Leveraging the ΦX174 Protein Antibiotic to Study MraY Structure, Function, and Regulation

Thesis by
Anna Karen Orta

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
Defended May 20th, 2024

Anna Karen Orta
ORCID: 0000-0002-8526-0383

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Bil Clemons, whose mentorship and guidance throughout graduate school kept me grounded. Bil taught me to always be skeptical and to pursue difficult questions with curiosity and drive. I'm grateful that he did not hesitate to take me on as a student and trusted me with what became an incredible story. My mentors at the University of Texas at El Paso, Professors Ricardo Bernal and Keith Pannell, gifted me with an undergraduate research experience so enriching that it dictated the next several years of my life.

I'm grateful for the Clemons group members before the great merge of 2020, especially Dr. Hyun-Gi Yun and Dr. Nadia Riera, who guided me at the beginning of my graduate journey. I was fortunate to work closely with Nadia for half of my time at Caltech; she became a second advisor to me and an even closer friend. Nadia taught me that your greatest rival is yourself. From vision boards in the basement of Broad to seeing this story published and celebrated, I'll always be grateful that we shared this experience.

Halfway through my time as a graduate student, Bil joined forces with Professor Doug Rees, who soon became an advisor to me and the rest of our group. Doug and Bil have collected an incredible team of scientists, like infinity stones. I was fortunate to be taught cryoEM by Dr. Naima Sharaf, who patiently sat with me for hours on end at the Arctica microscope, teaching me to take experimental wins with excitement and losses with levity. Thank you to the then-graduate students I overlapped with, Dr. Michelle Fry, Dr. Chengcheng Fan, and Dr. Ailiena Maggiolo, for your support and patience as I learned to navigate membrane protein expression, crystallography, and cryoEM. I'm grateful to all the current members of the Rees-Clemons group, particularly Victor Garcia and Alex Barlow, with whom I had the pleasure of sharing most of my graduate journey. Finally, I'm grateful to have learned how to be a mentor through several rotation students and then-undergraduate Helen Brackney, who remained driven and patient as I learned how to teach.

Caltech would not be the institution it is without the administration and staff. Thank you to Allen Lee, Courtney Oaida, Ann Mao, Dr. Songye Chen, Dr. Jens Kaiser, and Welison Floriano. The facilities staff at Broad and Braun labs, and the staff at Red Door and Broad Cafe, who always greeted me with the biggest smiles. To the CCID, especially Yazmin Gonzalez, who, along with Victor, Sal, Eric, and Amanda,

built an incredible community for us Latinos at Caltech.

Thank you to my friends, both before and during graduate school. Thank you to my oldest friends Soraya, Soraida, and Jasmine, and my college friends Evelyn, Leila, Jaime, Aldo, Robert, and Omar. They supported me throughout my bachelor's degree, witnessed me nervously apply to grad school, move to California, and cheered me on through the finish line. Thank you to my Caltech friends, from orientation week to our thesis defenses, it has been so encouraging to see you all succeed, and you served as a great source of inspiration to keep going. I cherish all the late nights, cabin trips, and our brave attempts at making jorts fashionable again. To Catherine, Dan, Dom, and Josh, for making a home away from home with me, and to Victor for keeping me sane inside and outside of the lab. Lastly, I would like to thank Alessandro for all the encouragement, for brainstorming my experiments, and for proofreading everything I wrote. I'm beyond grateful for you.

Finally, and most importantly, thank you to the largest support system I have, my family. You have made sure to be there for all the milestones in my life. This work and everything I've done up to this day is dedicated to my lifelong mentors, Ana De Santiago and Miguel Orta, and my lifelong partners in crime, Miguel Angel and Andres Orta. Tengo presente que este título es el fruto de un sacrificio multigeneracional, el cual tengo el privilegio de representar. No tengo palabras para agradecerles sus esfuerzos, sacrificios, cariño y apoyo desde mi primer día en el kínder hasta mi último día como doctorando. Mis éxitos son un reflejo directo de ustedes como padres y hermanos.

ABSTRACT

The overuse of antibiotics has escalated the prevalence of bacterial resistance to existing treatments, posing a significant threat to global health. This rise in antimicrobial resistance (AMR) has spurred research into innovative therapeutic approaches. Among the most promising strategies is the use of viruses of bacteria for 'phage therapy'. This thesis delves into the interplay between antibacterial resistance and peptidoglycan biosynthesis, highlighting the pivotal role of the membrane protein MraY. We present the first structure of MraY from a pathogenic species, revealing its inhibition by the lysis protein from the bacteriophage ΦX174, protein E. Additionally, we analyze lipidic interactions with MraY, proposing a previously unexplored allosteric feedback mechanism for regulating its enzymatic activity. Building on these insights, we expand the application of protein E to non-native hosts, offering new avenues for the development of targeted antibiotic interventions. This work not only advances our understanding of the structural and functional dynamics of MraY but also paves the way for novel antibacterial strategies.

PUBLISHED CONTENT AND CONTRIBUTIONS

Marmont, L. S., Orta, A. K., Baileeves, B. W., Sychantha, D., Fernández-Galliano, A., Li, Y. E., Greene, N. G., Corey, R. A., Stansfeld, P. J., Clemons Jr, W. M., et al. (2024). Synthesis of lipid-linked precursors of the bacterial cell wall is governed by a feedback control mechanism in *Pseudomonas aeruginosa*. *Nature Microbiology*, 1–13. https://doi.org/10.1038/s41564-024-01603-2.

Contributions: A.K. Orta expressed and purified the protein complex, and performed the cryoEM data collection, processing, and modeling. A.K. Orta participated in preparing the figures pertaining to the cryoEM data, and assisted in editing the manuscript.

Orta, A. K., Riera, N., Li, Y. E., Tanaka, S., Yun, H. G., Klaic, L., & Clemons Jr, W. M. (2023). The mechanism of the phage-encoded protein antibiotic from Φx174. *Science*, *381*(6654), eadg9091. https://doi.org/10.1126/science.adg9091.

Contributions: A.K. Orta expressed, purified the protein complex, participated in the execution of the biochemical experiments, and performed the cryoEM data collection, processing, and modeling. A.K. Orta prepared the figures and participated in writing the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

Antimicrobial resistance (AMR) is a critical problem endangering global health and food security worldwide (Asokan et al., 2019). AMR defines when a microorganism evolves to resist the effects of medications that previously could treat them effectively. Upon exposure to antibiotics, bacteria can develop mechanisms for survival through mutagenesis, slowed growth, or interactions with neighboring microbes (reviewed in Schrader et al., 2023). Development of resistance is accelerated by the misuse of antimicrobials in human medicine, agriculture, and animal husbandry. As a result, common infections and minor injuries, which had been treatable for decades, can once again become life-threatening. The World Health Organization has highlighted AMR as a growing threat, noting that without significant action, we face a future where common infections could become intractable, leading to increased medical costs and mortality (Asokan et al., 2019).

Clinical antibiotics have been developed to target a variety of essential pathways in bacteria. For instance, bacterial ribosomes can be inhibited in several ways that culminate in halting protein synthesis. Examples include tetracyclines, which prevent the attachment of aminoacyl-tRNA to the ribosomal acceptor site, and macrolides, which inhibit peptide elongation (reviewed in (Lin et al., 2018)). Fluoroquinolones act by inhibiting bacterial DNA gyrase or topoisomerase IV, enzymes critical for DNA replication and transcription (Kampranis & Maxwell, 1998). Antibiotics like sulfonamides function as antimetabolites, blocking the synthesis of folic acid, a vitamin essential for bacterial DNA and RNA synthesis, thereby starving the bacterium of necessary growth factors (reviewed in (Ovung & Bhattacharyya, 2021)). Finally, most relevant for the work presented here are antibiotics targeting the cell wall. Cell wall synthesis can be inhibited by beta-lactams and glycopeptide antibiotics. Examples of commonly-used beta-lactams are penicillin and cephalosporin (Abraham et al., 1954; Fleming, 1929), which disrupt the formation of peptidoglycan cross-links in the bacterial cell wall. Additionally, glycopeptide antibiotics such as vancomycin can inhibit peptidoglycan synthesis by binding to the D-Ala-D-Ala motif in the peptidoglycan precursor (Mccormick et al., 1955). Altogether, treatment with these antibiotics result in the weakening of the bacterial cell wall, which is crucial for maintaining the structural integrity of the cell, and results in cell lysis during growth or cell division.

Peptidoglycan biosynthesis

The peptidoglycan biosynthesis pathway, described here for *E. coli*, begins in the cytoplasm initiated by the proteins MurA and MurB that convert a N-acetylglucosamine

(GlcNAc) molecule into N-acetylmuramic acid (MurNAc)(reviewed in Teo and Roper, 2015). This is followed by the sequential attachment of the amino acids L-alanine, D-glutamate, L-diamino acid, and two D-alanine residues, catalyzed respectively by proteins MurC, MurD, MurE, and MurF. The product of this series of reactions is UDP-MurNAc-pentapeptide, also referred to as Park's Nucleotide based on its discoverer (Park, 1952). The membrane steps begin with the translocase MraY, which transfers the phospho-MurNAc-pentapeptide from Park's Nucleotide onto the lipid undecaprenyl phosphate (C55P) to generate Lipid I (Fig. 1.1). The peripheral-membrane protein MurG catalyzes the transfer of a GlcNAc molecule from a UDP-carrier onto the MurNAc moiety of Lipid I generating Lipid II. The head group of Lipid II is flipped across the membrane into the periplasm by MurJ. Peptidoglycan is made by a variety of enzymes that catalyze the polymerization of the MurNAc-GlcNAc disaccharide and the periodic crosslinking of the pentapeptide resulting in the mesh-like structure. The carrier C55PP is dephosphorylated and then flipped back across the membrane bilayer for recycling (reviewed in Workman and Strynadka, 2020).

There remains a significant need for new antibiotics. The cytoplasmic proteins MurA through MurF have been structurally characterized. Despite advances in targeting the cytoplasmic steps of peptidoglycan biosynthesis, significant gaps remain, particularly in targeting enzymes involved in the membrane-associated steps. Although inhibitors have been discovered against the enzymes that catalyze the membrane steps of the pathway, there are no clinical antibiotics that target MraY, MurG, or MurJ (reviewed in Teo and Roper, 2015). MraY, or phospho-MurNAc-pentapeptide translocase, is especially critical as it catalyzes Lipid I the first membrane-associated step of peptidoglycan biosynthesis. The absence of clinical inhibitors targeting MraY highlights a significant opportunity for antibiotic development.

MraY is essential for cell growth

This work focuses on MraY which is essential for peptidoglycan biosynthesis and cell survival (Boyle & Donachie, 1998). MraY belongs to the polyprenyl-phosphate *N*-acetyl hexosamine 1-phosphate transferase (PNPT) superfamily. These are membrane proteins involved in transfering a sugar moeity onto a polyisoprenol lipid. The substrates of MraY are the UDP-MurNAc-pentapeptide (Park's nucleotide, Fig. 1.2A) and the carrier lipid undecaprenyl phosphate (C55P, Fig. 1.2B). MraY catalyzes the transfer of the phospho-MurNAc-pentapeptide onto the phosphate group in the precursor lipid, with UMP as a leaving gropu. Structural studies of MraY

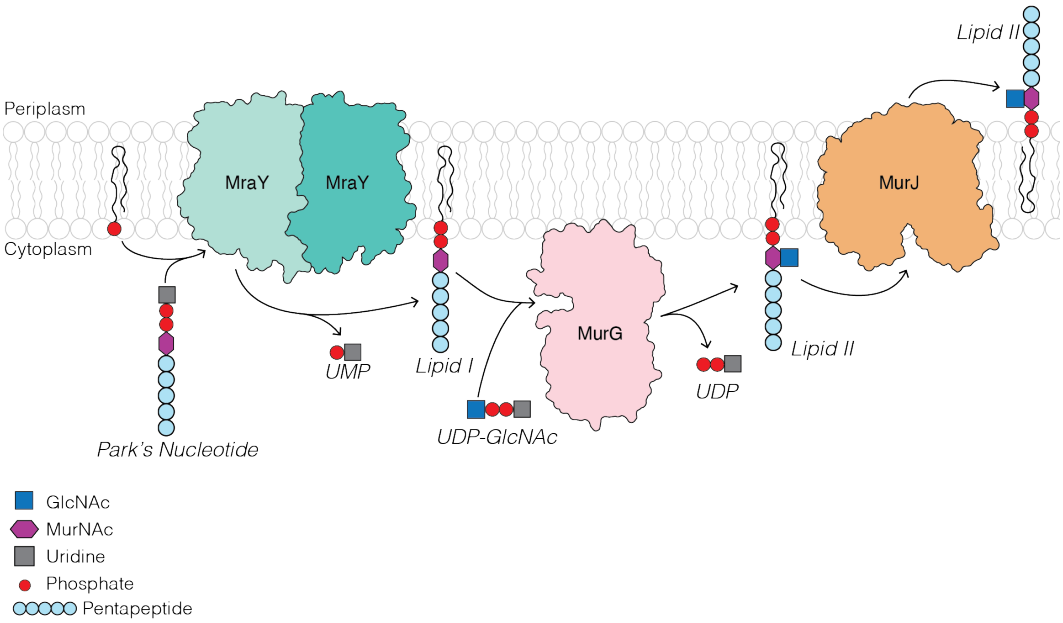


Figure 1.1: **The membrane steps of peptidoglycan biosynthesis**

from non-pathogenic species have determined that the enzymatically active form of MraY is a homodimer composed of ten transmembrane domains (TMDs) with five cytoplasmic loops that are involved in substrate binding (Fig. 1.2C). Notably, TM9 is broken into helices 9a and 9b, where TM9b is bent at a $\tilde{4}5$ degree angle relative to the membrane normal (Chung et al., 2013). The active site requires three catalytic aspartates: D115, D116, and D267 (Fig. 1.2C,D) (Lloyd et al., 2004). These residues were identified upon a CLUSTAL alignment with other prokaryotic homologs of MraY and WecA, and eukaryotic homolog DPAGT1. This class of phospho-GlcNAc transferases are membrane proteins that retain aspartate residues in the cytoplasmic loops. Additionally, a histidine patch present in Loop 9-10 (residues 326 to 328) is highly conserved in bacterial homologs but missing in eukaryotic homologs (Amer & Valvano, 2001; Bouhss et al., 1999). Studies using the *Aquifex aeolicus (Aa)* MraY homolog discovered mutants H324A, D117A, D118A, and D265 (H326, D115, D116, and D267 in *Ec*MraY) that resulted in significant loss of activity, and reduced activity with mutants H325A and H326A (H327A and H328A in *Ec*MraY) (Chung et al., 2013).

