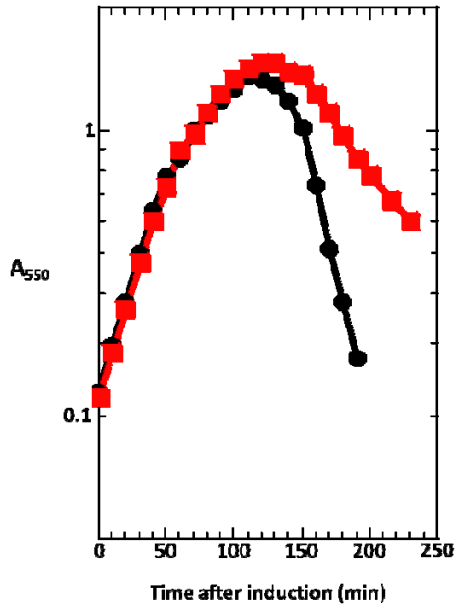

FIG. 5.1. SAR endolysins lyse cells rapidly with a holin. In the absence of a holin, SAR endolysins release from the membrane spontaneously resulting in gradual lysis. In the presence of a holin, lysis occurs rapidly. (●, black) MG1655 *ΔfhuA lacI*^{q} pQ pR'(S,R,Rz,Rz1)^{933W} (■, red) MG1655 *ΔfhuA lacI*^{q} pQ pR'(S_{am}R,RzRz1)^{933W}.

Materials and methods

Bacterial strains and growth conditions

The *Escherichia coli* strains XL1-Blue, MC4100, and MG1655 Δ*fhuA lacY lacI*^{q}, and RY8797 Δ*cheA* have been described (53, 127, 133). Standard conditions for the growth of cultures and the monitoring of lysis kinetics have been previously described (29, 127). All bacterial cultures were grown in standard LB medium, supplemented with 100 μg/ml ampicillin when appropriate. When indicated, isopropyl β-D-thiogalactopyranoside (IPTG) and dinitrophenol (DNP) were added to final concentrations of 1 mM and 10 mM, respectively. At the indicated time, 6-

nitroquinolin-8-yl 2,3,3,3-tetrafluoro-2-methoxypropanoate (ChemBridge, San Diego, CA), designated here as 67-J8, was added to the specified final concentration. At the indicated time after induction, an aliquot of cells were removed from an aerating culture and phase-contrast images were taken using a Nikon Digital Sight DS-5M camera on a Nikon OPTIPHOT microscope.

DNA procedures and plasmid construction

Procedures for the isolation of plasmid DNA, DNA amplification by PCR, PCR product purification, DNA transformation, site-directed mutagenesis, and DNA sequencing have been previously described (52, 126, 128). The construction of the plasmid pP1 Lyz, a derivative of pJF118EH, has been described previously (45, 159). The plasmid pZER^{933W}cHis, a derivative of the medium copy, IPTG-inducible vector pZE12 (86), was constructed by PCR amplifying R^{933W} from pET41b R^{933W} cHis (provided by Quingan Sun) and inserting it between the *Kpn*I and *Xba*I sites of pZE12 digested with the same enzymes. The construction of the plasmid pZER^{21} has been described previously (133). Derivative alleles were constructed using site-directed mutagenesis.

Subcellular fractionation

To determine whether a particular protein was present in the soluble or membrane fractions, 25 mls of an induced culture were collected by centrifugation at 5000 x g and resuspended in 2 ml of buffer (0.1M sodium phosphate, 0.1M KCl, 5 mM

EDTA, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, pH 7.0) and then disrupted by passage through a French pressure cell (Spectronic Instruments, Rochester, N.Y.) at 16,000 lb/in^2. The membrane and soluble fractions were separated by centrifugation at 100,000 x g for 60 min. Equivalent amounts of the fractions were examined by SDS-PAGE and Western blotting as described below.

SDS-PAGE and Western blotting

SDS-PAGE, Western blotting, and immunodetection experiments were performed as previously described (52). A mouse monoclonal antibody against the oligo-histidine epitope tag was purchased from Amersham and was used at a dilution of 1:3000. The anti-mouse IgG horseradish peroxidase-conjugated secondary antibody was supplied with the SuperSignal chemiluminescence kit (Pierce) and was used at a 1:5000 dilution. Blots were developed using the West Femto SuperSignal chemiluminescence kit (Pierce). Chemiluminescent signal was detected using a Bio-Rad ChemiDoc XRS.

***In vivo* high through-put inhibitor screening**

A unique, high-diversity library containing compounds filtered for drug-like properties was screening using high through-put methods. The library, developed by J.S. Sacchettini and T.R Ioerger, contains non-redundant compounds selected for their small size (< 600 Da) and lack of reactive chemotypes and aggregative properties. MG1655 Δ*fhuA lacY- lacI*^{q} cells containing pZER^{933W} cHis were grown to an optical density (O.D.) at 550nm of 0.2. 40 μL of the actively growing culture was added to each well of

384 well plate containing 10 μL of 4X LB, 4 mM IPTG and 1 μL 1 mM compound dissolved in DMSO. All liquid manipulations were performed using a Cybio Vario robot. An initial O.D. reading was taken at 550nm on a BMG Labtech Polarstar Omega plate reader after culture addition (and subsequent induction of R^{933W}) and readings were taken every hour for five hours. Hit compounds were ones that allowed growth as opposed to SAR endolysin mediated lysis and were rescreened by hand.

***In vitro* lysozyme assay**

Lysozyme activity was assayed with the EnzChek® Lysozyme Assay Kit (Invitrogen, Carlsbad, CA) which measures lysozyme activity on fluorescently-labeled *Micrococcus lysodeikticus* cell walls. Two μL of 2.5 μM R^{933W} in reaction buffer (0.1 M sodium phosphate, 0.1 M NaCl, pH 7.5, 2 mM sodium azide) was used in each experiment. 67-J8 was serial diluted, in a range from 6.5 μM to 0 μM, into wells. One hundred μL fluorescently-labeled *M. lysodeikticus* cell wall suspension was added as the substrate to start the reaction. Fluorescence was measured in a POLARStar Optima plate reader (BMG LabTech) using excitation/emission of 485/520 nm every 2 min.

***In vitro* lysozyme assay, $CHCl_3$ method**

Lysozyme activity against *E. coli* PG was determined using a modified *in vitro* lysozyme assay. *E. coli* cells were grown to an O.D. of 0.6 to 0.7. Cells were pelleted and resuspended in 5mL cold LB. $CHCl_3$ was added to a final volume of 10%. The mixture was inverted gently to mix and was allowed to settle at room temperature for

five min. $CHCl_3$ was removed and cells were washed twice more with $CHCl_3$. $CHCl_3$ treated cells were added to wells containing the indicated amount of purified protein or lysate. Hen egg white lysozyme (HEWL) and $CHCl_3$ treated cells alone served as assay controls. To prepare R^{933W} lysate, MG1655 Δ*fhuA lacY*⁻ $lacI^q$ cells containing $pZER^{933W}$ cHis were grown to an O.D. of 0.2 and were induced with IPTG. Cells were pelleted, resuspended in buffer (50 mM Tris, pH 8, 100 mM NaCl), and French pressed. Lysate was added in the indicated amounts to wells with and without compound. O.D. measurements at 550 nm were taken every 2 minutes using a Tecan Infinite 200 PRO plate reader.

Tethering assay and video recordings

RY8797, a Δ*cheA* derivative of RP437 (110) that rotates only counterclockwise, was used for all experiments. An overnight culture of RY8797 was grown in TB at 37°C. The culture was back diluted into fresh TB and grown to late log phase (A_{550}= 0.6-0.7). Cells were harvested, washed once with TB, and flagella were sheared in a Waring blender with four, seven second pulses. Cells were washed three times with TB and were incubated on a glass coverslip for 20 minutes with anti-flagellar filament antibody (provided by Dr. Mike Manson, Texas A&M University). Cells were observed by phase contrast microscopy in a PeCon closed cultivation perfusion chamber on a Zeiss Axio Observer.A1 scope affixed with an AxioCam HSm camera. Cells were equilibrated with TB from a Warner Model SH-27B inline heater pre-warmed to 37°C. Flow was stopped and video of six fields of view was taken for one min each. Flow was restored for five

min with TB plus 25 μM 67-J8. Flow was stopped and video was taken of six fields of view for one min each.

Image processing and statistics

Individual cells were chosen from a recorded field of view based on consistency of spinning. Cells that stopped frequently or did not spin horizontally relative to the slide were not counted. For both groups, before and after compound flow, all of the cells that met these criteria in a given field of view were counted. Most cells in any particular field of view were not spinning. The most likely reason for this was incomplete shearing. To achieve 20-30 countable cells, six fields of view were captured for each condition. After video acquisition, videos were trimmed to highlight individual spinning cells using the Cutter option in the AxioVision software. The number of rotations of individual cells was counted by hand. The lowest and highest data points were thrown out of each set as outliers. In the 'after-compound' data set, three cells were thrown out because they were larger than cells in the 'before compound' data set. A caveat of the tethering method is the synthesis of new flagella as cells grow and divide over time. Typically, this is overcome by the addition of chloramphenicol. Here, chloramphenicol was not used to avoid any possible drug interactions with 67-J8. The size of the cells thrown out may be coupled with flagellar synthesis which can alter, slow, or stop rotation. Histogram analysis was performed using KaleidaGraph. Rotational velocities were binned in 0.5 rotations/sec increments to account for the maximum standard error which was 0.46 in the 'after compound' data set.