Crystal structures solved in complex with inhibitors have shown conformational changes triggered by substrate binding. These occur at the hydrophilic cleft involving 34 conserved amino acids in the cytoplasmic side between TMD5, TMD8, TMD9b, and loops 5-6, 7-8 and 9-10 (Mashalidis et al., 2019). Of particular interest, the

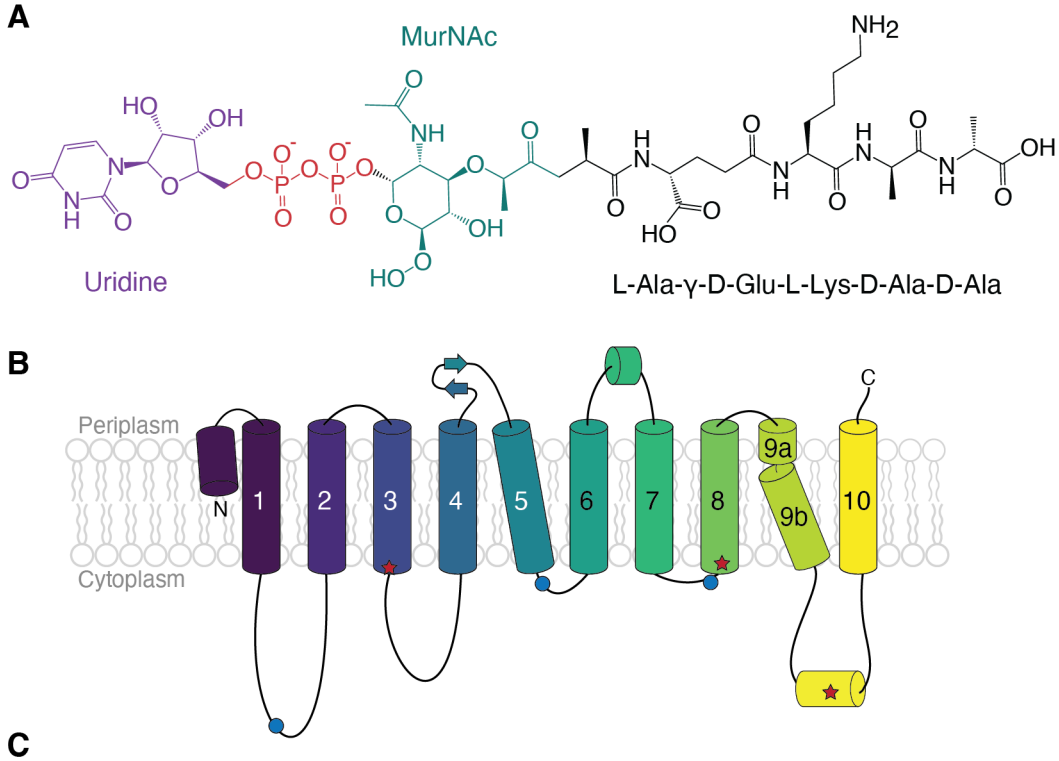
binding of inhibitors causes conformational changes of TM9b and Loop 9-10. The lipid binding site has been mapped to this same region, involving TM9, 5, and 4. Notably, the C55P precursor lipid binding site has been largely understudied due to the difficulty of targeting transmembrane helices through the means of conventional drugs.

An important aim is to developing structure-based inhibitors that do not inhibit the related human protein DPAGT1. These studies have mapped the active site of MraY to 6 hot spot regions, out of which inhibitors typically target two at any given time (Mashalidis et al., 2019). Known MraY inhibitors have been derived from *Streptomyces*, which generate secondary metabolites to out-compete other microorganisms (Procópio et al., 2012). These inhibitors are divided into five classes: muraymycins, tunicamycins, mureidomycins, capuramycins, and liposidomycins/caprazamycins. These molecules bind to a minimal of two "hot spot" regions leading to inhibition. All of these molecules share a uridine and a ribosyl substructure but vary in the secondary binding region. These inhibitors act as competitive inhibitors of Park's nucleotide by blocking the binding pocket (Mashalidis et al., 2019).

Muraymycins and tunicamycins both occlude the binding site for Park's nucleotide by hindering the soluble binding pocket in MraY(Brandish et al., 1996; McDonald et al., 2002). Structural work into mureidomycin-MraY interactions shows that these are competitive for both C55P and Park's nucleotide, with an interaction that extends from the lipid binding site to the uridine pocket (Mashalidis et al., 2019). Tunicamycins, although effective at inhibiting MraY, also inhibit the DPAGT1 (human) and Alg7 (yeast), preventing their clinical use as antibacterials (Kurosu, 2019).

ΦX174 and protein antibiotics

Bacteriophage ΦX174 is a small single-stranded DNA phage that played a historic role in many important milestones in molecular biology Benbow et al., 1971; Hutchison III and Sinsheimer, 1966; Lee and Sinsheimer, 1974; Luria, 1962; Sinsheimer, 1959; Smith et al., 2003; Tessman, 1959. It was the first genome to be fully sequenced by Sanger in 1977, which also identified the first occurrence of nested genes (two overlapping open reading frames) (Sanger et al., 1977). In the case of ΦX174, the lysis gene E is in a +1 reading frame within the ORF of the maturation gene D.



C

```
  1 MLVWLAEHLVKYYSGFNVFSYLTFRAIVSLLTALFISLWMGPRMIAHLQKLSFGQVVRNDGPESHFS  67
 68 KRGTPTMGGIMILTAIVISVLLWAYPSNPYVWCVLVVLVGYGVIGFVDDYRKVVRKDTKGLIARWKY 134
135 FWMSVIALGVAFALYLAGKDTPATQLVVPFFKDVMPQLGLFYILLAYFVIVGTGNAVNLTDGLDGLA 201
204 IMPTVFVAGGFALVAWATGNMNFASYLHIPYLRHAGELVIVCTAIVGAGLGFLWFNTYPAQVFMGDV 268
269 GSLALGGALGIIAVLLRQEFLLVIMGGVFVVETLSVILQVGSFKLRGQRIFRMAPIHHHYELKGWPE 335
336 PRVIVRFWIISLMLVLIGLATLKVR                                           360
```

Figure 1.2: **MraY substrates and key residues**.
A, Chemical structure of Park's nucleotide from *E. coli*. Uridine moiety is shown in purple, phosphate groups in red, MurNAc in green, pentapeptide in black. **B**, Topology diagram with secondary structure of MraY, periplasm and cytoplasm are labeled above and below the membrane respectively. Helices and beta sheets are colored in viridis and labeled 1-10. Residues highlighted in **C** are shown as blue circles (uridine-binding pocket) and red stars (essential residues). **C**, Protein sequence of *Ec*MraY. Transmembrane helices are colored in viridis and bold. Residues that conform the uridine-binding pocket are highlighted in blue. Residues essential for function are highlighted in red.

Single Gene Lysis mechanisms (Sgl)

Bacteriophages carry lytic genes that breach the cell wall of the host to allow escape of assembled virions. The lysis mechanisms can involve either multiple genes or, in the case of small single-stranded nucleotide genomes, single genes encode a lytic protein. These systems are referred to as single-gene lysis (SGL). For instance, the less than 6 kB genomes of ssRNA *Leviviricetes* and ssDNA *Microviridae* bacteriophages must use an Sgl as they cannot encode a multigene system in there small genomes. Sgl of known mechanism inhibit cell wall biosynthesis as virions are assembled, leading to host lysis and phage release similar to that seen for β-lactams (Bernhardt, Wang, et al., 2002).

The majority of Sgl that have been identified are found in the ssRNA *Leviviridae* class. Here, Sgl can be intergenic or, more commonly, are nested within one of the three core proteins of the phage genome. The maturation protein A_2 of the ssRNA bacteriophage Qβ plays multiple including acting as the Sgl that binds and inhibits MurA, the first protein in peptidoglycan biosynthesis (Reed et al., 2012). More recently discovered Sgl target the lipid II flippase MurJ, Sgl^{M} from phage M and Sgl^{PP7} from the *Pseudomonas* phage PP7 (Adler et al., 2023; Chamakura et al., 2017). Sgl that have been identified are the tip of an iceberg, 33 new Sgl were identified in a survey of ssRNA genomes (Chamakura et al., 2020).

Although most of the Sgl experimentally tested to date are derived from ssRNA phages, protein E from the ssDNA bacteriophage ΦX174 was the first identified and the most extensively studied. Bacteriophage ΦX174 belongs to the *Bullavirinae* subfamily within the *Microviridae* family of coliphages. It became the first genome ever sequenced that would reveal 11 genes, including gene E nested within gene D, which encodes a structural protein required for phage maturation (Dokland et al., 1999; Sanger et al., 1977). Protein E, encoded by gene E, contains a highly conserved transmembrane domain composed of the first 29 residues, followed by a long soluble C-terminal domain. Isoforms of protein E in the *Bullavirinae* subfamily can vary in length from 75 to 105 amino acids. Variability in the sequence among isoforms occurs mostly at the cytoplasmic C-terminus of the protein.

Protein E inhibits peptidoglycan biosynthesis assisted by SlyD

Early observations of the ΦX174 lysis phenotype showed rupture of the cell wall at the septal division site, similar to the phenotype caused by penicillin, suggesting inihibition of cell wall biogenesis (Bradley et al., 1969; Young & Young, 1982).

Expression of protein E alone is sufficient for host lysis (Henrich et al., 1982; Young & Young, 1982). This brought forth several theories on the mechanism of function of protein E, including the formation of an oligomeric pore spanning both bilayers, an undefined autolysis mechanism, and peptidoglycan biosynthesis inhibition (Bernhardt et al., 2000; Lubitz et al., 1984; Witte et al., 1990). Conclusive biochemical evidence showed that MraY is the target of protein E (Bernhardt et al., 2000). Surprisingly, it was demonstrated that wild-type protein E required the host chaperone SlyD for lysis (Roof et al., 1994). This chaperone was first discovered in a survival assay where recessive mutants in the *E. coli* chromosome were mapped to the *slyD* (sensitivity to lysis D) gene (Roof et al., 1994). The constitutively expressed SlyD is a peptidyl-prolyl cis-trans isomerase (PPI) composed of two domains, an FK506-Binding Protein (FKBP) prolyl-isomerase domain and an insertion flap (IF) domain that binds to extended-peptide (Hottenrott et al., 1997; Kovermann et al., 2013; Roof et al., 1994; Weininger et al., 2009). The C-terminus includes an extended metal-binding flexible tail that provides the metallochaperone role of SlyD (Löw et al., 2010; Zhang et al., 2005). Structures of SlyD from various species alone and bound to peptides (Pazicky et al., 2022; Quistgaard et al., 2016) show a typical FKBP domain and an IF domain that includes a cleft for binding an extended peptide by β augmentation. The domains are connected by a flexible linker (Weininger et al., 2009) and work together to act as an efficient nucleotide-free chaperone (Scholz et al., 2006). The role for SlyD in protein E-mediated lysis is unclear as SlyD is not essential under typical growth conditions (Mokhonov et al., 2018).

Protein E mutagenesis

Initial work toward revealing a functional mechanism for protein E focused on characterizing the role of specific residues (Bernhardt, Roof, et al., 2002; Buckley & Hayashi, 1986; Maratea et al., 1985; Tanaka & Clemons Jr, 2012; Witte et al., 1997). Early studies identified the transmembrane domain in its entirety (residues 1-29) is essential for inhibition of MraY (Buckley & Hayashi, 1986). In particular, residue P21 is a highly conserved kink in the transmembrane domain that is essential for inhibition (Tanaka & Clemons Jr, 2012; Witte et al., 1997). Loss of SlyD can be bypassed by increasing the total levels of protein E in the membrane through stabilizing mutants R3H and L19F suggesting that prolyl isomerization is not a critical part of the mechanism (Bernhardt, Roof, et al., 2002; Tanaka & Clemons Jr, 2012). Five MraY mutants (P170L, ΔL172, G186S, F288L, V291M) confer resistance to lysis by protein E, presumably hindering key binding sites for protein

E (Bernhardt et al., 2000; Zheng et al., 2008).

Heterologous expression of protein E yields *E. coli* 'bacterial ghosts'

While the *Microviridae* phages specifically infect *E. coli*, the inhibitory activity of protein E is not restricted to *E. coli* MraY. Protein E has been widely used to generate 'bacterial ghosts', bacteria that are lysed with recombinantly expressed protein E to leave only the membrane sac (Mayr et al., 2005; Witte et al., 1992). These have been pursued for a variety of uses including the development of vaccines (Szostak et al., 1996). To make ghosts, protein E must inhibit the native MraY in the desired species. Towards this end, many bacteria have been probed for ghost formation by expression of protein E, revealing that all Gram-negative bacteria previously tested resulted in the formation of 'ghosts' (Cai et al., 2013; Marchart et al., 2003; Muhammad et al., 2019; Simon et al., 2011; Szostak et al., 1996). For Gram-positives, protein E was unable to cause lysis in *Staphylococcus carnosus* (Halfmann et al., 1993). In fact, expression of the *Bacillus subtilis* MraY (*Bs*MraY)can prevent lysis by protein E in *E. coli*, suggesting *Bs*MraY is not inhibited by protein E (Zheng et al., 2008).

Outstanding questions behind the mechanism of lysis by protein E remained. From the structural and mechanistic standpoint, we first wondered how protein E was inhibiting MraY. Given the robust biochemical studies preceding our work, we took a structural approach to this question. In Chapter 2, we present the structure of protein E in complex with host proteins MraY and SlyD. Through the use of cryo-electron microscopy and site-directed mutagenesis, we conclusively demonstrate the mechanism through which protein E blocks peptidoglycan biosynthesis. The data presented in this study brought outstanding questions behind the effects of the lipidic environment on MraY. Concurrently, the Bernhardt group had identified hyperactive MraY mutants distant from the active site, and the Stansfeld group had developed a molecular dynamics platform to study MraY in its native bilayer. In Chapter 3, in collaboration with the Bernhardt and Stansfeld groups, we study the effects of the lipidic environment on MraY and provide evidence for periplasmic regulation of MraY. Leveraging the structures identified in Chapter 2, we explore the mutagenesis of protein E towards a development of a broad MraY inhibitor in Chapter 4. Along with the inhibition of Gram-positive MraY from *Bacillus subtilis*, we show protein E is amenable for mutagenesis beyond its naturally occurring mutations. Finally, Chapter 5 concludes this work by discussing the findings in this thesis, along with future directions towards the studies of MraY regulation and protein E-mediated

bacterial lysis.

References

Abraham, E. P., Newton, G. G. F., & Hale, C. W. (1954). Purification and some properties of cephalosporin N, a new penicillin. *Biochemical Journal*, *58*(1), 94–102. Retrieved May 11, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1269848/

Adler, B. A., Chamakura, K., Carion, H., Krog, J., Deutschbauer, A. M., Young, R., Mutalik, V. K., & Arkin, A. P. (2023). Multicopy suppressor screens reveal convergent evolution of single-gene lysis proteins [Publisher: Nature Publishing Group]. *Nature Chemical Biology*, *19*(6), 759–766. https://doi.org/10.1038/s41589-023-01269-7

Amer, A. O., & Valvano, M. A. (2001). Conserved amino acid residues found in a predicted cytosolic domain of the lipopolysaccharide biosynthetic protein WecA are implicated in the recognition of UDP-N-acetylglucosamine [Publisher: Microbiology Society,]. *Microbiology*, *147*(11), 3015–3025. https://doi.org/10.1099/00221287-147-11-3015

Asokan, G. V., Ramadhan, T., Ahmed, E., & Sanad, H. (2019). WHO Global Priority Pathogens List: A Bibliometric Analysis of Medline-PubMed for Knowledge Mobilization to Infection Prevention and Control Practices in Bahrain. *Oman Medical Journal*, *34*(3), 184–193. https://doi.org/10.5001/omj.2019.37

Benbow, R., Hutchison III, C., Fabricant, J., & Sinsheimer, R. (1971). Genetic map of bacteriophage Φx174. *J. Virol.*, *7*(5), 549–558.

Bernhardt, T. G., Roof, W. D., & Young, R. (2000). Genetic evidence that the bacteriophage Φx174 lysis protein inhibits cell wall synthesis. *Proc. Nat. Acad. Sci.*, *97*(8), 4297–4302.

Bernhardt, T. G., Roof, W. D., & Young, R. (2002). The escherichia coli fkbp-type ppiase slyd is required for the stabilization of the e lysis protein of bacteriophage φx174. *Molecular microbiology*, *45*(1), 99–108.

Bernhardt, T. G., Wang, N., Struck, D. K., & Young, R. (2002). Breaking free:"protein antibiotics" and phage lysis. *Res. Microbiol.*, *153*(8), 493–501.

Bouhss, A., Mengin-Lecreulx, D., Le Beller, D., & Van Heijenoort, J. (1999). Topological analysis of the mray protein catalysing the first membrane step of peptidoglycan synthesis. *Mol. Microbiol.*, *34*(3), 576–585.

Boyle, D. S., & Donachie, W. D. (1998). mraY Is an Essential Gene for Cell Growth in Escherichia coli [Publisher: American Society for Microbiology]. *Journal of Bacteriology*, *180*(23), 6429–6432. https://doi.org/10.1128/jb.180.23.6429-6432.1998

Bradley, D., Dewar, C. A., & Robertson, D. (1969). Structural changes in *Escherichia coli* infected with a Φx174 type bacteriophage. *J. Gen. Virol.*, *5*(1), 113–121.

Brandish, P. E., Kimura, K. I., Inukai, M., Southgate, R., Lonsdale, J. T., & Bugg, T. D. (1996). Modes of action of tunicamycin, liposidomycin B, and mureidomycin A: Inhibition of phospho-N-acetylmuramyl-pentapeptide translocase from Escherichia coli. [Publisher: American Society for Microbiology Journals]. *Antimicrobial Agents and Chemotherapy*, *40*(7), 1640–1644. https://doi.org/10.1128/AAC.40.7.1640

Buckley, K. J., & Hayashi, M. (1986). Lytic activity localized to membrane-spanning region of ϕx174 e protein. *Molecular and General Genetics MGG*, *204*(1), 120–125.

Cai, K., Zhang, Y., Yang, B., & Chen, S. (2013). *Yersinia enterocolitica* ghost with msbb mutation provides protection and reduces proinflammatory cytokines in mice. *Vaccine*, *31*(2), 334–340.

Chamakura, K. R., Sham, L.-T., Davis, R. M., Min, L., Cho, H., Ruiz, N., Bernhardt, T. G., & Young, R. (2017). A viral protein antibiotic inhibits lipid ii flippase activity. *Nature Micro.*, *2*(11), 1480–1484.

Chamakura, K. R., Tran, J. S., O'Leary, C., Lisciandro, H. G., Antillon, S. F., Garza, K. D., Tran, E., Min, L., & Young, R. (2020). Rapid *de novo* evolution of lysis genes in single-stranded rna phages. *Nat. Comm.*, *11*(1), 1–11.

Chung, B. C., Zhao, J., Gillespie, R. A., Kwon, D.-Y., Guan, Z., Hong, J., Zhou, P., & Lee, S.-Y. (2013). Crystal structure of mray, an essential membrane enzyme for bacterial cell wall synthesis. *Science*, *341*(6149), 1012–1016.

Dokland, T., Bernal, R. A., Burch, A., Pletnev, S., Fane, B. A., & Rossmann, M. G. (1999). The role of scaffolding proteins in the assembly of the small, single-stranded dna virus Φx174. *J. Mol. Biol.*, *288*(4), 595–608.

Fleming, A. (1929). On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. *British journal of experimental pathology*, *10*(3), 226–236. Retrieved May 11, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2048009/

Halfmann, G., Götz, F., & Lubitz, W. (1993). Expression of bacteriophage Φx174 lysis gene e in *Staphylococcus carnosus* tm300. *FEMS Microbiol. Lett.*, *108*(2), 139–143.

Henrich, B., Lubitz, W., & Plapp, R. (1982). Lysis of escherichia coli by induction of cloned Φx174 genes. *Mol. Gen. Genet.*, *185*(3), 493–497.

Hottenrott, S., Schumann, T., Plückthun, A., Fischer, G., & Rahfeld, J.-U. (1997). The *Escherichia coli* SlyD Is a Metal Ion-regulated Peptidyl-prolyl *cis/trans*-Isomerase*. *Journal of Biological Chemistry*, *272*(25), 15697–15701. https://doi.org/10.1074/jbc.272.25.15697

Hutchison III, C. A., & Sinsheimer, R. L. (1966). The process of infection with bacteriophage Φx174: X. mutations in a Φx lysis gene. *J. Mol. Biol.*, *18*(3), 429–IN2.

Kampranis, S. C., & Maxwell, A. (1998). The DNA Gyrase-Quinolone Complex: ATP HYDROLYSIS AND THE MECHANISM OF DNA CLEAVAGE*. *Journal of Biological Chemistry*, *273*(35), 22615–22626. https://doi.org/10.1074/jbc.273.35.22615

Kovermann, M., Schmid, F. X., & Balbach, J. (2013). Molecular function of the prolyl cis/trans isomerase and metallochaperone SlyD [Publisher: De Gruyter]. *Biological Chemistry*, *394*(8), 965–975. https://doi.org/10.1515/hsz-2013-0137

Kurosu, M. (2019). Structure-based drug discovery by targeting N-glycan biosynthesis, dolichyl-phosphate N-acetylglucosaminephosphotransferase. *Future Medicinal Chemistry*, *11*(9), 927–933. https://doi.org/10.4155/fmc-2018-0405

Lee, A. S., & Sinsheimer, R. L. (1974). A cleavage map of bacteriophage Φx174 genome. *Proc. Nat. Acad. Sci.*, *71*(7), 2882–2886.

Lin, J., Zhou, D., Steitz, T. A., Polikanov, Y. S., & Gagnon, M. G. (2018). Ribosome-Targeting Antibiotics: Modes of Action, Mechanisms of Resistance, and Implications for Drug Design. *Annual review of biochemistry*, *87*, 451–478. https://doi.org/10.1146/annurev-biochem-062917-011942

Lloyd, A. J., Brandish, P. E., Gilbey, A. M., & Bugg, T. D. H. (2004). Phospho-N-Acetyl-Muramyl-Pentapeptide Translocase from Escherichia coli: Catalytic Role of Conserved Aspartic Acid Residues. *Journal of Bacteriology*, *186*(6), 1747–1757. https://doi.org/10.1128/JB.186.6.1747-1757.2004

Löw, C., Neumann, P., Tidow, H., Weininger, U., Haupt, C., Friedrich-Epler, B., Scholz, C., Stubbs, M. T., & Balbach, J. (2010). Crystal structure determination and functional characterization of the metallochaperone slyd from *Thermus thermophilus*. *J. Mol. Biol.*, *398*(3), 375–390.

Lubitz, W., Halfmann, G., & Plapp, R. (1984). Lysis of *Escherichia coli* after infection with Φx174 depends on the regulation of the cellular autolytic system. *Microbiology*, *130*(5), 1079–1087.

Luria, S. (1962). Genetics of bacteriophage. *Ann. Rev. Microbiol.*, *16*(1), 205–240.

Maratea, D., Young, K., & Young, R. (1985). Deletion and fusion analysis of the phage Φx174 lysis gene e. *Gene*, *40*(1), 39–46.

Marchart, J., Rehagen, M., Dropmann, G., Szostak, M., Alldinger, S., Lechleitner, S., Schlapp, T., Resch, S., & Lubitz, W. (2003). Protective immunity against pasteurellosis in cattle, induced by *Pasteurella haemolytica* ghosts. *Vaccine*, *21*(13-14), 1415–1422.

Mashalidis, E. H., Kaeser, B., Terasawa, Y., Katsuyama, A., Kwon, D.-Y., Lee, K., Hong, J., Ichikawa, S., & Lee, S.-Y. (2019). Chemical logic of mray inhibition by antibacterial nucleoside natural products. *Nat. Comm.*, *10*(1), 1–12.

Mayr, U. B., Walcher, P., Azimpour, C., Riedmann, E., Haller, C., & Lubitz, W. (2005). Bacterial ghosts as antigen delivery vehicles. *Adv. Drug Deliv. Rev.*, *57*(9), 1381–1391.

Mccormick, M. H., Mcguire, J. M., Pittenger, G. E., Pittenger, R. C., & Stark, W. M. (1955). Vancomycin, a new antibiotic. I. Chemical and biologic properties. *Antibiotics Annual*, *3*, 606–611.

McDonald, L. A., Barbieri, L. R., Carter, G. T., Lenoy, E., Lotvin, J., Petersen, P. J., Siegel, M. M., Singh, G., & Williamson, R. T. (2002). Structures of the Muraymycins, Novel Peptidoglycan Biosynthesis Inhibitors [Publisher: American Chemical Society]. *Journal of the American Chemical Society*, *124*(35), 10260–10261. https://doi.org/10.1021/ja017748h

Mokhonov, V. V., Vasilenko, E. A., Gorshkova, E. N., Astrakhantseva, I. V., Novikov, D. V., & Novikov, V. V. (2018). Slyd-deficient escherichia coli strains: A highway to contaminant-free protein extraction. *Biochem. Biophys. Res. Comm.*, *499*(4), 967–972.

Muhammad, A., Kassmannhuber, J., Rauscher, M., Falcon, A. A., Wheeler, D. W., Zhang, A. A., Lubitz, P., & Lubitz, W. (2019). Subcutaneous immunization of dogs with *Bordetella bronchiseptica* bacterial ghost vaccine. *Front. Immunol.*, *10*, 1377.

Ovung, A., & Bhattacharyya, J. (2021). Sulfonamide drugs: Structure, antibacterial property, toxicity, and biophysical interactions. *Biophysical Reviews*, *13*(2), 259–272. https://doi.org/10.1007/s12551-021-00795-9

Park, J. T. (1952). Uridine-5'-Pyrophosphate Derivatives: II. A Structure Common to Three Derivatives. *Journal of Biological Chemistry*, *194*(2), 885–895. https://doi.org/10.1016/S0021-9258(18)55844-0

Pazicky, S., Werle, A.-L. A., Lei, J., Löw, C., & Weininger, U. (2022). Impact of distant peptide substrate residues on enzymatic activity of slyd. *Cell. Mol. Life Sci.*, *79*(3), 1–18.

Procópio, R. E. d. L., Silva, I. R. d., Martins, M. K., Azevedo, J. L. d., & Araújo, J. M. d. (2012). Antibiotics produced by Streptomyces. *The Brazilian Journal of Infectious Diseases: An Official Publication of the Brazilian Society of Infectious Diseases*, *16*(5), 466–471. https://doi.org/10.1016/j.bjid.2012.08.014

Quistgaard, E. M., Weininger, U., Ural-Blimke, Y., Modig, K., Nordlund, P., Akke, M., & Löw, C. (2016). Molecular insights into substrate recognition and catalytic mechanism of the chaperone and fkbp peptidyl-prolyl isomerase slyd. *BMC Biol.*, *14*(1), 1–25.

Reed, C. A., Langlais, C., Kuznetsov, V., & Young, R. (2012). Inhibitory mechanism of the Q$\Beta$ lysis protein A2. *Molecular Microbiology*, *86*(4), 836–844. https://doi.org/10.1111/mmi.12021

Roof, W. D., Horne, S. M., Young, K. D., & Young, R. (1994). *slyD*, a host gene required for Φx174 lysis, is related to the fk506-binding protein family of peptidyl-prolyl cis-trans-isomerases. *J. Biol. Chem.*, *269*(4), 2902–2910.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C., Slocombe, P. M., & Smith, M. (1977). Nucleotide sequence of bacteriophage Φx174 dna. *Nature*, *265*(5596), 687–695.

Scholz, C., Eckert, B., Hagn, F., Schaarschmidt, P., Balbach, J., & Schmid, F. X. (2006). Slyd proteins from different species exhibit high prolyl isomerase and chaperone activities. *Biochemistry*, *45*(1), 20–33.

Schrader, S. M., Botella, H., & Vaubourgeix, J. (2023). Reframing antimicrobial resistance as a continuous spectrum of manifestations. *Current Opinion in Microbiology*, *72*, 102259. https://doi.org/10.1016/j.mib.2022.102259

Simon, R., Tennant, S. M., Wang, J. Y., Schmidlein, P. J., Lees, A., Ernst, R. K., Pasetti, M. F., Galen, J. E., & Levine, M. M. (2011). *Salmonella enterica* serovar enteritidis core o polysaccharide conjugated to h: G, m flagellin as a candidate vaccine for protection against invasive infection with *S. enteritidis*. *Infec. Immun.*, *79*(10), 4240–4249.

Sinsheimer, R. L. (1959). A single-stranded deoxyribonucleic acid from bacteriophage Φx174. *J. Mol. Biol.*, *1*(1), 43–IN6.

Smith, H. O., Hutchison, C. A., Pfannkoch, C., & Venter, J. C. (2003). Generating a synthetic genome by whole genome assembly: ΦX174 bacteriophage from synthetic oligonucleotides. *Proc. Nat. Acad. Sci.*, *100*(26), 15440–15445.

Szostak, M. P., Hensel, A., Eko, F. O., Klein, R., Auer, T., Mader, H., Haslberger, A., Bunka, S., Wanner, G., & Lubitz, W. (1996). Bacterial ghosts: Non-living candidate vaccines. *J. Biotech.*, *44*(1-3), 161–170.

Tanaka, S., & Clemons Jr, W. M. (2012). Minimal requirements for inhibition of mray by lysis protein e from bacteriophage Φx174. *Molecular microbiology*, *85*(5), 975–985.

Teo, A. C., & Roper, D. I. (2015). Core steps of membrane-bound peptidoglycan biosynthesis: Recent advances, insight and opportunities. *Antibiotics*, *4*(4), 495–520.

Tessman, I. (1959). Some unusual properties of the nucleic acid in bacteriophages s13 and Φx174. *Virology*, *7*(3), 263–275.

Weininger, U., Haupt, C., Schweimer, K., Graubner, W., Kovermann, M., Brüser, T., Scholz, C., Schaarschmidt, P., Zoldak, G., Schmid, F. X., et al. (2009). Nmr solution structure of slyd from *Escherichia coli*: Spatial separation of prolyl isomerase and chaperone function. *J. Mol. Biol.*, *387*(2), 295–305.

Witte, A., Bläsi, U., Halfmann, G., Szostak, M., Wanner, G., & Lubitz, W. (1990). ΦX174 protein e-mediated lysis of escherichia coli. *Biochimie*, *72*(2-3), 191–200.

Witte, A., Wanner, G., Sulzner, M., & Lubitz, W. (1992). Dynamics of phix174 protein e-mediated lysis of escherichia coli. *Arch. Microbiol.*, *157*(4), 381–388.

Witte, A., Schrot, G., Schön, P., & Lubitz, W. (1997). Proline 21, a residue within the α-helical domain of Φx174 lysis protein e, is required for its function in escherichia coli. *Molecular microbiology*, *26*(2), 337–346.

Workman, S. D., & Strynadka, N. C. J. (2020). A Slippery Scaffold: Synthesis and Recycling of the Bacterial Cell Wall Carrier Lipid. *Journal of Molecular Biology*, *432*(18), 4964–4982. https://doi.org/10.1016/j.jmb.2020.03.025

Young, K. D., & Young, R. (1982). Lytic action of cloned Φx174 gene e. *J. Virol.*, *44*(3), 993–1002.

Zhang, J. W., Butland, G., Greenblatt, J. F., Emili, A., & Zamble, D. B. (2005). A role for slyd in the *Escherichia coli* hydrogenase biosynthetic pathway. *J. Biol. Chem.*, *280*(6), 4360–4366.

Zheng, Y., Struck, D. K., Bernhardt, T. G., & Young, R. (2008). Genetic analysis of mray inhibition by the Φx174 protein e. *Genetics*, *180*(3), 1459–1466.

Chapter 2

THE MECHANISM OF THE ΦX PHAGE ENCODED PROTEIN ANTIBIOTIC

Adapted from:

Orta, A. K., Riera, N., Li, Y. E., Tanaka, S., Yun, H. G., Klaic, L., & Clemons Jr, W. M. (2023). The mechanism of the phage-encoded protein antibiotic from Φx174. *Science*, *381*(6654), eadg9091. https://doi.org/10.1126/science.adg9091.

ABSTRACT

A key step in the bacteriophage life cycle is the requirement to breach the peptidoglycan layer of the bacterial cell wall. While a variety of lysis mechanisms have evolved, the simplest are found in single stranded DNA or RNA bacteriophages that, constrained by the small size of their genomes, encode a single gene lysis (SGL) protein. The first discovered and most studied example is protein E from ΦX174 in the *Microviridae* family; a 91 amino acid peptide with a single transmembrane domain at its N-terminus. Protein E expression, dependent on the host chaperone SlyD, is sufficient for lysis of bacteria via inhibition of the phospho-MurNAc-pentapeptide translocase MraY, an essential enzyme in the biosynthesis of peptidoglycan. Despite the historic importance of ΦX174, the lysis mechanism remains poorly defined. Using single particle electron cryo-microscopy, here we demonstrate that protein E forms a stable inhibitory complex with both *E. coli* MraY and SlyD by physically blocking access to the active site of MraY . Functional insight is gained for both SlyD and MraY. These results reveal a new path towards the development of bacterial antibiotics.