Results and discussion

R^{933W} is a SAR endolysin

When expressed from a plasmid, R^{933W} lysed cells independently of holin triggering (Fig. 5.2A). This holin independent lysis is indicative of SAR domain spontaneous release from the membrane. Since SAR domain release is synchronously triggered due to holin-mediated membrane depolarization, it follows that addition of energy poisons such as DNP would have the same effect. As expected, DNP addition 20 min after induction results in premature lysis presumably due to R^{933W} release (Fig. 5.2A). A subcellular fractionation of cells expressing R^{933W} showed soluble and membrane bound protein of the same apparent molecular mass, characteristic of SAR proteins (74, 133, 159) (Fig. 5.2B). SAR domains rely on a preponderance of weakly hydrophobic, neutral, or polar residues within the predicted transmembrane sequences (159). In previous studies of SAR domains, replacement of 2-3 consecutive Gly or Ala residues eliminated the ability of the SAR domain to exit the bilayer. Accordingly, two consecutive Ala residues (Ala18, Ala19) were converted to Leu residues by site-directed mutagenesis of R^{933W}. Subcellular fractionation of cells induced for the double mutant revealed that the protein was restricted to the membrane fraction (Fig. 5.2C). Moreover, the endolysin was non-lytic (Fig. 5.2A). Taken together, these data demonstrate unambiguously that R^{933W} is a *bona fide* SAR endolysin.

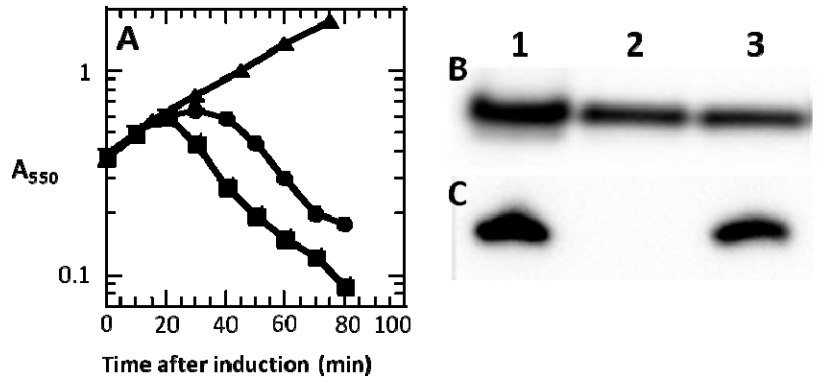

FIG. 5.2. R^{933W} is a SAR endolysin. (*A*) Induction of R^{933W} and derivative alleles. (●) pZE12 R^{933W}, (■) pZE12 R^{933W} + 10mM DNP at min after induction, (▲) pZE12 R^{933W} A18,19L. (B-C) Localization of R^{933W} and derivatives after induction of the indicated plasmids. Lanes 1, 2, and 3 represent the total, soluble, and membrane fractions, respectively. (B) pZE12 R^{933W} , (C) pZE12 R^{933W} A18,19L.

High through-put inhibitor screening

To identify lysis inhibitors, we utilized an *in vivo* approach. Bacterial cultures harboring an IPTG inducible plasmid, pR^{933W}, were grown to mid-log phase. The culture was added to wells containing inducer and compound. Using our method of *in vivo* screening, there are three possible outcomes. If the compound has no effect, R^{933W} will be secreted and released from the IM lysing the cells. Addition of a toxic compound will result in cell growth cessation. During screening, several toxic compounds were found. The third outcome is that the compound inhibits SAR endolysin mediated lysis and results in cell growth. . This method was appealing because it demands compound entry into the cell or at least past the OM and PG, since R^{933W} functions in the periplasm. If the compound was unable to pass the OM or PG, then it would have no effect and would not be realistic as a drug. Our *in vivo* method also required that the compound is sustained in

cells. If the compound enters the cells but is metabolized, its efficacy would be reduced, making it a less attractive drug candidate. Nearly 8,000 compounds were screened in 384-well format. One compound, designated 67-J8 by plate and well location, respectively, prevented lysis (Fig. 5.3A). The lysis inhibiting capacity of 67-J8 was successfully scaled up from 40 µL in 384 well plates to 25mL shaker flasks in order to determine mode of action (Fig. 5.3B).

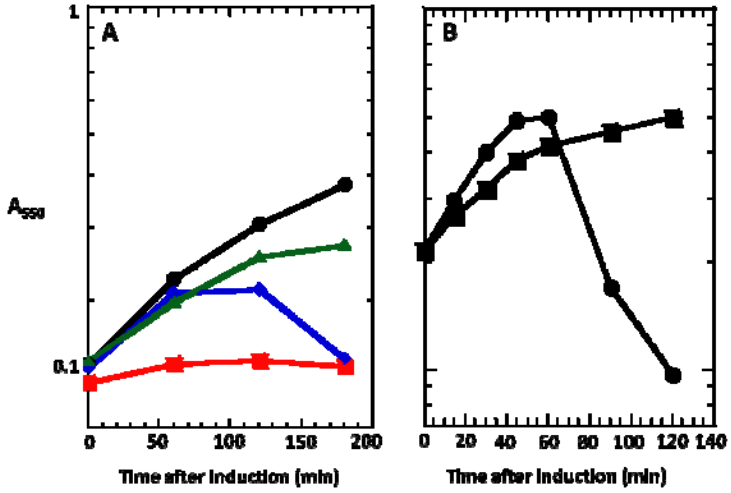

FIG. 5.3. Inhibitor screening curves. (A) *In vivo* lysis inhibitor screening was performed in high through-put format in 384 well plates as described in the text. Of 8000 compounds, a toxic compound was found (■, red). Many compounds had no effect (◆, blue) compared to an uninduced well (●, black), and one prevented lysis (designated 67-J8) (▲, green). (B) The lysis inhibiting ability of 67-J8 was scaled up to 25mL in a shaker flask. Cells treated with 25 µM 67-J8 at time 0 (■) did not lyse compared to cells with no compound (●).

67-J8 does not inhibit R^{933W} activity

Our initial motivation for high through-put screening of the compound library was to find compounds that inhibit the muralytic activity of R^{933W}. To test if 67-J8 was a

lysozyme inhibitor, an *in vitro* lysozyme assay was performed (Fig. 5.4). The assay measures lysozyme activity using fluorescein-labeled *M. lysodeikticus* cell walls as a substrate. The cell walls are highly labeled thereby quenching fluorescence. Cell wall hydrolysis relieves the quenching resulting in an increase in fluorescence. Activity is directly proportional to the increase in fluorescence over time. Lysozyme activity of pure R^{933W} (provided by Quingan Sun of the Sacchettini Lab, Texas A&M University) was tested in the presence of various concentrations of 67-J8. The compound had no effect on the activity of purified R^{933W} versus the fluorescent substrate *in vitro*, suggesting that it is not a lysozyme inhibitor. No change in fluorescence was observed suggesting that 67-J8 is not a lysozyme inhibitor.

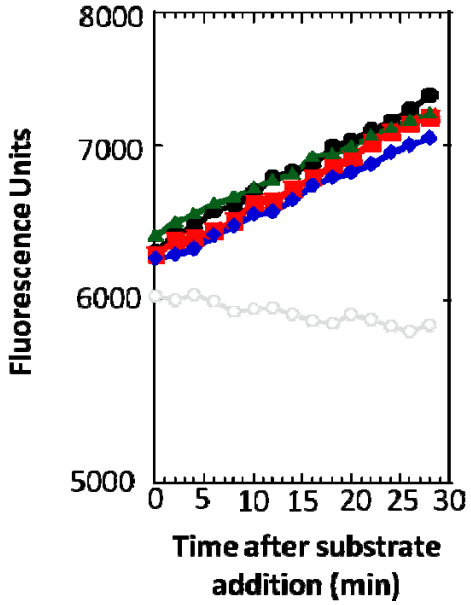

FIG. 5.4. *In vitro* lysozyme activity assay of R^{933W} treated with 67-J8. Pure R^{933W}, at 0.05 µM final protein concentration (provided by Quingan Sun), was tested for lysozyme activity in the presence of 67-J8. (●, black) 6.5 µM 67-J8, (■, red) 3.25 µM 67-J8, (◆, blue) 1.625 µM 67-J8, (▲, green) 0 compound, (○, grey) 0 R^{933W}.

67-J8 does not inhibit access to *E. coli* PG

The above described *in vitro* lysozyme assay tested R^{933W} activity against purified, derivatized murein from the Gram-positive bacterium, *M. lysodeikticus*. The chemical structure of PG shows considerable species-specific variation (122). Among Gram negatives, the variability is in the extent of crosslinking (91) and the length of the glycan strands (143). Nothing is known about R^{933W} substrate specificity. To test whether the compound's inhibitory character was PG-specific, a modified *in vitro* lysozyme activity assay was performed (Fig. 5.5). 0.1 μg (0.05 μM final protein concentration) of pure R^{933W} (provided by Quingan Sun) was added to wells containing various amounts of 67-J8. To prepare the substrate, *E. coli* cells were harvested, concentrated, and treated with $CHCl_3$. Presumably, after $CHCl_3$ treatment, the *E. coli* PG is exposed and susceptible to exogenous lysozyme addition. Lysozyme degradation of the PG results in cell lysis as determined by a decrease in turbidity over time. In the presence and absence of compound, no change in lysis was observed. This suggests that the compound does not hinder R^{933W}'s access to the *E. coli* PG.