2.1 Introduction

Phage therapy provides a potential solution to the multi-drug resistant bacteria problem (Matsuzaki et al., 2014; Schooley et al., 2017) and to fully realize this requires a fundamental understanding of the mechanisms viruses use to kill their host (Duan et al., 2022; Ferriol-González & Domingo-Calap, 2021; Holtappels et al., 2021; Kortright et al., 2019). The first use of phage to treat a bacterial infection occurred soon after phage were discovered by Felix d'Herelle (Herelle, 1948). Quickly, phage cocktails were developed that contained a variety of viruses including the historically important phage ΦX174 (Herelle, 1948), which has been at the center of many critical discoveries in molecular biology (Benbow et al., 1971; Hutchison III & Sinsheimer, 1966; Lee & Sinsheimer, 1974; Luria, 1962; Sinsheimer, 1959; Smith et al., 2003; Tessman, 1959) since its isolation in 1935 (Sertic & Boulgakov, 1935). The coliphage ΦX174 belongs to an abundant family of phages found broadly in environments that contain colibacteria (*e.g. E. coli*) including in the human gut (Kim et al., 2011). Of many first discoveries, one of the most notable was the 1977 publication of its complete genome (Sanger et al., 1977), a milestone for both genomics and sequencing technologies. This revealed that the small genome, 11 genes in 5.4 kilobase pairs, included a single lysis gene, called *E*, that encoded the small protein E embedded within the open reading frame of the scaffolding Protein D (Barrell et al., 1976; Denhardt & Sinsheimer, 1965; Hutchison III & Sinsheimer, 1963). The life cycle of ΦX174 has been extensively studied (Cherwa & Fane, 2011; Sun et al., 2014), yet the mechanism for the key step of cell lysis remains to be determined (K. Chamakura & Young, 2019).

Early observations of the lysis mechanism of ΦX174 showed rupture of the cell wall at the septal division site, similar to that of peptidoglycan biogenesis inhibitors (Bradley et al., 1969; Young & Young, 1982). Subsequent work demonstrated that expression of gene *E* is sufficient for bacterial lysis (Henrich et al., 1982; Young & Young, 1982). Protein E is part of a growing group of single gene lysis (SGL) proteins that are found in small single-stranded DNA and RNA phages, few of which have been mechanistically characterized (Bernhardt et al., 2001; K. Chamakura & Young, 2019; K. R. Chamakura et al., 2017; K. R. Chamakura & Young, 2020). Various proposed mechanisms for protein E include an undefined autolysis system, oligomerization of protein E spanning both membrane layers into an oligomeric tunnel, and lysis by inhibition of peptidoglycan synthesis (Bernhardt et al., 2000; Lubitz et al., 1984; Witte et al., 1990). The latter is the most likely, evidenced by the demonstration that protein E directly inhibited the phospho-

MurNAc-pentapeptidetranslocase MraY, an integral membrane protein that is a key enzyme in the peptidoglycan biosynthesis pathway, resulting in disruption of the cell wall (Bernhardt et al., 2000; Bouhss et al., 2007). A surprising result was the demonstration that protein E inhibition of MraY was dependent on the cytoplasmic chaperone SlyD (**S**ensitivity to **ly**sis D) (Maratea et al., 1985; Roof et al., 1994).

The constitutively expressed SlyD is a peptidyl-prolyl cis-trans isomerase (PPI) composed of two domains, an FK506-Binding Protein (FKBP) prolyl-isomerase domain and an insertion flap (IF) domain (Roof et al., 1994; Weininger et al., 2009). The extended C-terminus includes a metal binding region key to the metallochaperone role of SlyD (Löw et al., 2010; Zhang et al., 2005). Structures of SlyD from various species alone and bound to peptides (Pazicky et al., 2022; Quistgaard et al., 2016) show a typical FKBP domain and an IF domain that includes a cleft for binding an extended peptide by β augmentation. The domains are connected by a flexible linker (Weininger et al., 2009) and work together to act as an efficient nucleotide free chaperone (Scholz et al., 2006). The role for SlyD in protein E mediated lysis is unclear as SlyD is not essential under typical growth conditions (Mokhonov et al., 2018).

The essential peptidoglycan layer (PG), a sugar-peptide mesh polymer, is a key component of the cell wall that is unique to eubacteria, maintaining cell shape and providing protection against changes in osmotic pressure (Egan et al., 2020; Vollmer et al., 2008). PG biogenesis (reviewed in (Teo & Roper, 2015)) has long been an attractive target for antibiotic development, yet clinical inhibitors of the membrane steps, which includes MraY, remain to be discovered (Bugg et al., 2011). MraY catalyzes the transfer of a phospho-MurNAc-pentapeptide from the sugar donor of 'Park's nucleotide' (UDP-MurNAc-pentapeptide) onto the phosphate of the 55 carbon polyisoprenyl lipid known as bactoprenol (C55P) generating Lipid I. The antibacterial mechanism for inhibitors that target PG biogenesis is a loss of peptidoglycan precursors resulting in a weakened cell wall that leads to rupture of the membrane during growth. Three PG biogenesis enzymes are targets of viral SGL proteins, MurA, MurJ, and, most relevant here, MraY and likely there will be more (K. R. Chamakura et al., 2020; K. R. Chamakura & Young, 2020).

MraY belongs to the polyprenyl-phosphate-N-acetylhexosamine-1-phosphate-transfe -rase superfamily (Lehrman, 1994; Price & Momany, 2005). Known inhibitors of MraY are nucleoside derived and compete for the Park's nucleotide binding site (Winn et al., 2010). Structures of MraY have been solved by X-ray crystallography

in detergent from either the thermophilic Gram-negative *Aquifex aeolicus* (*Aa*MraY) or the Gram-positive *Enterocloster bolteae* (*Eb*MraY formerly known as *Clostridium bolteae* (Haas & Blanchard, 2020)) (Chung et al., 2013; Hakulinen et al., 2017; Mashalidis et al., 2019). Although neither species is pathogenic or from a model species for MraY studies, the structures revealed a conserved architecture of a homodimer with ten transmembrane domains (TMDs). While the exact catalytic mechanism remains to be resolved, the structures localized the active site in a vestibule on the cytoplasmic side of a cleft formed by TMs 4, 5, & 9 which includes the predicted access site for the lipid substrate (Al-Dabbagh et al., 2008). There are several prominent features in the structure associated with the active site. TM5 is steeply angled relative to the membrane. TM9 contains a sharp kink breaking the TM9 helix with TM9a roughly perpendicular to the membrane and TM9b bent at a 45 degree angle relative to the membrane normal (Chung et al., 2013). Several conserved loops surround the active site. Loop 9-10, the loop connecting TM9 and TM10, being the most distinct as it contains important catalytic His residues (Al-Dabbagh et al., 2008), and adopts different conformations in the crystal structures (Mashalidis et al., 2019). Loop 1-2 is disordered in the crystal structures and contains a number of conserved residues that may be important for function. Structures bound to inhibitors (Hakulinen et al., 2017; Mashalidis et al., 2019) localized the UDP binding site of Park's nucleotide and suggested a path for the acyl-chain of the lipid substrate that would go from the active site and along the cleft, wrapping around TM9b.

Protein E, found in the *Bullavirinae* viral subfamily within the *Microviridae* family, contains a conserved N-terminal transmembrane domain (TMD) and a cytoplasmic C-terminal domain (CTD) that includes a positively charged predicted amphipathic helix and an unstructured tail (Fig. 2.3a, 2.1). The TMD of protein E spans residues 5-22 and, when appended to a C-terminal soluble protein, is sufficient for host lysis (Buckley & Hayashi, 1986; Maratea et al., 1985; Tanaka & Clemons Jr, 2012). However, it is noteworthy that the CTD contains a number of conserved residues including multiple positive charges (Fig. 2.3a, 2.1).

Initial work toward a functional mechanism for protein E focused on characterizing the role of specific residues (Bernhardt et al., 2002; Buckley & Hayashi, 1986; Maratea et al., 1985; Tanaka & Clemons Jr, 2012; Witte et al., 1997). Mutational analysis of the TMD identified a face of the helix that is important for efficient lysis (Tanaka & Clemons Jr, 2012) and that the residue P21 is essential for function

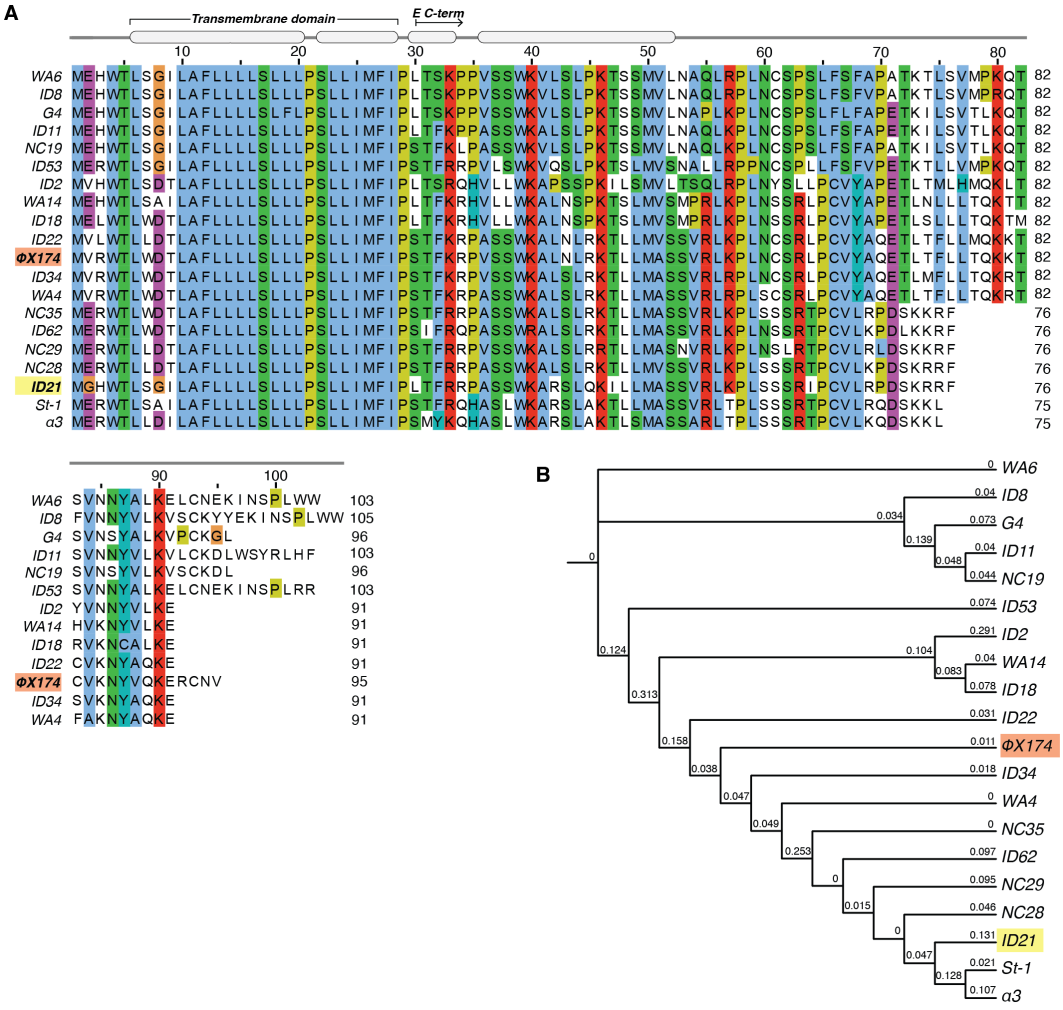


Figure 2.1: **Broad comparison of protein E isoforms from *Bullavirinae*. A,** Sequence alignment of protein E species from UniRef90 (Steinegger & Söding, 2018) using ClustalW. Helices are indicated above the alignment. The TMD and E C-term for protein E are indicated. ΦX174 and ID21 are highlighted. **b,** Phylogenetic tree of protein E isoforms with branch lengths annotated.

(Tanaka & Clemons Jr, 2012; Witte et al., 1997). The first direct mechanistic results identified SlyD as required for protein E-mediated lysis (Roof et al., 1994). Studies have posited various roles for SlyD in lysis, such as the isomerization of P21 (Witte et al., 1997) or inducing oligomerization into a pore (Mezhyrova et al., 2021). Contrasting an enzymatic requirement, loss of SlyD can be bypassed by increasing the total levels of protein E in the membrane (Bernhardt et al., 2002; Tanaka & Clemons Jr, 2012). Despite decades since the observation that MraY and SlyD are key players in lysis mediated by protein E, the mechanism through which these proteins work together to facilitate rupture of the cell wall remains unknown.

Here we demonstrate that the viral protein E bridges the two bacterial proteins, MraY and SlyD, to form the dimeric heterotrimer YES complex (*Ec*Mra**Y**, Protein **E**, *Ec***S**lyD). This results in inhibition of peptidoglycan biogenesis by obstruction of the MraY active site and loss of Lipid I production. We experimentally validate this result for two different viral species, providing a clear model for bacterial lysis by this subfamily of phage. Through these studies we additionally characterize the first MraY structure from an important pathogen, the most complete model to date, revealing new features of this essential enzyme. The structure of SlyD is the first bound to a protein highlighting important functional interfaces. Finally, the EM maps allow visualization of extensive lipid interactions. These results highlight a novel mechanism for killing bacteria and provides a potential route towards the development of therapeutics.

2.2 Results

The structure of the YES complex

With a number of *Bullavirinae* phage to choose from (Fig. 2.1), we performed the following studies using the protein E sequence from either the original phage ΦX174 or the shorter protein E isoform from phage ID21 (91 and 76 residues respectively). We first established that protein E and MraY formed a stable complex by co-expressing an affinity tagged ΦX174 protein E and a wild-type *Ec*MraY. After purification, we identified a stable complex that ran as a single peak by size-exclusion chromatography (SEC) but resolved as four bands on a gel. Two bands were the expressed proteins and the other two were determined to be the endogenous *Ec*SlyD (Fig. 2.2). The higher SlyD band is likely a dimer formed by the multiple histidines and cysteines in the disordered C-terminus, as without this region we see a single band on a gel (Fig. 2.3B) (Löw et al., 2010; Roof et al., 1994). To avoid this disordered region, except where indicated, for all subsequent work we used the conserved core of SlyD without the C-terminus, *Ec*SlyD that was truncated after residue 154. Importantly, this truncation rescued the lysis activity in a Δ*slyD* strain (Fig. 2.2) and had previously been shown to retain chaperone activity (Martino et al., 2009).

The two protein E isoforms, from either ΦX174 or ID21, were able to induce lysis at a similar efficiency when expressed in a wild-type *E. coli* strain (Fig. 2.3A & C). For purification, all three genes in the YES complex (wild-type *Ec*MraY, protein E, and truncated SlyD) were recombinantly expressed together in the Δ*slyD E. coli* strain (Roof et al., 1994) with a C-terminal affinity tag on protein E. The complex was extracted in detergent and purified resulting in a single peak by size exclusion chromatography (SEC) with all three proteins in an apparent stoichiometric complex (Fig. 2.3B).

Structures for both the $YES_{\Phi X174}$ and YES_{ID21} complexes were solved using single particle electron cryo-microscopy (Fig. 2.3D, 2.4, & 2.5). The final density maps were obtained following several rounds of data processing with heterogeneous and homogeneous refinements (Figs. 2.4, 2.5). The final refined overall resolution was 3.4Å for the YES_{ID21} complex and 3.6Å for the $YES_{\Phi X174}$ complex (Fig. 2.6). Final statistics are provided in Table 2.1. For both structures, the resolution was higher for regions in and nearest the membrane and 90% of the protein could be built unambiguously (Fig. 2.4 & 2.5). Sequence differences between ΦX174 and ID21 are visible in the density (Fig. 2.6), but are generally in regions that are

solvent or membrane exposed. Due to the higher resolution and completeness of the YES_{ID21} complex, except where noted, it is used as the reference structure. Within the density map, we could clearly distinguish two copies of each member of the YES complex (six separate proteins). When contoured to remove the detergent micelle, densities for 22 TMDs are clear, 20 of which are accounted for by the back-to-back dimer of *Ec*MraY (Fig. 2.3D, E). The majority of the cytoplasmic density can be accounted for by two SlyD molecules, which show the most flexibility (Fig. 2.7D). The remaining protein density is the two protein E molecules that each contain a bent TMD and soluble domains that bridge between MraY and SlyD. Density for lipids are visible bound around the membrane exposed surface of the MraY dimer (Fig. 2.3D).

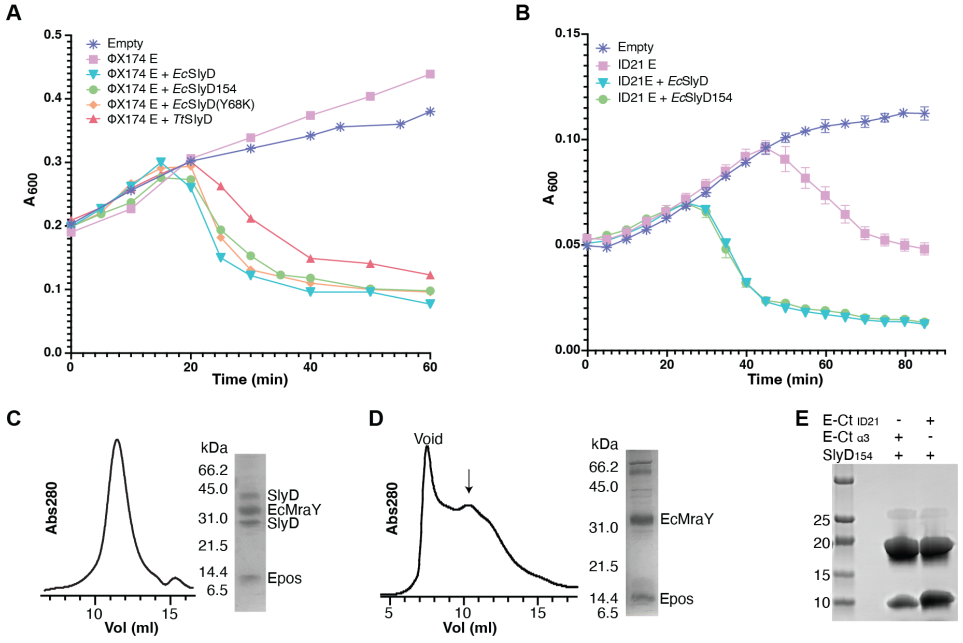


Figure 2.2: **Stable complex formation of SlyD with protein E. A,** Lysis assay in a Δ*slyD* background. Either empty vector (∗), protein $E_{\Phi X174}$ alone (□), or protein $E_{\Phi X174}$ with either *Ec*SlyD (▽), *Ec*$SlyD_{154}$ (○), *Ec*SlyD Y68K (◇), or *Tt*SlyD (△). **B,** Similar to (**A**), lysis assay in a Δ*slyD* background with empty vector (▽), protein E_{ID21} alone (○) or protein E_{ID21} with either *Ec*SlyD (◇) or *Ec*$SlyD_{154}$ (□). **C,** Co-expressed Epos ($E_{\Phi X174}$ R3H, L19F) with *Ec*MraY in a wild-type *E. coli* background purified by SEC and the corresponding SDS-PAGE. Protein bands are labeled. **D,** Co-expressed Epos with *Ec*MraY in Δ*slyD* background purified by SEC and the corresponding SDS-PAGE. Arrow highlights the fraction run on the gel. **E,** SDS-PAGE of purified protein E C-term from either ID21 or α3 co-expressed with $SlyD_{154}$.

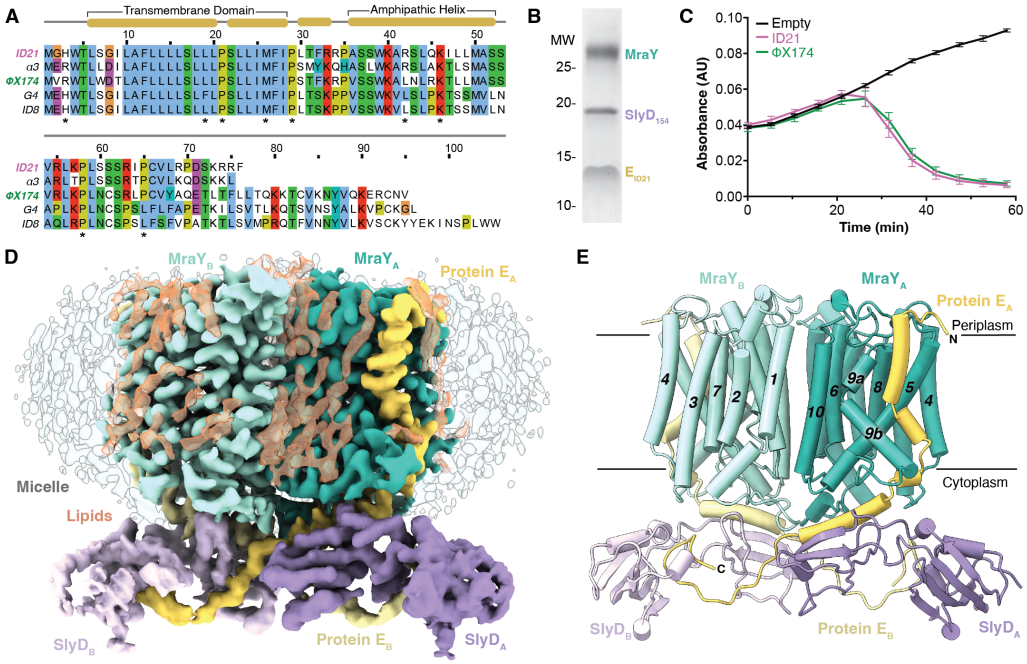


Figure 2.3: **The structure of the YES complex. A,** An alignment of a representative subset of protein E isoforms. Residue coloring is based on ClustalW (Madeira et al., 2022). Secondary structure elements are shown above the sequence. Sequences are ordered as in Fig. 2.1. **B,** SDS-PAGE gel of the purified YES_{ID21} complex.**C,** A lysis assay for protein E expression. Cells containing either empty (black) or the protein E genes for ID21 (pink) or ΦX174 (green) were induced at time 0 and the absorbance at 600 nm was monitored over time. **D,** Overview density maps of the YES_{ID21} complex viewed in the plane of the membrane. The map in grey highlights the detergent micelle. The higher contoured map shows the six components of the two-fold complex with density for *E. coli* MraY (cyan), protein E (yellow), and *E. coli* SlyD (purple) highlighted. The pairs of each protein are distinguished by the B-subunit colored lighter and the general coloring scheme is maintained for images of the complex. Density that is likely lipid is shown in transparent orange. **E,** Cartoon representation of the YES_{ID21} complex oriented and colored as in (**D**). Foreground TMDs for the MraYs are numbered and the bilayer is represented by lines.

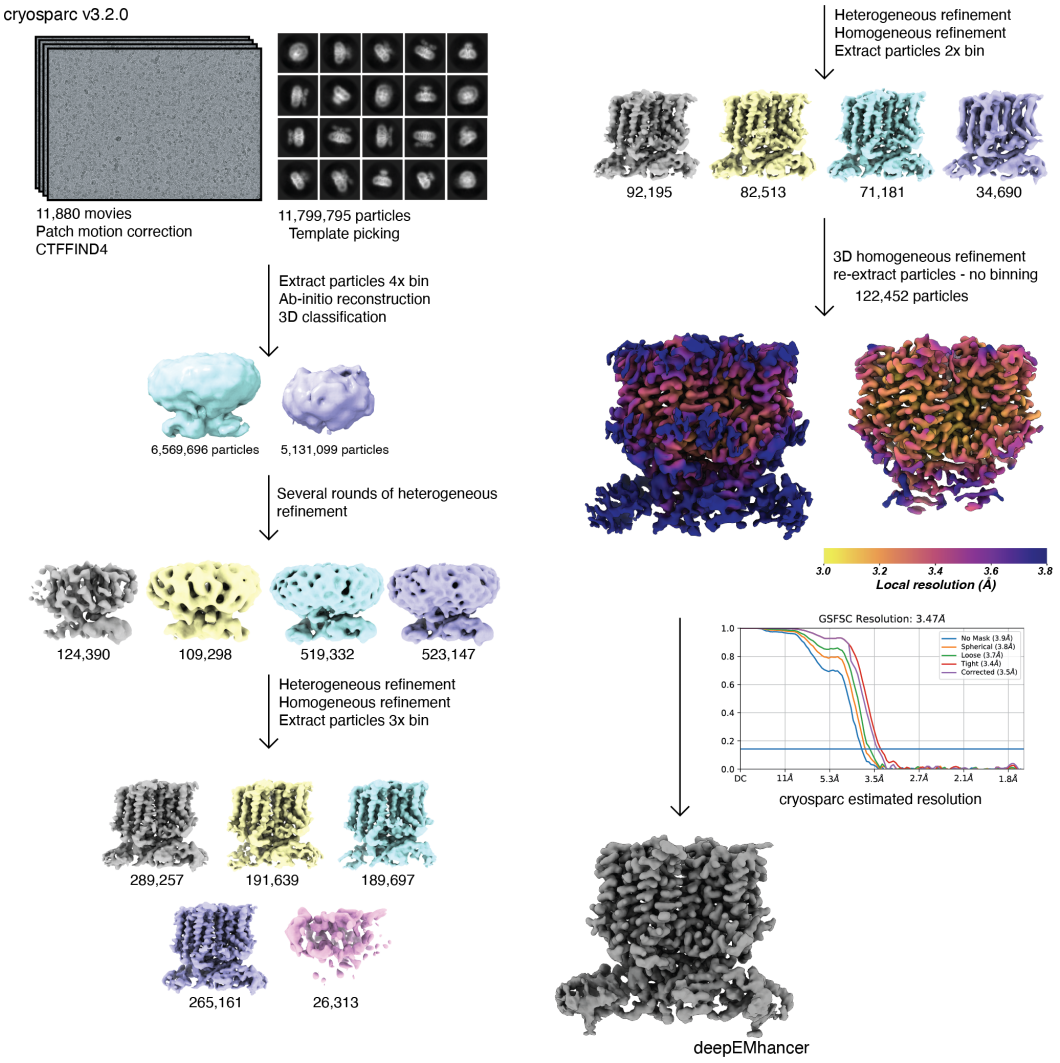


Figure 2.4: **CryoEM processing pipeline for YES_{ID21} complex.** Processing was done using cryosparc v3.2.0. The order of processing follows the arrows. The number of particles after each step is shown. Local resolution is shown in a Plasma color scheme ranging from 3.0Å(yellow) to 3.8Å(purple) resolution. Overall GSFSC resolution was determined on cryosparc. Final model was sharpened using DeepEMhancer (Sanchez-Garcia et al., 2021).

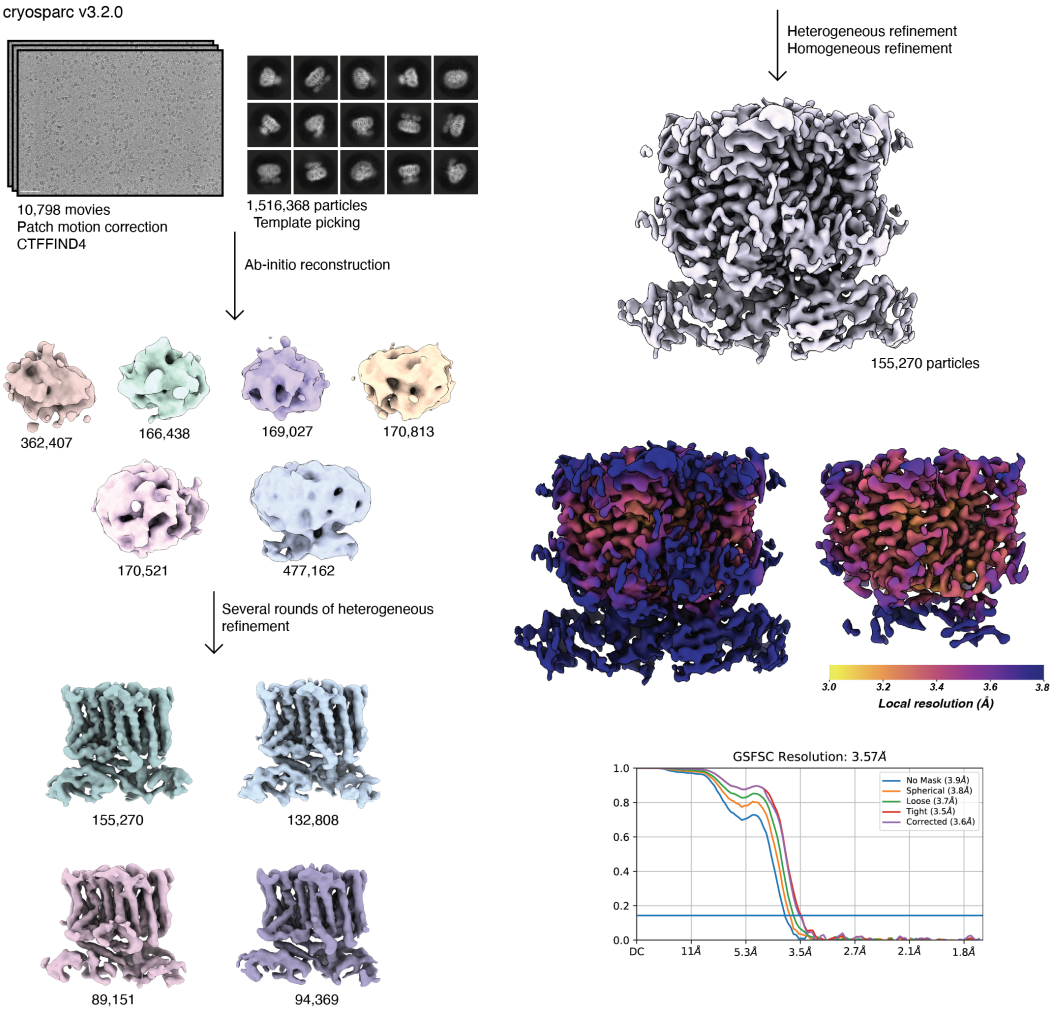


Figure 2.5: **CryoEM processing pipeline for $YES_{\Phi X174}$ complex.** Processing was done using cryosparc v3.2.0. The order of processing follows the arrows. The number of particles after each step is shown. Local resolution is shown in a Plasma color scheme ranging from 3.0Å(yellow) to 3.8Å(purple) resolution. Overall GSFSC resolution was determined on cryosparc.