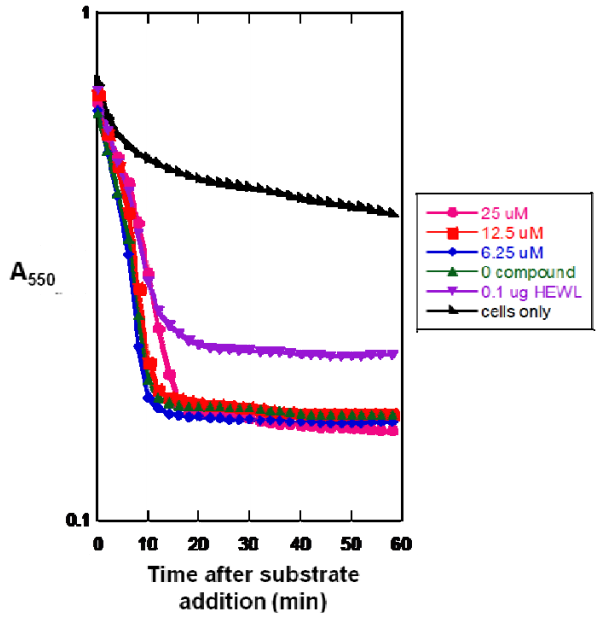

FIG. 5.5. 67-J8 mediated lysis inhibition is not PG specific. 0.1 µg pure R^{933W} (provided by Quingan Sun) was added to wells containing the indicated amount of 67-J8. At time 0, 100µL $CHCl_3$ treated *E. coli* cells were added. O.D. readings were taken every 2 minutes in a Tecan Infinite 200 PRO plate reader. 0.1 µg HEWL was added for a control.

67-J8 does not inhibit R^{933W} expression or accumulation

The experiments described above show that 67-J8 does not affect R^{933W} activity or PG access. There are many other testable targets; for example, the compound could prevent lysis by altering protein expression. Another possibility is that SAR endolysin secretion or accumulation in the IM is blocked. To determine if the compound was altering protein expression or accumulation, samples of cells expressing R^{933W} in the presence or absence of compound were assayed for protein via Western blot (Fig. 5.6A). No change could be seen between conditions, indicating that 67-J8 does not affect protein translation or accumulation. Periodically, the morphology of cells expressing

R^{933W} in the presence or absence of 25 μM 67-J8 was examined (Fig. 5.6B). After 60 minutes of induction, cells expressing R^{933W} in the presence of compound began to round up and were fully spherical in shape after 120 min. MG1655 cells treated with 25 μM 67-J8 for 120 min produced no round cells (Fig. 5.6C). The loss of rod-shape indicates the loss of PG integrity presumably by R^{933W} activity. Furthermore, the effect of 67-J8 could be immediately reversed by the addition of 10 mM DNP (Fig. 5.6D). Taken together, the above results suggest that R^{933W} secretion and activation is not affected by 67-J8 treatment.

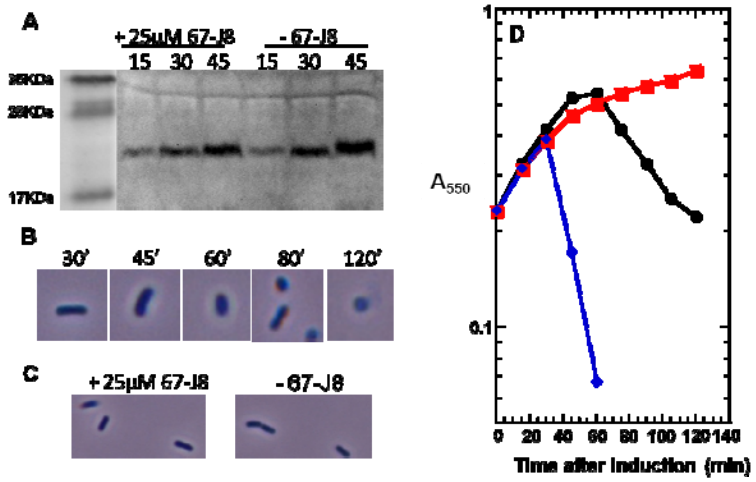

FIG. 5.6. 67-J8 does not alter protein expression or active protein accumulation. (A) At the indicated time after induction, cells expressing pR^{933W} in the presence or absence of 25 μM 67-J8 were TCA precipitated. Western blot analysis indicated no change in total protein yield. (B) At the indicated time, pictures were taken of representative cells expressing R^{933W} in the presence of 25 μM 67-J8. (C) MG1655 Δ*fhuA lacY*⁻ $lacI^q$ cells after 120 min of growth in the presence or absence of 25 μM 67-J8. (D) The addition of DNP reverses the effect of 67-J8. MG1655 Δ*fhuA* $lacI^q$ pQ pR' (SamR/Rz/Rz1) cells expressing $pZER^{933W}$ with: (●, black) no compound, (■, red) 25 μM 67-J8 added at time 0, (◆, blue) 25 μM 67-J8 added at time 0, 10mM DNP added at 30 min after induction.

67-J8 is a general inhibitor of SAR endolysin mediated lysis

We were interested in knowing if the pathway affected by 67-J8 was specific for R^{933W} or if it was shared with other SAR endolysins. SAR endolysins can be broken into two sub-classes based on the presence or absence of a Cys in the SAR domain. Those that contain a Cys, like P1 Lys from phage P1, are proposed to be regulated by a disulfide bond isomerization (158). Those without, like R^{21}, are proposed to be regulated by a conformational restructuring of the active site upon SAR domain release from the membrane (133). R^{933W} has no SAR domain Cys, so it is member of the R^{21}-like class of SAR endolysins. The lysis inhibiting capacity of 67-J8 was tested against P1 Lyz and R^{21} (Fig. 5.7). Both P1 Lyz and R^{21} mediated lysis was prevented by compound addition. Thus the lytic step targeted by 67-J8 is common to SAR endolysins regardless of their regulation type.

67-J8 mediated lysis inhibition is temporally dependent

To determine how soon after induction 67-J8 acts to prevent lysis, the time at which the compound was added was varied (Fig. 5.8). Cells containing pR^{933W} were induced with IPTG. At various times after induction, aliquots were removed and grown from then on in the presence of 25 μM 67-J8. Cells in which the compound was added at induction grew to a higher O.D. than those in which the compound was added later. Interestingly, when compound was added 40 min after induction, cells appeared to die. From this, we conclude that in order to be effective, the compound must be added soon after induction.

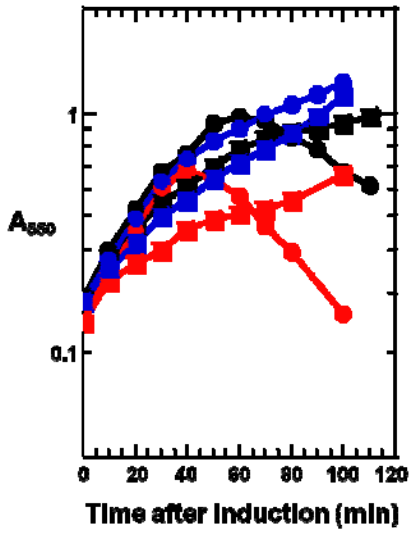

FIG. 5.7. 67-J8 prevents lysis by other SAR endolysins. MG1655 Δ*fhuA lacY*⁻ *lacI*^{q} cells containing pZER[21] (black) were induced in the presence (■) or absence (●) of 25 μM 67-J8 at time 0. MC4100 cells containing pP1 Lyz (red curves) were induced in the presence (■) or absence (●) of 25 μM 67-J8 at time 0. Blue curves represent empty MG1655 Δ*fhuA lacY*⁻ *lacI*^{q} cells with (■) and without (●)+ 25 μM 67-J8 at time 0.

67-J8 increases the proton motive force

Our next hypothesis was that 67-J8 was increasing the PMF. Presumably, the retention of the SAR domain is dependent on the PMF, since depletion of the PMF results in early lysis of cells expressing a SAR endolysin. It follows that an increase in PMF would retain the SAR domain in the membrane, possibly explaining the lysis block in compound treated cells expressing SAR endolysins. In order to test this, we monitored the PMF in individual *E. coli* cells as a function of time by recording the rotational speeds of cells tethered to a surface by a flagellum (125). The speed of rotation is proportional to the PMF over a dynamic physiological range (44).

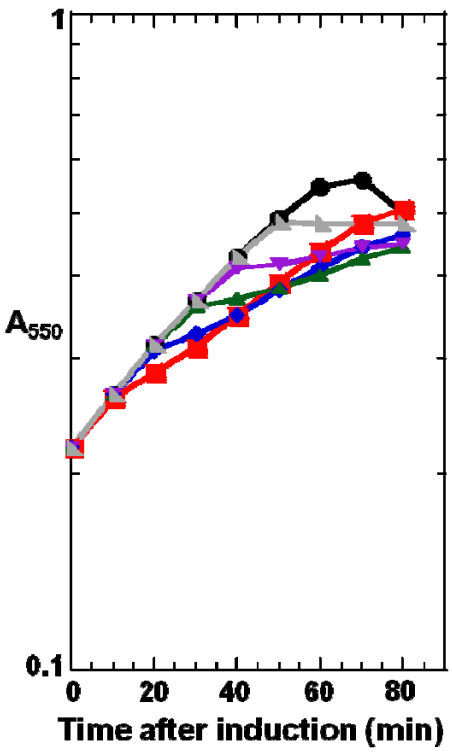

FIG. 5.8. The effectiveness of 67-J8 depends on time of addition. At the indicated time point, compound was added to a final concentration of 25 μM to cells expressing R^{933W}. (●, black) no compound, (■, red) 0 min, (◆, blue) 10 min, (▲, green) 20 min, (▼, purple) 30 min, (◣, grey) 40 min.