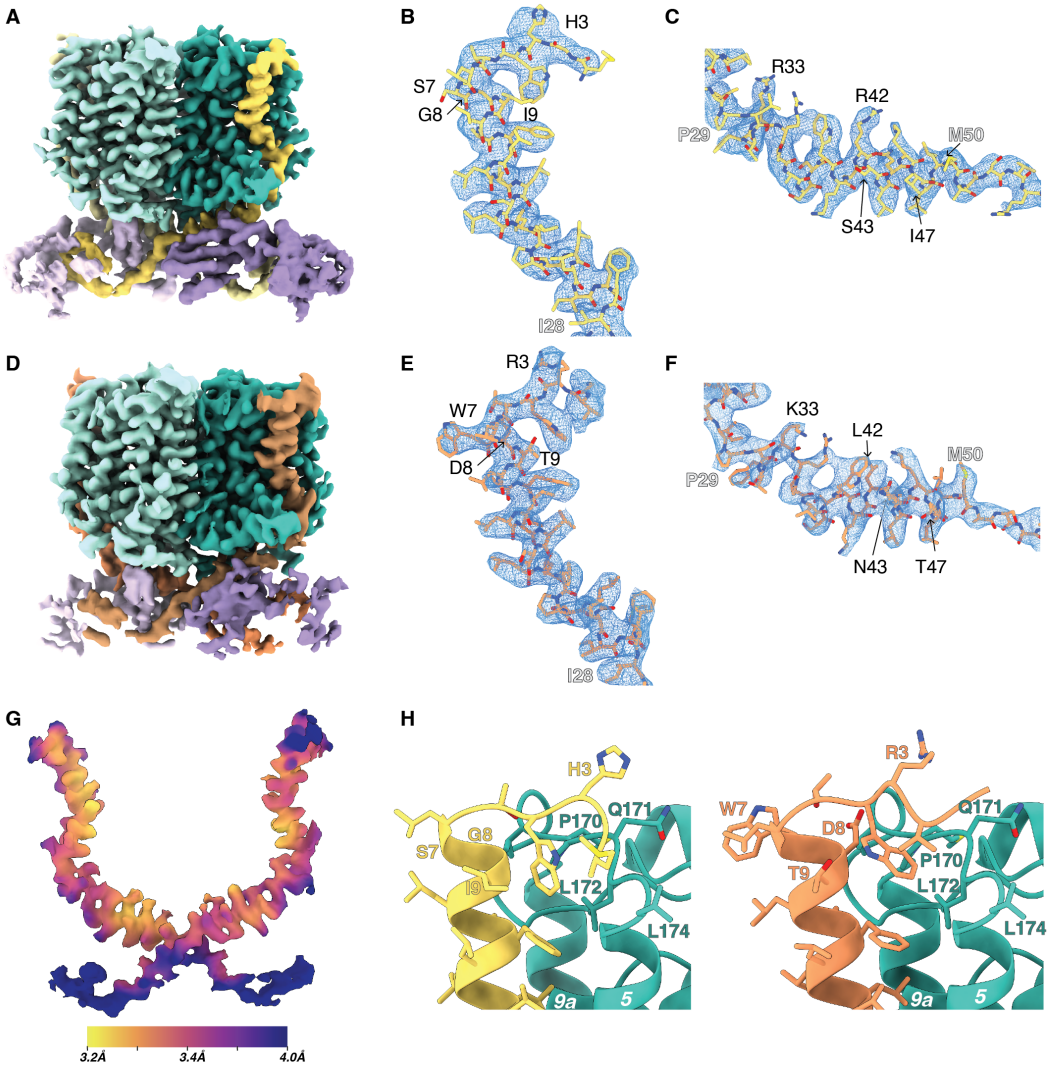


Figure 2.6: **Comparison of the YES_{ID21} complex and the $YES_{\Phi X174}$ complex. A,** Similar to Fig. 2.3a for protein density of the YES_{ID21} complex. **B,** Density maps (blue mesh) for protein E_{ID21} with protein in stick representation highlighting the regions that contain the TMD. **C**, the amphipathic cytoplasmic helix. **D-F**, similar to (**A**-**C**) except for the $YES_{\Phi X174}$ complex. Here protein E is shown in orange. **G,** The N-terminal interactions of protein E_{ID21} with *Ec*MraY. Proteins are shown in cartoon with relevant side chains as sticks. **H,** As in (**G**), for protein $E_{\Phi X174}$. In **B, C, & E-H** residues that vary between the two isoforms are labeled. Residues I28, P29, and M50 are labeled for orientation.

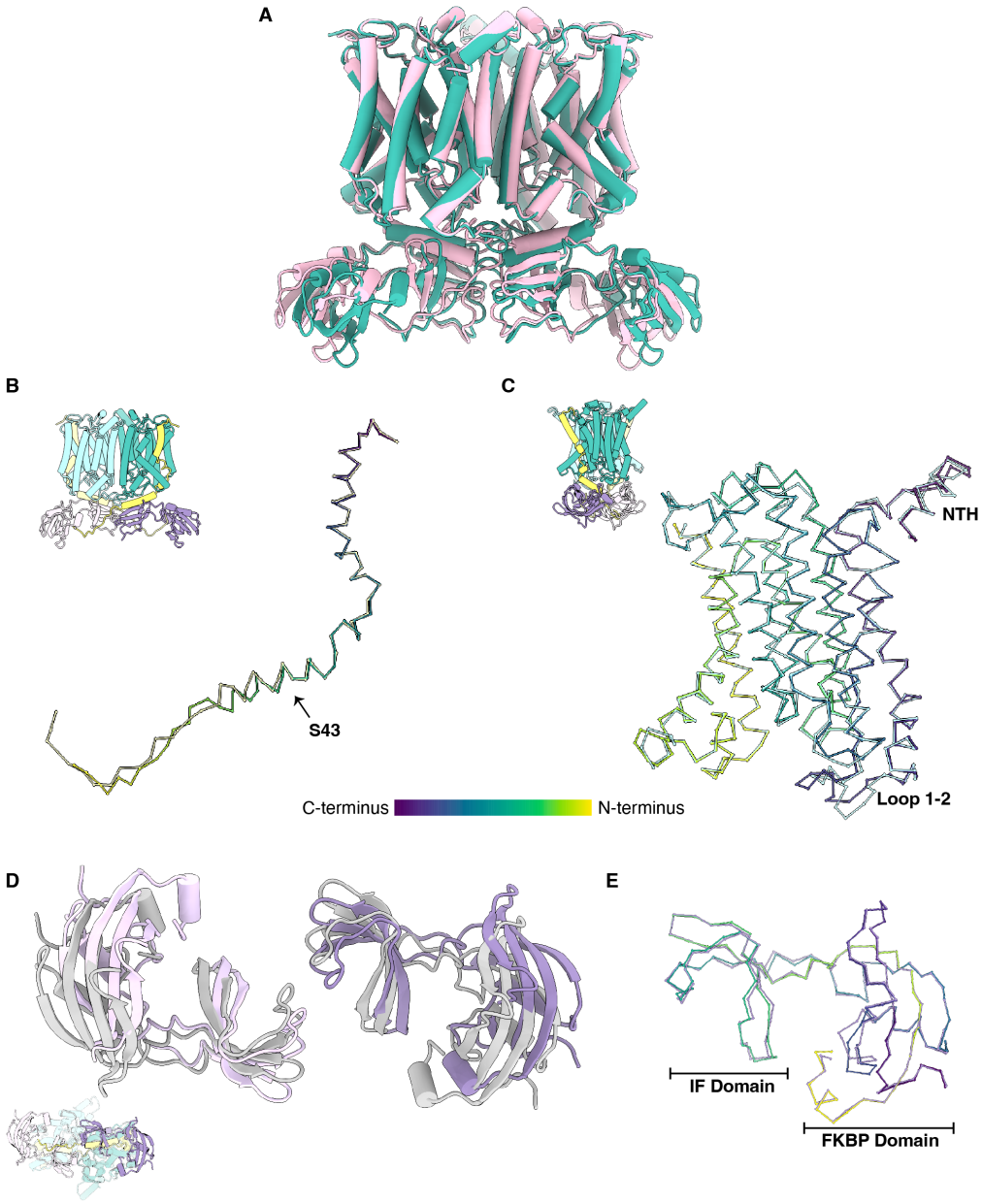


Figure 2.7: **Symmetry in the YES complex. A,** Overlay of a symmetry imposed, 180°rotation, of the YES_{ID21} complex. The proteins are shown as cartoons in either cyan or pink. **B,** Cα ribbon of the two structurally aligned protein E_{ID21} monomers (A: Viridis, B: yellow). Symmetry is broken after residue S43. The inset shows the orientation relative to A. **C,** Cα ribbon of the two structurally aligned MraY monomers (A: Viridis, B: cyan). The inset shows the orientation relative to A as in (**B**). The NTH and Loop 1-2 are labeled. **D,** Viewed from the cytoplasm, cartoon models of SlyDs from the aligned YES complex in (**A**) highlighting the change in orientation. **E,** Cα ribbon of the structurally aligned SlyD monomers (A: Viridis, B: light purple).

For protein E, following from the N-terminus in the periplasm, the TMD binds in the groove formed between TM5 and TM9 of MraY ending in the active-site pocket on the cytoplasmic side where it makes a sharp turn (Fig. 2.3E & 2.8A). The TMD is followed by an amphipathic helix that crosses the active-site groove, parallel to the membrane, presenting a positive face toward MraY and a hydrophobic face towards a SlyD IF-domain. The remaining C-terminus is in an extended conformation that primarily interacts with the second SlyD. This results in a cross over point between the two protein Es with each contacting both SlyDs. Overall, the dimeric complex (two of each of the heterotrimers) has a near two fold symmetry perpendicular to the plane of the membrane. While the membrane and periplasmic facing regions overlay perfectly, the symmetry is broken at the cytoplasmic face where two protein E C-termini cross each other at different residues and the SlyD adopt slightly different orientations (Figs. 2.3E & 2.7). For the end of protein E, we can see continuous backbone density that positions proline 65 in the active site of the FKBP domain. Beyond that, the density is insufficient to resolve the sequence and we see little difference between ΦX174 and ID21 (Fig. 2.6).

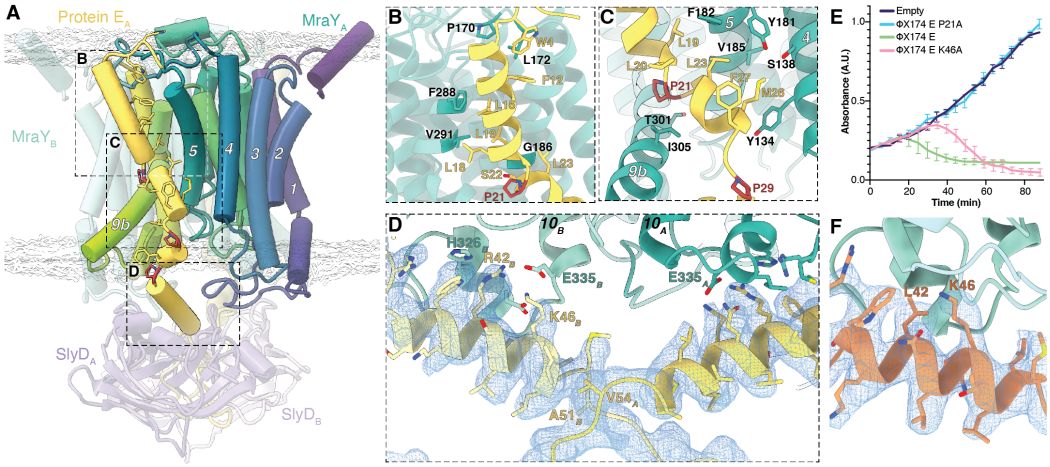


Figure 2.8: **The interaction of protein E with *Ec*MraY**. **A,** Cartoon of the YES$_{ID21}$ complex as in Fig. 2.3e with a 90°rotation except that MraY$_A$ is colored with the Viridis color scheme (Saladi et al., 2020). Prolines are shown in red. Bilayer is shown in white. The B-subunits and SlyDs are faded. Boxes indicate regions highlighted in panels (**B**-**D**). Side chains for protein E residues that contact MraY are shown as sticks. **B,** As in (**A**) with protein E resistance mutants in MraY highlighted (dark cyan). **C,** As in (**A**) with residues at the interface highlighted as sticks. **D,** The region where the two protein E molecules cross highlighting the asymmetry colored as in (**A**) with interacting residues shown as sticks. Density for the amphipathic helix of protein E shown as a blue mesh. SlyDs are removed for clarity. **E,** Lysis assay of expressed protein E$_{\Phi X174}$ variants. **F,** Similar to (**D**) for protein E$_{\Phi X174}$.

The interaction of protein E with MraY

Based on the clear density for protein E, we can analyze the results from previous work. Functional studies have consistently revealed the requirement for a proline at position 21 (Tanaka & Clemons Jr, 2012; Witte et al., 1997; Zheng et al., 2008). The structure allows an elegant explanation for this requirement. The TMD binds in the groove connecting to the active site of MraY initially entering under the kink of TM9 (Fig. 2.8A). The proline at position 21 breaks the hydrogen bonding of the TMD to facilitate the bend around TM9b following the groove to the active site. Mutation of P21 to even a small change, such as alanine, would result in loss of the kink favoring a straight TMD that could not bind in the groove. Residue P29 is the other completely conserved proline (Fig. 2.3A & 2.1) and allows for a second kink that completes the wrap around TM9. A P29A mutation also leads to a lysis phenotype supporting the importance of the shape complementarity (Tanaka & Clemons Jr, 2012; Witte et al., 1997). Additional protein E alanine mutations in the TMD identified residues that result in a slower rate of lysis by protein E, postulated to be due to lowering binding efficiency to MraY (Tanaka & Clemons Jr, 2012). In the structure, most of these residues (L19, L20, L23, & M26) (Fig. 2.8C) make direct contact with MraY. The other mutant, F27, appears to sterically position M26 into a tight interaction with Y134 in MraY, which is conserved in most Gram-negative bacteria (Fig. 2.9).

At the cytoplasmic interface, protein E residues A36 through M50 form an amphipathic helix spanning the width of an MraY subunit. The hydrophilic face of this helix points towards the membrane in the MraY active site. The helix contains conserved positive charged residues that interact with conserved negative charged residues in MraY (Fig. 2.8D). An example is the K46 salt bridge where, in our lysis assay, a K46A mutation results in delayed lysis (Fig. 2.8E). This supports the importance of this region for inhibition in the wild-type context. While the two protein E isoforms are very similar, this region highlights an interesting difference. Position 42 is a leucine in ΦX174 while it is an arginine in ID21 (Fig. 2.3). In the latter, R42 forms a salt bridge interaction with E335 in MraY likely providing additional stabilization. It is in this region that the asymmetry can be highlighted where the crossover occurs at residue V54 in protein E_A and A51 in protein E_B. Near the crossover, there are differing interactions with residues in MraY, such as the essential H326.

Two mutations in protein E were identified that allowed phage propagation in a

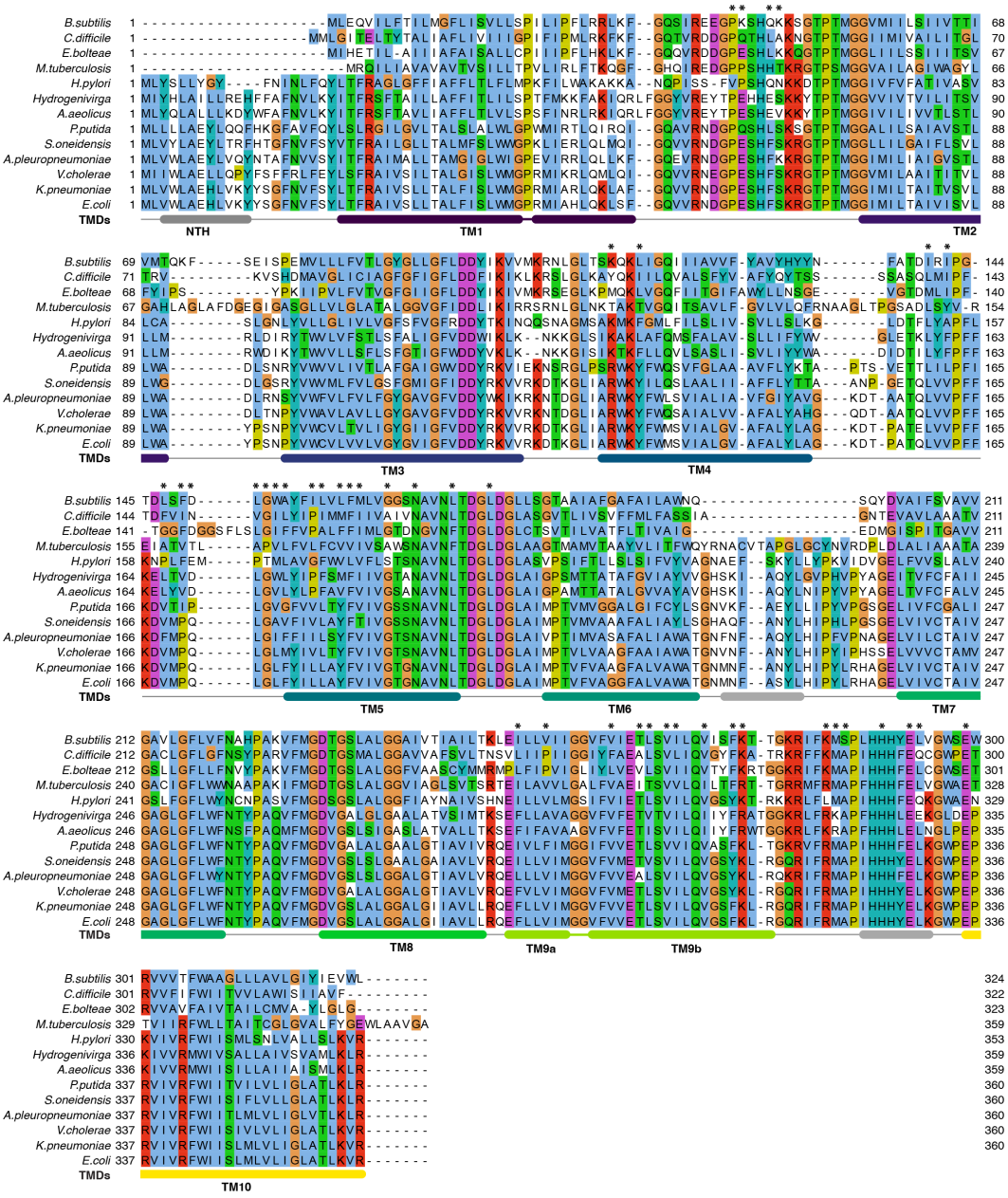


Figure 2.9: **Multiple sequence alignment of representative MraY homologs.** Structure based multiple sequence alignment using Promals3D with ClustalW coloring for residues. Species were selected by using the phylogenetic tree selecting from the cutoff shown in Fig. 2.15. Uniprot ID of species (in order): O26830, Q9WY77, Q03521, Q182Y8, P9WMW7, O25235, A8UQI5, Q88N79, Q8E9P5, B0BRH4, Q9KPG4, A6T4N0, P0A6W3. Secondary structure based *Ec*MraY from the YES complex structure is shown below the sequences, TMDs are colored as in Fig. 2.8a. Side chains in MraY that contact protein E are labeled with a (*).