Motile cells were tethered to a glass coverslip by a single flagellum by using an antibody against the flagellar filament. Wild type motile cells frequently stop and change rotational direction. To avoid these pauses, a *cheA* mutant, in which the flagella spin only counterclockwise, was used in these experiments. Tethered cells were equilibrated for 5 min in a perfusion chamber by washing with TB from an in line heater pre-warmed to 37°C. Flow was stopped and video was taken for six fields of view for one minute each. In any one field of view, only ten to twenty percent of the cells rotate, as has been observed previously (125). Twenty-five μM compound in TB was then flowed in for five

min after which flow was stopped and video was taken as described above. The rotations of individual cells spinning at a constant rate were counted by hand during a 20 second span of video and rotational velocity (rotations/sec) was calculated (Fig. 5.9). Criteria for inclusion in the dataset and removal of outliers are described in materials and methods. Before compound treatment, the average rotational velocity was 2.6 rotations per second. After five min of compound treatment, the average rotational speed increased to 3.4 rotations per second. Although a large variance was observed within each dataset, a general trend can be seen. After the addition of compound, the histogram suggests that the number of cells at the lower rotational velocities (between 1.5 and 4.5 rotations per sec) begins to drop and the number of cells higher than 4.5 rotations per second begins to increase. The increase in rotational speed after compound treatment suggests an increase in PMF.

Conclusions

Of nearly 8,000 compounds screened with our *in vivo* assay, only one compound, designated as 67-J8, was successful in preventing SAR endolysin-mediated lysis. However, when the compound was tested for inhibition of R^{933W} activity against *M. lysodeikticus* and *E. coli* cell walls, no inhibition was detected over the concentration range examined. 67-J8 was shown to increase the average rotational velocity of cells suggesting that it increases cellular PMF. SAR domains are maintained as meta-stable transmembrane domains within the context of an energized bilayer until PMF depletion by holin.

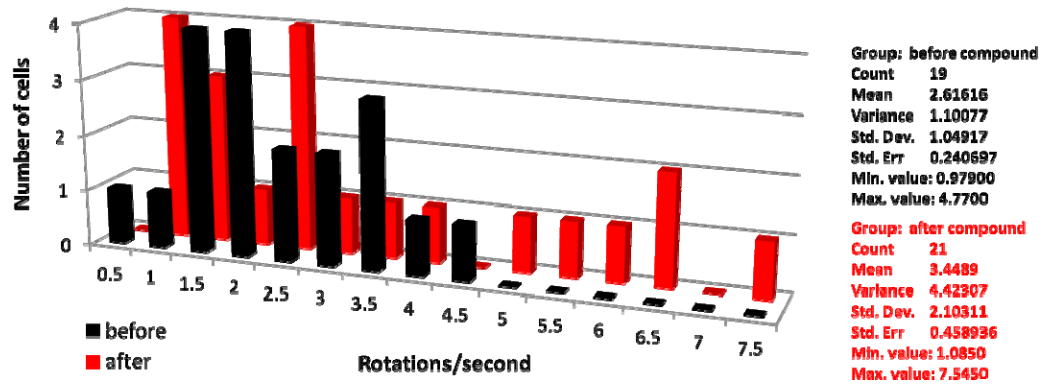

FIG. 5.9. 67-J8 increases proton motive force. Cells were tethered to a glass slide and rotational velocity of individual cells was calculated before and after compound treatment. The number of cells displaying the indicated rotational velocity before compound (black) and after compound (red) is plotted as a combined, overlaid histogram. The statistics for each data set are given.

Membrane poisons such as DNP and cyanide mimic the depolarization effect of the holin and have been shown to prematurely release SAR endolysins (74, 133, 159). In the absence of holin, SAR endolysins spontaneously release from the membrane resulting in gradual lysis. If the PMF is stabilized or increased, presumably, the SAR domain would be retained in the membrane longer. The increase in PMF accounts for the effectiveness of 67-J8 against other SAR endolysins, regardless of their regulation style. The compound has little to no effect on wild type cell growth and morphology (Fig. 5.7, 5.6C), but shows a growth defect in cells expressing SAR endolysins.

How the increase in PMF relates to the observed growth defect is not clear. Interestingly, deletion of the *E. coli atp* operon causes a 20% increase in PMF and is coupled with a

growth defect (65). The mode by which 67-J8 increases the PMF has yet to be established.

CHAPTER VI

SUMMARY AND FUTURE DIRECTIONS

The first SAR endolysin evidence was published in 2004 (159). Before this relatively recent discovery, it was assumed that all dsDNA phages of Gram-negative hosts used the classic lysis system typified by phage lambda. This system requires a holin to form large, non-specific lesions in the membrane. The lesions allow passage of the soluble, cytoplasmic endolysin into the periplasm, which then leads to degradation of the peptidoglycan. In the absence of holin function, the endolysin would be trapped in the cytoplasm and the intracellular infection process would continue indefinitely, with no lysis occurring. Thus the common view was that the endolysin is merely a reporter for holin activity, and the latter supported the key step in lysis. Thus although the T4 lysozyme was arguably one of the most studied enzymes, there was little focus on its role in lysis. Instead, it was mainly a model for protein folding and stability (7). The notion that all dsDNA phages use the classic holin-endolysin paradigm began to give way with the discovery that some Gram-positive phages had secreted endolysins that could lyse cells in a holin independent manner (118). However, in the absence of a facile experimental system with genetics in place, this exception to the holin-endolysin paradigm received little attention.

SAR endolysins

The new lysis paradigm for dsDNA phages of Gram-negative hosts began to emerge through the study of P1 Lyz of the lambdoid phage P1 (63). The work of Dr. Min Xu and Dr. Arockiasamy Arulandu and Stephanie Swanson of the Sacchettini lab showed that induction of a plasmid expressing Lyz lysed cells without a holin (159). P1 Lyz was shown to contain a putative N-terminal secretion signal. A sub-cellular fractionation, however, showed that membrane bound and periplasmic Lyz was the same apparent molecular mass. This suggested that the endolysin maintains the secretion signal that then anchors the protein to the cytoplasmic membrane, and then allows for its release, thus this domain was termed a Signal Anchor Release (SAR) domain (159). Since the protein is produced before the production of the first phage and its catalytic domain was present in the periplasm, it was critical that there be a regulatory system for preventing premature lysis. It was determined through genetic, biochemical, and structural analysis that in its membrane bound, inactive form, P1 Lyz is inactivated by a disulfide bond involving its catalytic Cys (158). The inactivating disulfide is broken by a Cys in the SAR domain after its release from the membrane.

Shifting the SAR endolysin paradigm

The work in this dissertation has followed from that work and has broadened the field in many ways. First, it has shown that a sub-class of disulfide bond regulated SAR endolysins exists. This sub-class is typified by Lyz^{103} from the *Erwinia amylovora* phage ERA103. In its membrane-bound, inactive form, Lyz^{103} was shown to have a disulfide

bond between Cys residues flanking the catalytic Glu residue (74). Although we do not have structural information, we can posit that the disulfide bond effectively acts as a 'cage' to sequester the catalytic Glu. The activity of Lyz^{103} was shown to be dependent on a Cys in the SAR domain. Consequently, we hypothesize that after SAR domain release from the membrane, a disulfide bond isomerization, mediated by the Cys in the SAR domain breaks the cage, activating the enzyme. An inference from bioinformatic analysis is that covalent inactivation is required for SAR endolysins whose catalytic residues lie farther from the SAR domain as these enzymes are capable of reaching the peptidoglycan while tethered to the membrane.

Second, this work has demonstrated, in collaboration with Quingan Sun of the Sacchettini Lab, that a second class of SAR endolysins, typified by R^{21}, exists (133). R^{21} was shown to be regulated by dynamic membrane topology in which the SAR domain actively participates in the proper folding and creation of the active site. Unlike P1 Lyz and Lyz^{103}, whose SAR domains contribute only a sulfhydryl, R^{21}'s SAR domain was shown to be essential for activity. Further, at least among sequenced phages, my analysis indicates that members of the R^{21} class compose the majority (80%) of SAR endolysins.

Moreover, SAR endolysins represent 25% of all endolysins identified in the genomes of phages of Gram-negative hosts. This work also revealed some SAR endolysins that are not homologous to T4 E, a true lysozyme. Before this dissertation, it was thought that all SAR endolysins were homologs of E. E has an E-5x-D-8x-T catalytic triad and uses a hydrolysis mechanism. Bcep22 gp*79*, a glycoside hydrolase family 24 protein, and BcepMu gp*22*, a transglycosylase, serve as prototypes of new

SAR endolysins. Additionally, phylogenetic analysis has allowed us to develop a hypothesis regarding the evolution of dsDNA phage lysis of Gram-negative hosts. Since the SAR endolysin can lyse cells in the absence of a holin, it is the simplest, self-contained lysis protein. As such, we have proposed that the earliest lysis systems would have composed only a SAR endolysin. This SAR endolysin-only system would have developed a more fine tuned system by acquiring a holin to allow synchronous release of the SAR endolysin. Progression to the classic lysis system would require the loss of the SAR domain and development of a large-hole forming holin for cytoplasmic endolysin release into the periplasm.

Future directions

Several questions regarding SAR endolysins remain. These questions involve the nature of the SAR domain itself. Presumably, a SAR domain supports release from the membrane due to the over-representation of weakly hydrophobic amino acids such as Ala, Gly, Cys, and Thr. A transmembrane domain (TMD) on the other hand contains strong hydrophobic amino acids such as Leu, Ile, and Val. It is thought that the hydrophobic nature of a TMD allows for a thermodynamically favorable transfer of the sequence from the translocon channel into the lipid membrane (154). A threshold of hydropathy between TMDs and SAR domains has not yet been established. The research in this dissertation shows that a SAR domain can be converted into a conventional TMD by titration of hydrophobic residues. This feature was shown to be SAR domain specific, i.e. P1 Lyz requires three Gly to Leu mutations (74) while Lyz^{103} and R^{21} (133) require

two Gly to Leu mutations. This suggests that the ability of a SAR domain to release from the membrane may be a function of its overall weak hydrophobicity.

SAR domains share characteristics of both TMDs and secreted signal sequence containing proteins. A SAR domain is hydrophobic enough to be recognized by the Sec translocon and partitioned into the membrane, although is not hydrophobic enough to stay there. A remaining question pertains to SAR domain transport to the membrane. TMDs are recognized and transported to the membrane in a Signal Recognition Particle (SRP) dependent pathway (85). This pathway occurs co-translationally. The pre-protein's signal peptide is recognized by the SRP as it emerges from the translating ribosome. The SRP guides the pre-protein to the SecYEG translocon and is then translocated by SecA. Secreted proteins are not as hydrophobic as TMDs and do not need co-translational secretion. These proteins use the SecB pathway. It is SRP-independent, occurs post-translationally, and requires a chaperone, SecB. After the pre-protein is released from the ribosome, its signal sequence is bound by SecB. SecB may function to protect the pre-protein from aggregating due to the hydrophobic signal sequence before it is translocated. SecB guides the pre-protein to the SecYEG translocon where it is translocated (142). Expression of a SAR endolysin in SRP (33) or SecB (72) depletion strains may allude to which pathway is used.

Another question that has yet to be answered is whether or not the SAR endolysin requires host factors for membrane release and subsequent refolding. Several chaperones function in the periplasm and any of them could play a role in the removal of the SAR domain from the membrane or refolding the protein to accommodate the SAR

domain. This may be a possibility as large conformational changes occur during P1 Lyz and release and refolding. Requirement of host factors can be determined using a genetic approach by analyzing lysis defective host mutants obtained by EMS mutagenesis. The possibility exists however, that the SAR domain requires no additional factor to release from the membrane or fold properly. Support for this comes from the immediacy by which a SAR domain leaves the membrane and is activated after the addition of energy poisons such as DNP. If this is the case, then it is possible that the amount of time a SAR domain spends in the membrane before releasing is directly proportional to its overall hydrophobicity.

Concluding remarks

Overall, this dissertation has shifted the paradigm of the paradigm-shifting SAR endolysin. The SAR endolysin forced us to change how we look at endolysins. This work presents alternate forms of SAR endolysin regulation and displays the diversity of SAR endolysin mechanisms. It has also shown the prevalence of the elegant SAR endolysin among dsDNA phages of Gram-negative hosts. Further study of SAR endolysins may even reshape how we think of protein folding, stability, and dynamics.

REFERENCES

1. **Abedon, S. T., T. D. Herschler, and D. Stopar.** 2001. Bacteriophage latent-period evolution as a response to resource availability. Appl. Environ. Microbiol. **67:**4233-41.

2. **Adams, P. D., R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, and T. C. Terwilliger.** 2002. PHENIX: building new software for automated crystallographic structure determination. Acta. Crystallogr. D: Biol. Crystallogr. **58:**1948-54.

3. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389-402.

4. **Angly, F., M. Youle, B. Nosrat, S. Srinagesh, B. Rodriguez-Brito, P. McNairnie, G. Deyanat-Yazdi, M. Breitbart, and F. Rohwer.** 2009. Genomic analysis of multiple Roseophage SIO1 strains. Environ. Microbiol. **11:**2863-73.

5. **Arbeloa, A., C. V. Oates, O. Marches, E. L. Hartland, and G. Frankel.** 2011. Enteropathogenic and enterohemorrhagic *Escherichia coli* type III secretion effector EspV induces radical morphological changes in eukaryotic cells. Infect. Immun. **79:**1067-76.

6. **Asakura, M., A. Hinenoya, M. S. Alam, K. Shima, S. H. Zahid, L. Shi, N. Sugimoto, A. N. Ghosh, T. Ramamurthy, S. M. Faruque, G. B. Nair, and S. Yamasaki.** 2007. An inducible lambdoid prophage encoding cytolethal distending toxin (Cdt-I) and a type III effector protein in enteropathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. U S A **104:**14483-8.

7. **Baase, W. A., L. Liu, D. E. Tronrud, and B. W. Matthews.** 2010. Lessons from the lysozyme of phage T4. Protein Sci. **19:**631-41.

8. **Bardwell, J. C., K. McGovern, and J. Beckwith.** 1991. Identification of a protein required for disulfide bond formation *in vivo*. Cell **67:**581-9.

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

10. **Bell, J. A., K. P. Wilson, X. J. Zhang, H. R. Faber, H. Nicholson, and B. W. Matthews.** 1991. Comparison of the crystal structure of bacteriophage T4 lysozyme at low, medium, and high ionic strengths. Proteins **10:**10-21.

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

12. **Berry, J., E. J. Summer, D. K. Struck, and R. Young.** 2008. The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. Mol. Microbiol. **70:**341-51.

13. **Bienkowska-Szewczyk, K., B. Lipinska, and A. Taylor.** 1981. The *R* gene product of bacteriophage lambda is the murein transglycosylase. Mol. Gen. Genet. **184:**111-4.

14. **Black, L. W., and D. S. Hogness.** 1969. The lysozyme of bacteriophage lambda. I. Purification and molecular weight. J. Biol. Chem. **244:**1968-75.

15. **Blackburn, N. T., and A. J. Clarke.** 2001. Identification of four families of peptidoglycan lytic transglycosylases. J. Mol. Evol. **52:**78-84.

16. **Bläsi, U., C. Y. Chang, M. T. Zagotta, K. B. Nam, and R. Young.** 1990. The lethal lambda *S* gene encodes its own inhibitor. EMBO J **9:**981-9.

17. **Bläsi, U., K. Nam, D. Hartz, L. Gold, and R. Young.** 1989. Dual translational initiation sites control function of the lambda *S* gene. EMBO J **8:**3501-10.

18. **Bläsi, U., and R. Young.** 1996. Two beginnings for a single purpose: the dual-start holins in the regulation of phage lysis. Mol. Microbiol. **21:**675-82.

19. **Bochtler, M., S. G. Odintsov, M. Marcyjaniak, and I. Sabala.** 2004. Similar active sites in lysostaphins and D-Ala-D-Ala metallopeptidases. Protein Sci. **13:**854-61.

20. **Bonovich, M. T., and R. Young.** 1991. Dual start motif in two lambdoid *S* genes unrelated to lambda *S*. J. Bacteriol. **173:**2897-905.

21. **Braid, M. D., J. L. Silhavy, C. L. Kitts, R. J. Cano, and M. M. Howe.** 2004. Complete genomic sequence of bacteriophage B3, a Mu-like phage of *Pseudomonas aeruginosa*. J. Bacteriol. **186:**6560-74.

22. **Braun, V.** 1975. Covalent lipoprotein from the outer membrane of *Escherichia coli*. Biochim. Biophys. Acta **415:**335-77.

23. **Briers, Y., G. Volckaert, A. Cornelissen, S. Lagaert, C. W. Michiels, K. Hertveldt, and R. Lavigne.** 2007. Muralytic activity and modular structure of the endolysins of *Pseudomonas aeruginosa* bacteriophages phiKZ and EL. Mol. Microbiol. **65:**1334-44.

24. **Broome-Smith, J. K., and B. G. Spratt.** 1986. A vector for the construction of translational fusions to TEM beta-lactamase and the analysis of protein export signals and membrane protein topology. Gene **49:**341-349.

25. **Canchaya, C., G. Fournous, S. Chibani-Chennoufi, M. L. Dillmann, and H. Brussow.** 2003. Phage as agents of lateral gene transfer. Curr. Opin. Microbiol. **6:**417-24.

26. **Carr, P. D., D. Verger, A. R. Ashton, and D. L. Ollis.** 1999. Chloroplast NADP-malate dehydrogenase: structural basis of light-dependent regulation of activity by thiol oxidation and reduction. Structure **7:**461-75.

27. **Ceyssens, P. J., A. Brabban, L. Rogge, M. S. Lewis, D. Pickard, D. Goulding, G. Dougan, J. P. Noben, A. Kropinski, E. Kutter, and R. Lavigne.** 2010. Molecular and physiological analysis of three *Pseudomonas aeruginosa* phages belonging to the "N4-like viruses". Virology **405:**26-30.

28. **Chang, C. Y., K. Nam, U. Bläsi, and R. Young.** 1993. Synthesis of two bacteriophage lambda S proteins in an *in vivo* system. Gene **133:**9-16.

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

30. **Cheng, X., X. Zhang, J. W. Pflugrath, and F. W. Studier.** 1994. The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. Proc. Natl. Acad. Sci. U S A **91:**4034-8.

31. **Clarke, V. A., N. Platt, and T. D. Butters.** 1995. Cloning and expression of the beta-N-acetylglucosaminidase gene from Streptococcus pneumoniae. Generation of truncated enzymes with modified aglycon specificity. J. Biol. Chem. **270:**8805-14.

32. **Cramer, W. A., and S. K. Phillips.** 1970. Response of an *Escherichia coli*-bound fluorescent probe to colicin E1. J. Bacteriol. **104:**819-25.