Δ*slyD* background, R3H and L19F, termed Epos (**P**laques **o**n Δ*slyD*) (Bernhardt et al., 2002). The protein E R3H mutant in ΦX174 results in a silent mutation in Protein D and is found native to other species such as ID21 (Fig. 2.10). Previous work had supported that this variant results in higher levels of protein E in the membrane (Bernhardt et al., 2002). In the structure (Fig. 2.6G & H), this residue does not make specific contacts and it is likely that loss of the arginine would reduce the positive charge in the periplasm favoring the correct orientation of the TMD due to the positive-inside rule (von Heijne, 1986).

Additionally, L19F does not result in higher levels of protein E and it likely increases the affinity of protein E to MraY (Tanaka & Clemons Jr, 2012). Consistent with the structure, phenylalanine residues within transmembrane domains are known to be stabilizing by filling in pockets between side chains in neighboring helices (Steindorf & Schneider, 2017). Another phenylalanine mutation at the interface, L23F, increases the rate of lysis as well; although this is not a general rule as other leucine to phenylalanine mutants did not improve lysis rates (Tanaka & Clemons Jr, 2012). Both L19 and L23 are near the conserved F182 in MraY and may add additional stability through aromatic π interactions (Chourasia et al., 2011) (Fig. 2.8C). The opposite mutation can show loss of binding, for example, the F288L mutation in MraY (Bernhardt et al., 2000) is at the interface with protein E and results in a loss of lysis, likely due to lower affinity. While the L19F mutation in protein E should be favored, as it improves the rate of lysis, it is only found in one species (Fig. 2.1). An explanation may be that this mutation results in an amino acid change in Protein D that would be unfavorable for the virus, discussed below.

The structure allows for a clear explanation of the mutations in MraY that allow resistance to protein E mediated lysis (Bernhardt et al., 2000; Zheng et al., 2008). All of the mutations in MraY (P170L, ΔL172, G186S, F288L, V291M) are at the interface with protein E (Fig. 2.8B). Mutant F288L, as noted above, likely results in lowered affinity with protein E. Residue G186 is at the nearest approach between the two proteins and a mutation to serine would prevent protein E binding. Mutant V291M lies directly at the interface near L19 in protein E; although a specific effect for this mutation is not clear. Finally, P170L and ΔL172 are located within the periplasmic loop 4-5 and, although more resistant to lysis, they are predicted to still bind protein E, albeit with lower affinity (Zheng et al., 2008).

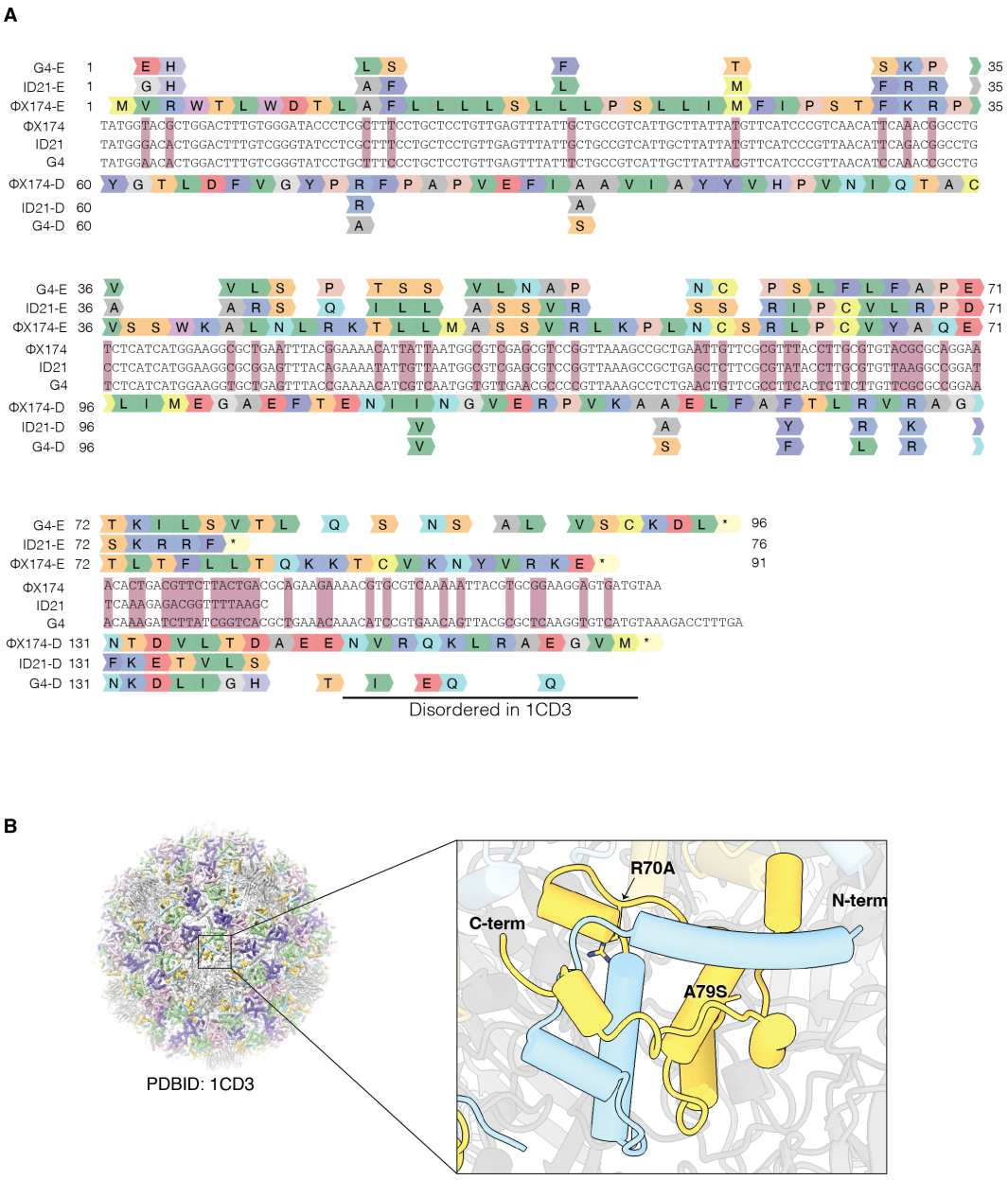


Figure 2.10: **Sequence variability for the coding region of Proteins E from ΦX174, ID21, and G4. A,** The DNA sequence of gene E and the overlapping gene D in the bacteriophages ΦX174, ID21, and G4. Nucleotide differences in the sequences are highlighted in dark pink. Translated protein sequence for ΦX174 protein E is shown above the DNA sequences and Protein D. Amino acids are colored from Benchling (Biology Software,2022). For ID21 or G4, only residues that differ from ΦX174 in the two proteins are shown. **B,** Left, the structure of the ΦX174 pro-virus capsid (PDB:1CD3) with proteins shown as cartoon and colored to highlight the symmetry related Protein Ds (Dokland et al., 1999). Right, a zoomed in view of one Protein D monomer colored either light blue or yellow for the region that overlaps with protein E in yellow. The positions where sequence changes in protein E result in sequence changes in Protein D are shown as sticks.

The mechanism of inhibition of MraY by protein E

The YES complex structure allows us to propose a simple mechanism for inhibition of MraY by protein E. Previous structures of MraY revealed a likely path for the lipid substrate (Hakulinen et al., 2017; Mashalidis et al., 2019). Modeling small molecules on the YES complex (Fig. 2.11) shows the predicted path of the isoprenyl chain of C55P is the groove formed between TM5 and TM9, which is occluded by protein E. Therefore, one mechanism of inhibition is that protein E prevents access of the lipid substrate to the active site (Fig. 2.11A). Protein E is a noncompetitive inhibitor of Park's Nucleotide (Zheng et al., 2009) and, consistent with this, the pocket that binds the nucleoside is fully accessible in the structure (Fig. 2.11B). Loop 9-10 in MraY contains catalytic histidines that must move toward the binding pocket to facilitate catalysis by completing the active site (Al-Dabbagh et al., 2008). The cytoplasmic helix of protein E separates loop 9-10 from the rest of the active site blocking this transition providing a second mechanism of inhibition. Overall, protein E blocks access of the lipid substrate and prevents formation of the active site upon substrate binding.

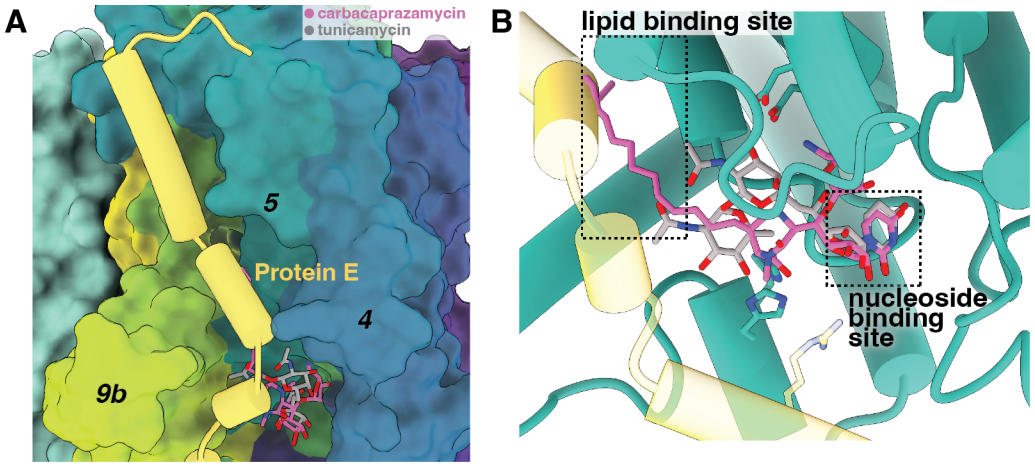


Figure 2.11: **The mechanism of inhibition by protein E**. **A,** Accessible surface of MraY (colored as in Fig. 2.8A) viewed from the cytoplasm looking towards the active site cleft. Protein E is shown in cartoon. The inhibitors tunicamycin (gray) and carbacaprazamycin (pink) are modeled as sticks based on their respective complex structures (PDBID:5JNQ & 6OYH). **B,** Similar to **A** from a slightly different angle highlighting the catalytic pocket. MraY (dark cyan) is shown in cartoon. The two substrate binding sites are highlighted by dashed boxes. Predicted catalytic residues in MraY are shown as sticks.

Key structural features of MraY

The YES complex contains the first *E. coli* MraY structure (Fig. 2.12) which is also the first from either an important human pathogen or model organism. Overall, comparing the YES complex MraY to those solved by crystallography, there is significant agreement when comparing monomers with RMSDs around 1Å(Fig. 2.13A & 2.14). Our EM structure of *Ec*MraY reveals additional regions of MraY previously disordered in the crystal structures. Most notably, we are able to resolve all of the cytoplasmic loops that enclose the active site. We model for the first time loop 1-2, which likely adopts distinct conformations during the catalytic cycle and indeed has two slightly different conformations in our dimer (Figs. 2.7C). In this structure, loop 9-10 adopts a conformation that is similar to the other inhibited MraY structures known to date (Chung et al., 2016; Hakulinen et al., 2017; Mashalidis et al., 2019), distinct from the uninhibited structure, and may be a general feature of an inhibited MraY (Fig. 2.13 & 2.14).

An unexpected structural feature in *Ec*MraY occurs at the N-terminus. In published structures, the N-terminus either begins at TM1 or is a helix that projects away from the structure in an orientation incompatible with the bilayer (Fig. 2.13B). In the YES complex, the N-terminus of the first helix (NTH) hydrogen-bonds to the C-terminal end of TM2, effectively helical stacking (Fig. 2.13B & 2.12C). This is a unique structural feature that, to our knowledge, has not been observed in any protein before. A multiple sequence alignment across bacteria shows that this feature of MraY is conserved across Gram-negative bacteria, but missing in Gram-positives (Fig. 2.9 & 2.15). While this feature is not found in the crystal structures, AlphaFold predicts N-terminal helical stacking for *E. coli* and other Gram-negative bacteria, although the hydrogen bonding and orientation is slightly different from the EM structure (Fig. 2.15) (Jumper et al., 2021). For the *Aa*MraY structures, the positioning of the N-terminus is likely a product of crystallization, as AlphaFold predicts N-terminal helical stacking for this and the related *Hydrogenivirga* species (Fig. 2.15) (Chung et al., 2013; Mashalidis et al., 2019). For Gram-positive bacteria, both the *Eb*MraY and predicted structures begin with a slightly longer TM1 and lack the N-terminal helical stacking (Fig. 2.13B & 2.15) (Hakulinen et al., 2017).