33. **de Gier, J. W., P. Mansournia, Q. A. Valent, G. J. Phillips, J. Luirink, and G. von Heijne.** 1996. Assembly of a cytoplasmic membrane protein in *Escherichia coli* is dependent on the signal recognition particle. FEBS Lett. **399:**307-9.

34. **DeMartini, M., S. Halegoua, and M. Inouye.** 1975. Lysozymes from bacteriophages T3 and T5. J. Virol. **16:**459-61.

35. **Dereeper, A., V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, J. F. Dufayard, S. Guindon, V. Lefort, M. Lescot, J. M. Claverie, and O. Gascuel.** 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. **36:**W465-9.

36. **Desiere, F., W. M. McShan, D. van Sinderen, J. J. Ferretti, and H. Brussow.** 2001. Comparative genomics reveals close genetic relationships between phages from dairy bacteria and pathogenic *Streptococci*: evolutionary implications for prophage-host interactions. Virology **288:**325-41.

37. **Dewey, J. S., C. G. Savva, R. L. White, S. Vitha, A. Holzenburg, and R. Young.** 2010. Micron-scale holes terminate the phage infection cycle. Proc. Natl. Acad. Sci. U S A **107:**2219-23.

38. **Dobbins, A. T., M. George, Jr., D. A. Basham, M. E. Ford, J. M. Houtz, M. L. Pedulla, J. G. Lawrence, G. F. Hatfull, and R. W. Hendrix.** 2004. Complete genomic sequence of the virulent *Salmonella* bacteriophage SP6. J. Bacteriol. **186:**1933-44.

39. **Emsley, P., and K. Cowtan.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D: Biol Crystallogr. **60:**2126-32.

40. **Engel, H., A. J. Smink, L. van Wijngaarden, and W. Keck.** 1992. Murein-metabolizing enzymes from *Escherichia coli*: existence of a second lytic transglycosylase. J. Bacteriol. **174:**6394-403.

41. **Engelman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. **15:**321-53.

42. **Fischetti, V. A.** 2010. Bacteriophage endolysins: a novel anti-infective to control Gram-positive pathogens. Int. J. Med. Microbiol. **300:**357-62.

43. **Fukatsu, T., N. Nikoh, R. Kawai, and R. Koga.** 2000. The secondary endosymbiotic bacterium of the pea aphid *Acyrthosiphon pisum* (Insecta: homoptera). Appl. Environ. Microbiol. **66:**2748-58.

44. **Fung, D. C., and H. C. Berg.** 1995. Powering the flagellar motor of *Escherichia coli* with an external voltage source. Nature **375:**809-12.

45. **Furste, J. P., W. Pansegrau, R. Frank, H. Blocker, P. Scholz, M. Bagdasarian, and E. Lanka.** 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range tacP expression vector. Gene **48:**119-31.

46. **Gan, L., S. Chen, and G. J. Jensen.** 2008. Molecular organization of Gram-negative peptidoglycan. Proc. Natl. Acad. Sci. U S A **105:**18953-7.

47. **Garcia, E.** 1996. On the misuse of the term 'lysozyme'. Mol. Microbiol. **21:**885.

48. **Garland, T., Jr., and R. Diaz-Uriarte.** 1999. Polytomies and phylogenetically independent contrasts: examination of the bounded degrees of freedom approach. Syst. Biol. **48:**547-58.

49. **Ghuysen, J. M., J. Lamotte-Brasseur, B. Joris, and G. D. Shockman.** 1994. Binding site-shaped repeated sequences of bacterial wall peptidoglycan hydrolases. FEBS Lett. **342:**23-8.

50. **Graschopf, A., and U. Bläsi.** 1999. Functional assembly of the lambda S holin requires periplasmic localization of its N-terminus. Arch. Microbiol. **172:**31-9.

51. **Gründling, A., U. Bläsi, and R. Young.** 2000. Genetic and biochemical analysis of dimer and oligomer interactions of the lambda S holin. J. Bacteriol. **182:**6082-90.

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

53. **Gründling, A., M. D. Manson, and R. Young.** 2001. Holins kill without warning. Proc. Natl. Acad. Sci. U.S.A. **98:**9348-9352.

54. **Gründling, A., D. L. Smith, U. Bläsi, and R. Young.** 2000. Dimerization between the holin and holin inhibitor of phage lambda. J. Bacteriol. **182:**6075-81.

55. **Hahn, M., M. Hennig, B. Schlesier, and W. Hohne.** 2000. Structure of jack bean chitinase. Acta Crystallogr. D: Biol Crystallogr. **56:**1096-9.

56. **Halling, C., R. Calendar, G. E. Christie, E. C. Dale, G. Deho, S. Finkel, J. Flensburg, D. Ghisotti, M. L. Kahn, K. B. Lane, and et al.** 1990. DNA sequence of satellite bacteriophage P4. Nucleic Acids Res. **18:**1649.

57. **Hardies, S. C., A. M. Comeau, P. Serwer, and C. A. Suttle.** 2003. The complete sequence of marine bacteriophage VpV262 infecting *Vibrio parahaemolyticus* indicates that an ancestral component of a T7 viral supergroup is widespread in the marine environment. Virology **310:**359-371.

58. **Hardy, L. W., and A. R. Poteete.** 1991. Reexamination of the role of Asp20 in catalysis by bacteriophage T4 lysozyme. Biochemistry **30:**9457-9463.

59. **Heineman, R. H., I. J. Molineux, and J. J. Bull.** 2005. Evolutionary robustness of an optimal phenotype: re-evolution of lysis in a bacteriophage deleted for its lysin gene. J. Mol. Evol. **61:**181-91.

60. **Heo, Y. J., I. Y. Chung, K. B. Choi, G. W. Lau, and Y. H. Cho.** 2007. Genome sequence comparison and superinfection between two related *Pseudomonas aeruginosa* phages, D3112 and MP22. Microbiology **153:**2885-95.

61. **Herold, S., J. Siebert, A. Huber, and H. Schmidt.** 2005. Global expression of prophage genes in *Escherichia coli* O157:H7 strain EDL933 in response to norfloxacin. Antimicrob. Agents Chemother. **49:**931-44.

62. **Holm, L., and C. Sander.** 1994. Structural similarity of plant chitinase and lysozymes from animals and phage. An evolutionary connection. FEBS Lett. **340:**129-32.

63. **Iida, S., and W. Arber.** 1977. Plaque forming specialized transducing phage P1: isolation of P1CmSmSu, a precursor of P1Cm. Mol. Gen. Genet. **153:**259-69.

64. **Imada, M., and A. Tsugita.** 1971. Amino-acid sequence of lambda phage endolysin. Nature New Biology **233:**230-231.

65. **Jensen, P. R., and O. Michelsen.** 1992. Carbon and energy metabolism of *atp* mutants of *Escherichia coli*. J. Bacteriol. **174:**7635-41.

66. **Jensen, S. E., and J. N. Campbell.** 1976. Amidase activity involved in peptidoglycan biosynthesis in membranes of *Micrococcus luteus* (sodonensis). J. Bacteriol. **127:**319-26.

67. **Jolliffe, L. K., R. J. Doyle, and U. N. Streips.** 1981. The energized membrane and cellular autolysis in *Bacillus subtilis*. Cell **25:**753-63.

68. **Jones, D. T., W. R. Taylor, and J. M. Thornton.** 1994. A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry **33:**3038-49.

69. **Kim, W. S., and K. Geider.** 2000. Characterization of a viral EPS-depolymerase, a potential tool for control of fire blight. Phytopathol. **90** 1263-1268.

70. **Kim, W. S., H. Salm, and K. Geider.** 2004. Expression of bacteriophage phiEa1h lysozyme in *Escherichia coli* and its activity in growth inhibition of *Erwinia amylovora*. Microbiology **150:**2707-2714.

71. **Komor, E., H. Weber, and W. Tanner.** 1979. Greatly decreased susceptibility of nonmetabolizing cells towards detergents. Proc. Natl. Acad. Sci. U S A **76:**1814-8.

72. **Kumamoto, C. A., and J. Beckwith.** 1985. Evidence for specificity at an early step in protein export in *Escherichia coli*. J. Bacteriol. **163:**267-74.

73. **Kuroki, R., L. H. Weaver, and B. W. Matthews.** 1993. A covalent enzyme-substrate intermediate with saccharide distortion in a mutant T4 lysozyme. Science **262:**2030-3.

74. **Kuty, G. F., M. Xu, D. K. Struck, E. J. Summer, and R. Young.** 2010. Regulation of a phage endolysin by disulfide caging. J. Bacteriol. **192:**5682-7.

75. **Lacombe-Harvey, M. E., T. Fukamizo, J. Gagnon, M. G. Ghinet, N. Dennhart, T. Letzel, and R. Brzezinski.** 2009. Accessory active site residues of *Streptomyces* sp. N174 chitosanase: variations on a common theme in the lysozyme superfamily. FEBS J. **276:**857-69.

76. **Lei, S. P., H. C. Lin, S. S. Wang, J. Callaway, and G. Wilcox.** 1987. Characterization of the *Erwinia carotovora pelB* gene and its product pectate lyase. J. Bacteriol. **169:**4379-83.

77. **Li, S. C., N. K. Goto, K. A. Williams, and C. M. Deber.** 1996. Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. Proc. Natl. Acad. Sci. U S A **93:**6676-81.

78. **Lingwood, C. A.** 2003. Shiga toxin receptor glycolipid binding. Pathology and utility. Methods Mol. Med. **73:**165-86.

79. **Little, J. W.** 1984. Autodigestion of lexA and phage lambda repressors. Proc. Natl. Acad. Sci. U S A **81:**1375-9.

80. **Liu, L. P., and C. M. Deber.** 1998. Guidelines for membrane protein engineering derived from *de novo* designed model peptides. Biopolymers **47:**41-62.

81. **Livny, J., and D. I. Friedman.** 2004. Characterizing spontaneous induction of *Stx* encoding phages using a selectable reporter system. Mol. Microbiol. **51:**1691-1704.

82. **Loeffler, J. M., D. Nelson, and V. A. Fischetti.** 2001. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. Science **294:**2170-2.

83. **Loessner, M. J., K. Kramer, F. Ebel, and S. Scherer.** 2002. C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. Mol. Microbiol. **44:**335-49.

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

85. **Luirink, J., G. von Heijne, E. Houben, and J. W. de Gier.** 2005. Biogenesis of inner membrane proteins in *Escherichia coli*. Annu. Rev. Microbiol. **59:**329-55.

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

87. **Makino, K., K. Yokoyama, Y. Kubota, C. H. Yutsudo, S. Kimura, K. Kurokawa, K. Ishii, M. Hattori, I. Tatsuno, H. Abe, T. Iida, K. Yamamoto, M. Onishi, T. Hayashi, T. Yasunaga, T. Honda, C. Sasakawa, and H. Shinagawa.** 1999. Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak. Genes Genet. Syst. **74:**227-39.

88. **Mannisto, R. H., H. M. Kivela, L. Paulin, D. H. Bamford, and J. K. Bamford.** 1999. The complete genome sequence of PM2, the first lipid-containing bacterial virus to be isolated. Virology **262:**355-63.

89. **Marcyjaniak, M., S. G. Odintsov, I. Sabala, and M. Bochtler.** 2004. Peptidoglycan amidase MepA is a LAS metallopeptidase. J. Biol. Chem. **279:**43982-9.

90. **Matsumura, M., and B. W. Matthews.** 1989. Control of enzyme activity by an engineered disulfide bond. Science **243:**792-794.

91. **Meroueh, S. O., K. Z. Bencze, D. Hesek, M. Lee, J. F. Fisher, T. L. Stemmler, and S. Mobashery.** 2006. Three-dimensional structure of the bacterial cell wall peptidoglycan. Proc. Natl. Acad. Sci. U S A **103:**4404-9.

92. **Mikoulinskaia, G. V., I. V. Odinokova, A. A. Zimin, V. Y. Lysanskaya, S. A. Feofanov, and O. A. Stepnaya.** 2009. Identification and characterization of the metal ion-dependent L-alanoyl-D-glutamate peptidase encoded by bacteriophage T5. FEBS J. **276:**7329-42.

93. **Miller, J. H.** 1992. A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria. Cold Spring Harbor Press, NY.

94. **Moffatt, B. A., and F. W. Studier.** 1987. T7 lysozyme inhibits transcription by T7 RNA polymerase. Cell **49:**221-7.

95. **Monzingo, A. F., E. M. Marcotte, P. J. Hart, and J. D. Robertus.** 1996. Chitinases, chitosanases, and lysozymes can be divided into procaryotic and eucaryotic families sharing a conserved core. Nat. Struct. Biol. **3:**133-40.

96. **Mooers, B. H., and B. W. Matthews.** 2006. Extension to 2268 atoms of direct methods in the *ab initio* determination of the unknown structure of bacteriophage P22 lysozyme. Acta Crystallogr. D: Biol. Crystallogr. **62:**165-76.

97. **Morreale, G., H. Lanckriet, J. C. Miller, and A. P. Middelberg.** 2003. Continuous processing of fusion protein expressed as an *Escherichia coli* inclusion body. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. **786:**237-46.

98. **Navarre, W. W., H. Ton-That, K. F. Faull, and O. Schneewind.** 1999. Multiple enzymatic activities of the murein hydrolase from *staphylococcal* phage phi11. Identification of a D-alanyl-glycine endopeptidase activity. J. Biol. Chem. **274:**15847-56.

99. **Nieva-Gomez, D., and R. B. Gennis.** 1977. Affinity of intact *Escherichia coli* for hydrophobic membrane probes is a function of the physiological state of the cells. Proc. Natl. Acad. Sci. U S A **74:**1811-5.

100. **Nilsson, I., A. E. Johnson, and G. von Heijne.** 2003. How hydrophobic is alanine? J. Biol. Chem. **278:**29389-93.

101. **O'Brien, A. D., L. R. Marques, C. F. Kerry, J. W. Newland, and R. K. Holmes.** 1989. Shiga-like toxin converting phage of enterohemorrhagic *Escherichia coli* strain 933. Microb. Pathog. **6:**381-90.

102. **Ochman, H., J. G. Lawrence, and E. A. Groisman.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405:**299-304.

103. **Ochoa, T. J., J. Chen, C. M. Walker, E. Gonzales, and T. G. Cleary.** 2007. Rifaximin does not induce toxin production or phage-mediated lysis of Shiga toxin-producing *Escherichia coli*. Antimicrob. Agents Chemother. **51:**2837-41.

104. **Oh, C. S., and S. V. Beer.** 2005. Molecular genetics of *Erwinia amylovora* involved in the development of fire blight. FEMS Microbiol. Lett. **253:**185-92.

105. **Oliver, D. B., R. J. Cabelli, K. M. Dolan, and G. P. Jarosik.** 1990. Azide-resistant mutants of *Escherichia coli* alter the SecA protein, an azide-sensitive component of the protein export machinery. Proc. Natl. Acad. Sci. U S A **87:**8227-31.

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

107. **Pang, T., C. G. Savva, K. G. Fleming, D. K. Struck, and R. Young.** 2009. Structure of the lethal phage pinhole. Proc. Natl. Acad. Sci. U S A **106:**18966-71.

108. **Park, T., D. K. Struck, C. A. Dankenbring, and R. Young.** 2007. The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. J. Bacteriol. **189:**9135-9.

109. **Park, T., D. K. Struck, J. F. Deaton, and R. Young.** 2006. Topological dynamics of holins in programmed bacterial lysis. Proc. Natl. Acad. Sci. U S A **103:**19713-19718.

110. **Parkinson, J. S., and S. E. Houts.** 1982. Isolation and behavior of *Escherichia coli* deletion mutants lacking chemotaxis functions. J. Bacteriol. **151:**106-13.

111. **Payne, K., Q. Sun, J. Sacchettini, and G. F. Hatfull.** 2009. Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. Mol. Microbiol. **73:**367-81.

112. **Perry, L. L., P. SanMiguel, U. Minocha, A. I. Terekhov, M. L. Shroyer, L. A. Farris, N. Bright, B. L. Reuhs, and B. M. Applegate.** 2009. Sequence analysis of *Escherichia coli* O157:H7 bacteriophage PhiV10 and identification of a phage-encoded immunity protein that modifies the O157 antigen. FEMS Microbiol. Lett. **292:**182-6.

113. **Plunkett, G., 3rd, D. J. Rose, T. J. Durfee, and F. R. Blattner.** 1999. Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. J. Bacteriol. **181:**1767-78.

114. **Quevillon, E., V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler, and R. Lopez.** 2005. InterProScan: protein domains identifier. Nucleic Acids Res. **33:**W116-20.

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

116. **Rasko, D. A., D. R. Webster, J. W. Sahl, A. Bashir, N. Boisen, F. Scheutz, E. E. Paxinos, R. Sebra, C. S. Chin, D. Iliopoulos, A. Klammer, P. Peluso, L. Lee, A. O. Kislyuk, J. Bullard, A. Kasarskis, S. Wang, J. Eid, D. Rank, J. C. Redman, S. R. Steyert, J. Frimodt-Moller, C. Struve, A. M. Petersen, K. A. Krogfelt, J. P. Nataro, E. E. Schadt, and M. K. Waldor.** 2011. Origins of the *Escherichia coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. N. Engl. J. Med. **365:**709-17.

117. **Rennell, D., S. E. Bouvier, L. W. Hardy, and A. R. Poteete.** 1991. Systematic mutation of bacteriophage T4 lysozyme. J. Mol. Biol. **222:**67-88.

118. **Sao-Jose, C., R. Parreira, G. Vieira, and M. A. Santos.** 2000. The N-terminal region of the *Oenococcus oeni* bacteriophage fOg44 lysin behaves as a bona fide signal peptide in *Escherichia coli* and as a cis-inhibitory element, preventing lytic activity on oenococcal cells. J. Bacteriol. **182:**5823-5831.

119. **Sato, T., T. Shimizu, M. Watarai, M. Kobayashi, S. Kano, T. Hamabata, Y. Takeda, and S. Yamasaki.** 2003. Genome analysis of a novel Shiga toxin 1 (Stx1)-converting phage which is closely related to Stx2-converting phages but not to other Stx1-converting phages. J. Bacteriol. **185:**3966-71.

120. **Scheurwater, E., C. W. Reid, and A. J. Clarke.** 2008. Lytic transglycosylases: bacterial space-making autolysins. Int. J. Biochem. Cell Biol. **40:**586-91.

121. **Scheurwater, E. M., and A. J. Clarke.** 2008. The C-terminal domain of *Escherichia coli* YfhD functions as a lytic transglycosylase. J. Biol. Chem. **283:**8363-73.

122. **Schleifer, K. H., and O. Kandler.** 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol. Rev. **36:**407-77.

123. **Scholl, D., J. Kieleczawa, P. Kemp, J. Rush, C. C. Richardson, C. Merril, S. Adhya, and I. J. Molineux.** 2004. Genomic analysis of bacteriophages SP6 and K1-5, an estranged subgroup of the T7 supergroup. J. Mol. Biol. **335:**1151-71.

124. **Shimizu, T., Y. Ohta, and M. Noda.** 2009. Shiga toxin 2 is specifically released from bacterial cells by two different mechanisms. Infect. Immun. **77:**2813-23.

125. **Silverman, M., and M. Simon.** 1974. Flagellar rotation and the mechanism of bacterial motility. Nature **249:**73-4.

126. **Smith, D. L., C. Y. Chang, and R. Young.** 1998. The λ holin accumulates beyond the lethal triggering concentration under hyper-expression conditions. Gene Expression **7:**39-52.

127. **Smith, D. L., D. K. Struck, J. M. Scholtz, and R. Young.** 1998. Purification and biochemical characterization of the lambda holin. J. Bacteriol. **180:**2531-40.

128. **Smith, D. L., and R. Young.** 1998. Oligohistidine tag mutagenesis of the lambda holin gene. J. Bacteriol. **180:**4199-211.

129. **Steiner, M., W. Lubitz, and U. Bläsi.** 1993. The missing link in phage lysis of Gram-positive bacteria: gene 14 of *Bacillus subtilis* phage phi 29 encodes the functional homolog of lambda S protein. J. Bacteriol. **175:**1038-42.

130. **Stojkovic, E. A., and L. B. Rothman-Denes.** 2007. Coliphage N4 N-acetylmuramidase defines a new family of murein hydrolases. J. Mol. Biol. **366:**406-19.

131. **Strynadka, N. C., and M. N. James.** 1996. Lysozyme: a model enzyme in protein crystallography. EXS **75:**185-222.

132. **Summer, E. J., J. Berry, T. A. Tran, L. Niu, D. K. Struck, and R. Young.** 2007. Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts. J. Mol. Biol. **373:**1098-112.

133. **Sun, Q., G. F. Kuty, A. Arockiasamy, M. Xu, R. Young, and J. C. Sacchettini.** 2009. Regulation of a muralytic enzyme by dynamic membrane topology. Nat. Struct. Mol. Biol. **16:**1192-4.

134. **Suttle, C. A.** 2007. Marine viruses--major players in the global ecosystem. Nat. Rev. Microbiol. **5:**801-12.

135. **Takac, M., A. Witte, and U. Bläsi.** 2005. Functional analysis of the lysis genes of *Staphylococcus aureus* phage P68 in *Escherichia coli*. Microbiology **151:**2331-42.

136. **Takahashi, K., K. Narita, Y. Kato, T. Sugiyama, N. Koide, T. Yoshida, and T. Yokochi.** 1997. Low-level release of Shiga-like toxin (verocytotoxin) and endotoxin from enterohemorrhagic *Escherichia coli* treated with imipenem. Antimicrob. Agents Chemother. **41:**2295-6.

137. **Tanford, C.** 1973. The hydrophobic effect: formation of micelles and biological membranes. Wiley, New York,.

138. **Taylor, A.** 1971. Endopeptidase activity of phage lamba-endolysin. Nat. New Biol. **234:**144-5.

139. **Thunnissen, A. M., A. J. Dijkstra, K. H. Kalk, H. J. Rozeboom, H. Engel, W. Keck, and B. W. Dijkstra.** 1994. Doughnut-shaped structure of a bacterial muramidase revealed by X-ray crystallography. Nature **367:**750-3.

140. **Thunnissen, A. M., N. W. Isaacs, and B. W. Dijkstra.** 1995. The catalytic domain of a bacterial lytic transglycosylase defines a novel class of lysozymes. Proteins **22:**245-58.

141. **Torti, S. V., and J. T. Park.** 1976. Lipoprotein of Gram-negative bacteria is essential for growth and division. Nature **263:**323-6.

142. **Valent, Q. A., P. A. Scotti, S. High, J. W. de Gier, G. von Heijne, G. Lentzen, W. Wintermeyer, B. Oudega, and J. Luirink.** 1998. The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. EMBO J. **17:**2504-12.

143. **Vollmer, W., D. Blanot, and M. A. de Pedro.** 2008. Peptidoglycan structure and architecture. FEMS Microbiol. Rev. **32:**149-67.

144. **Wagner, P. L., M. N. Neely, X. Zhang, D. W. Acheson, M. K. Waldor, and D. I. Friedman.** 2001. Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain. J. Bacteriol. **183:**2081-5.

145. **Waldor, M. K., and D. I. Friedman.** 2005. Phage regulatory circuits and virulence gene expression. Curr. Opin. Microbiol. **8:**459-65.

146. **Wang, I. N.** 1996. The evolution of phage lysis timing. Evolutionary Ecology 545-558.

147. **Wang, I. N., D. E. Dykhuizen, and L. B. Slobodkin.** 1996. The evolution of phage lysis timing. Evol. Ecol. **10:**545-558.

148. **Wang, I. N., D. L. Smith, and R. Young.** 2000. Holins: the protein clocks of bacteriophage infections. Annu. Rev. Microbiol. **54:**799-825.

149. **Wang, S. L., and W. T. Chang.** 1997. Purification and characterization of two bifunctional chitinases/lysozymes extracellularly produced by *Pseudomonas aeruginosa* K-187 in a shrimp and crab shell powder medium. Appl. Environ. Microbiol. **63:**380-6.

150. **Weaver, L. H., M. G. Grutter, S. J. Remington, T. M. Gray, N. W. Isaacs, and B. W. Matthews.** 1984. Comparison of goose-type, chicken-type, and phage-type lysozymes illustrates the changes that occur in both amino acid sequence and three-dimensional structure during evolution. J. Mol. Evol. **21:**97-111.

151. **White, R., S. Chiba, T. Pang, J. S. Dewey, C. G. Savva, A. Holzenburg, K. Pogliano, and R. Young.** 2011. Holin triggering in real time. Proc. Natl. Acad. Sci. U S A **108:**798-803.

152. **White, R., T. A. Tran, C. A. Dankenbring, J. Deaton, and R. Young.** 2010. The N-terminal transmembrane domain of lambda S is required for holin but not antiholin function. J. Bacteriol. **192:**725-33.

153. **White, R. L.** 2009. What makes the lysis clock tick? A study of the bacteriophage holin. Doctoral dissertation. Texas A&M University, College Station.

154. **White, S. H., and G. von Heijne.** 2004. The machinery of membrane protein assembly. Curr. Opin. Struct. Biol. **14:**397-404.

155. **White, S. H., and W. C. Wimley.** 1999. Membrane protein folding and stability: physical principles. Annu. Rev. Biophys. Biomol. Struct. **28:**319-65.

156. **Wu, J., and J. T. Watson.** 1997. A novel methodology for assignment of disulfide bond pairings in proteins. Protein Science **6:**391-398.

157. **Xu, M.** 2005. Bacteriophage P1: a new paradigm for control of phage lysis. Doctoral dissertation. Texas A&M University, College Station, TX.

158. **Xu, M., A. Arulandu, D. K. Struck, S. Swanson, J. C. Sacchettini, and R. Young.** 2005. Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. Science **307:**113-117.

159. **Xu, M., D. K. Struck, J. Deaton, I. N. Wang, and R. Young.** 2004. A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. Proc. Natl. Acad. Sci. U. S. A. **101:**6415-20.

160. **Yarnell, W. S., and J. W. Roberts.** 1992. The phage lambda gene *Q* transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase. Cell **69:**1181-9.

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

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

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

164. **Zdobnov, E. M., and R. Apweiler.** 2001. InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17:**847-8.

165. **Zegans, M. E., J. C. Wagner, K. C. Cady, D. M. Murphy, J. H. Hammond, and G. A. O'Toole.** 2009. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. J. Bacteriol. **191:**210-9.

166. **Zhang, N., and R. Young.** 1999. Complementation and characterization of the nested *Rz* and *Rz1* reading frames in the genome of bacteriophage lambda. Mol. Gen. Genet. **262:**659-67.

167. **Zhang, X., A. D. McDaniel, L. E. Wolf, G. T. Keusch, M. K. Waldor, and D. W. Acheson.** 2000. Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice. J. Infect. Dis. **181:**664-70.

168. **Zhang, X., and F. W. Studier.** 2004. Multiple roles of T7 RNA polymerase and T7 lysozyme during bacteriophage T7 infection. J. Mol. Biol. **340:**707-30.

169. **Zuber, B., M. Chami, C. Houssin, J. Dubochet, G. Griffiths, and M. Daffe.** 2008. Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. J. Bacteriol. **190:**5672-80.

VITA

Gabriel Faith Kuty received her Bachelor of Science degree in genetics from Texas A&M University in 2005. She entered the doctoral program in the Biochemistry and Biophysics Department at Texas A&M University in 2005 and joined the laboratory of Dr. Ryland Young in April of 2006. She received her Ph.D. in December 2011. Her research focuses on SAR endolysin membrane dynamics, their regulation in phage lysis, and their role in pathogenicity. While pursuing her Ph.D., Ms. Kuty served as the Biochemistry Graduate Association (BGA) representative on the Biochemistry and Biophysics Department Graduate Recruiting and Admissions Committee, and served one term as the Vice President of the BGA. Since then, she has been an active member of the organization. In addition to teaching a recitation course for Biochemistry 410 and assisting in a laboratory course for Biochemistry 432, Ms. Kuty also supervised an undergraduate as part of the Research Experience for Undergraduates (REU) program funded by the National Science Foundation and mentored several graduate students.

Ms. Kuty may be reached at Texas A&M University, MS 2128, Bio/Bio 315, College Station, TX, 77843. Her email address is: gabbykuty@tamu.edu.