Protein E as a general antibacterial protein

Protein E arrived late in the evolution of ΦX174 and was overprinted into a +1 reading frame in the ORF for gene *D* (Fiddes & Godson, 1979; Pavesi, 2021). The structure supports that this embedding constrains the evolution of protein E (Barrell

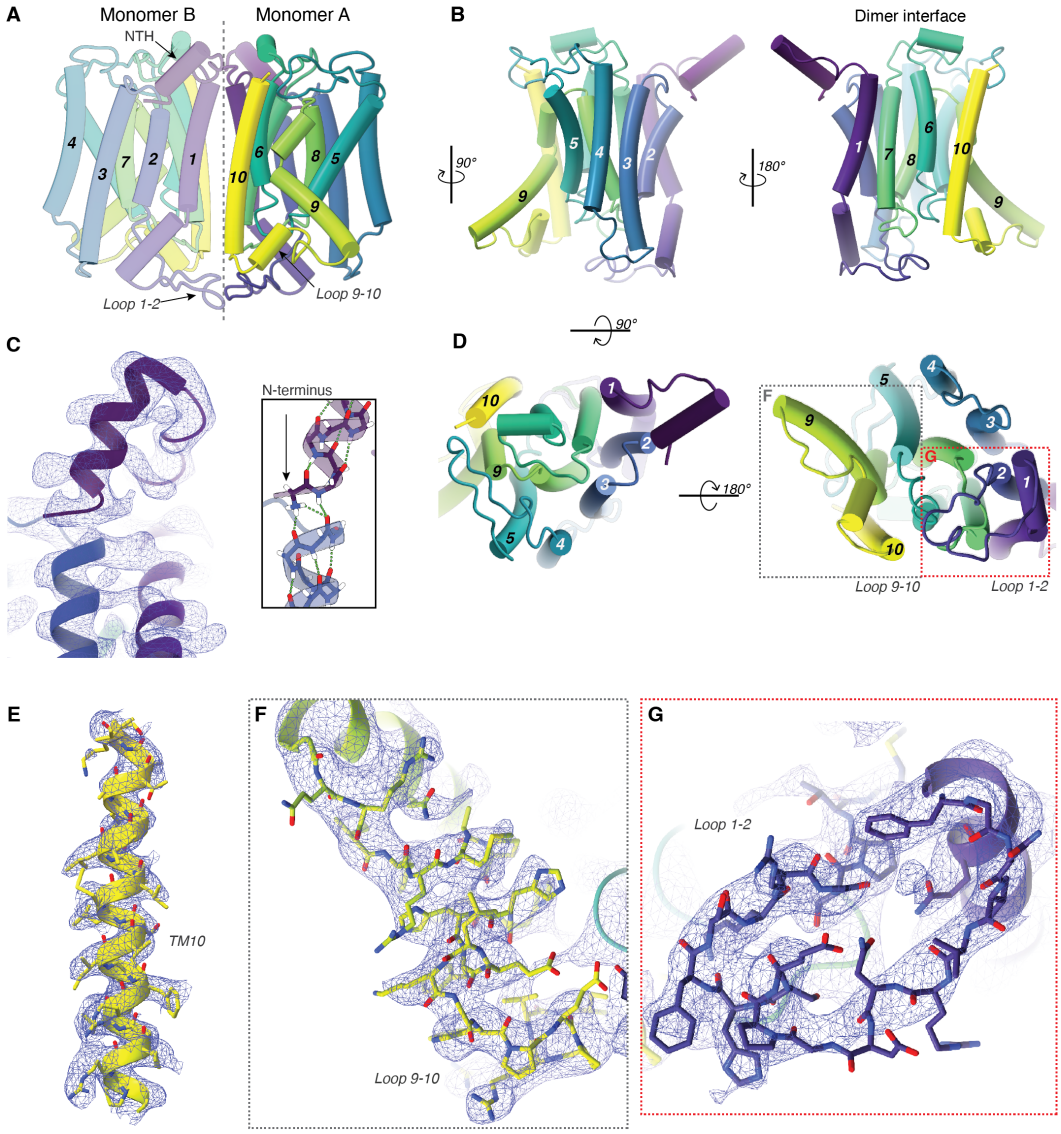


Figure 2.12: **Structural features of *Ec*MraY A,** Cartoon representation of *Ec*MraY colored in Viridis by monomeric unit A and B. Foreground transmembrane helices are labeled 1-10. Loops 1-2, 9-10, and the NTH are labeled for reference. **B,** Monomer of MraY viewed from the active site cleft (left) or the dimer interface (right). **C,** Cartoon representation of the N-terminal helix of MraY with N-terminal and TM2 residues shown as sticks. Density map is shown as a blue mesh. **D,** Periplasmic (left) and cytoplasmic (right) view of the MraY monomer. Transmembrane helices are labeled. **E,** Cartoon representation with stick side chains of TM10. Densities shown as a blue mesh. **F,** Stick model of Loop 9-10 with density (blue mesh) oriented as box in (**D**). **G,** As (**F**) for Loop 1-2 .

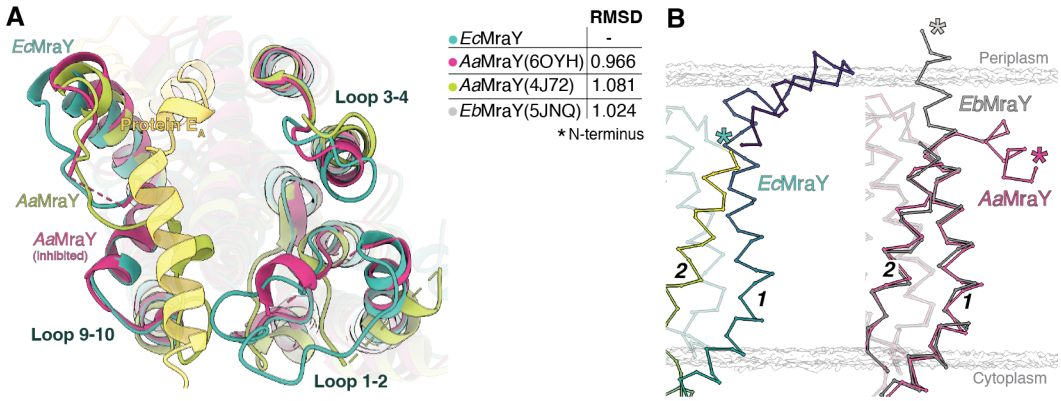


Figure 2.13: **New features observed in the EM structure of *E. coli* MraY. A,** Cytoplasmic view of a structural alignment of *Ec*MraY (dark cyan) against uninhibited *Aa*MraY (green, PDB:4J72) and carbacaprazamycin inhibited *Aa*MraY (pink, PDB:6OYH). RMSDs to monomer A of *Ec*MraY are shown. The color scheme for the various MraY crystal structures is used throughout the figures. **B,** A view in the plain of the membrane showing the region that includes TM1 and TM2 in backbone ribbons. Left, the *Ec*MraY structure colored in Viridis. Right, the two crystal structures shown in (**A**) and the crystal structure from *Eb*MraY (green, PDB:4J72). Each structure is aligned to *Ec*MraY. The location of each N-terminus is indicated by an asterisk.

et al., 1976). Considering the sequence changes across protein E isoforms, we note that from the N terminus through residue 70, with few exceptions, each position that is not completely conserved is either silent or that there is only a slight change in protein D (Fig. 2.10A). The silent changes vary the codon's second position, which is the wobble position in the overlapping codon for protein D. For example, residue W7 in ΦX174 is replaced with a Ser or Leu in other species. Although seemingly substantial changes, each is coded by the sequence UXG and all three variations at this position are sampled. These variable positions generally do not contact MraY. An exception is found in the G4 isoform where a phenyalanine occurs at position 19, as in the Epos mutant, which results in a change from alanine to serine at position 79 in protein D. This places a polar residue in the hydrophobic core of protein D (Fig. 2.10B). This single nucleotide change increases protein E affinity to MraY but likely lowers the stability of protein D and the overall fitness of the virus. Introduction of this mutation results in smaller plaque formation relative to the wild type (Bernhardt et al., 2002). The C terminus of protein D and the extensions in the longer isoforms of protein E are likely disordered and less constrained, hence the increased sequence variability (Fig. 2.10A & B).

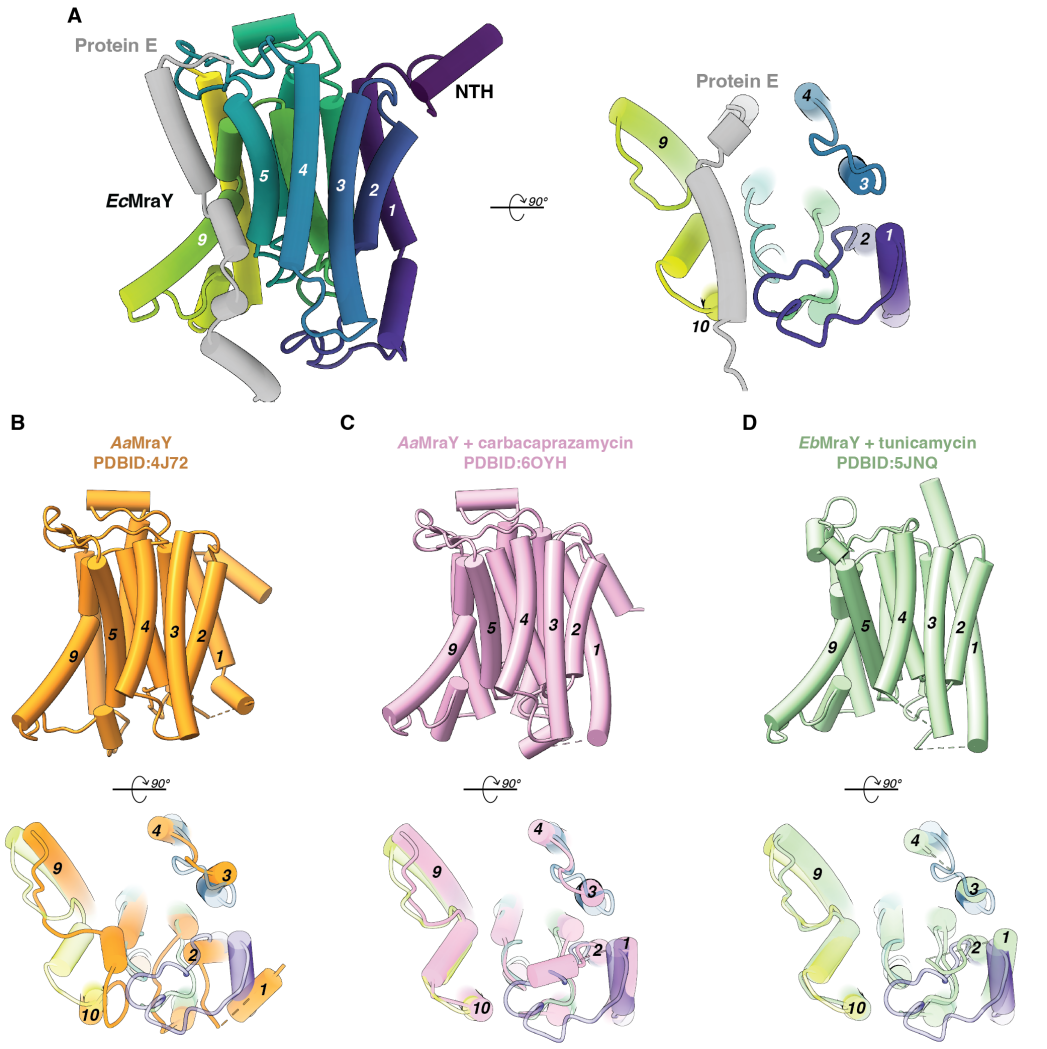


Figure 2.14: **Structural alignment of experimental MraY structures.** For each panel, viewed from the active site cleft and from the cytoplasm. **A,** *Ec*MraY is shown in Viridis and transmembrane domains are labeled. Protein E_{ID21} is shown in gray. **B,** Uninhibited *Aa*MraY (PDB:4J72) in orange. **C,** Carbacaprazamycin inhibited *Aa*MraY (PDB:6OYH) in pink. **D,** Tunicamycin inhibited *Eb*MraY (PDB:5JNQ) in green.

The role of SlyD in protein E mediated lysis

The amphipathic helix of protein E bridges the two *E. coli* proteins, MraY and SlyD, which make no specific contacts to each other (Fig. 2.8A). The IF domain sits on the hydrophobic face of this protein E helix and contacts the extended C-terminus of the opposing protein E which then continues on to bind the FKBP domain. This results in a bow-tie like interaction with each SlyD binding both protein E soluble domains (Fig. 2.16B). These interfaces validate structural studies of SlyD where

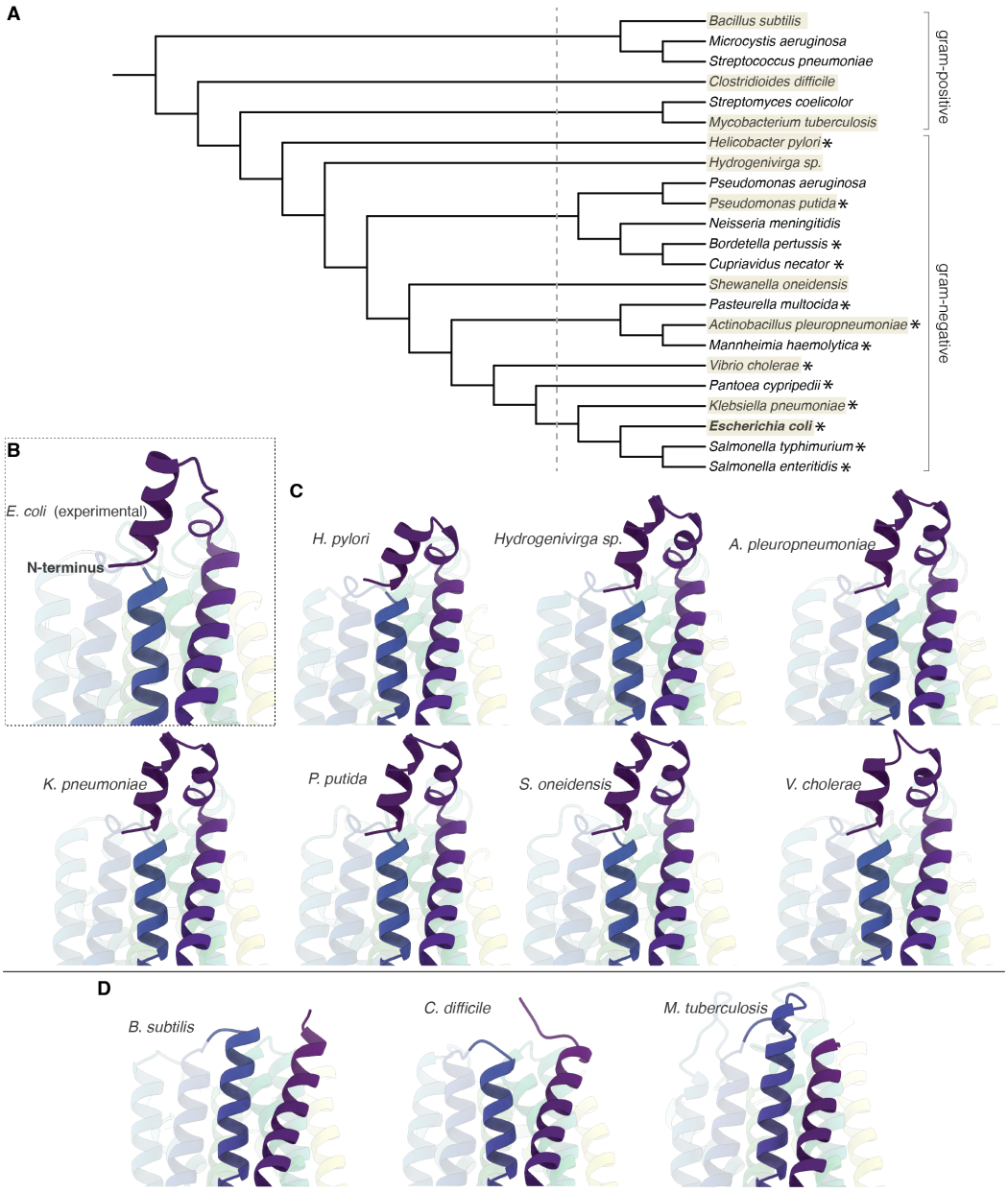


Figure 2.15: **Conservation of the N-terminal helix stacking.** **A,** Phylogenetic tree of MraY from representative bacterial species. Species demonstrated to from ghosts when protein E is expressed are highlighted by an asterisk (*). Selected species are highlighted in yellow. **B,** *Ec*MraY cryoEM structure colored in Viridis with the N-terminal helices highlighted. **C,** AlphaFold prediction of MraY structures from Gram-negative species oriented and shown as in (**A**). **D,** AlphaFold prediction of MraY structures from Gram-positive species oriented and shown as in (**A**).

extended peptides bind to each of the interfaces seen here including β-augmentation in a groove in the IF domain (Fig. 2.16C & D) (Quistgaard et al., 2016). This is the first evidence of the IF domain binding to an α-helix, which may be an important chaperoning interaction. The FKBP domain accounts for the highest variability in our particles (Fig. 2.17) and the lowest resolution (Fig. 2.4). This flexibility is consistent with NMR structures of SlyD (Weininger et al., 2009) and the only contact of this domain to the rest of the YES complex is the flexible linker to the IF domain and binding to the extended C-terminus of protein E. At the FKBP active site, P65 is not completely conserved (Fig. 2.1), although other species have a proline at residue 63 which could reach the active site with additional tilting of the FKBP domain. A proline near this position may help to localize the chaperone to the complex.

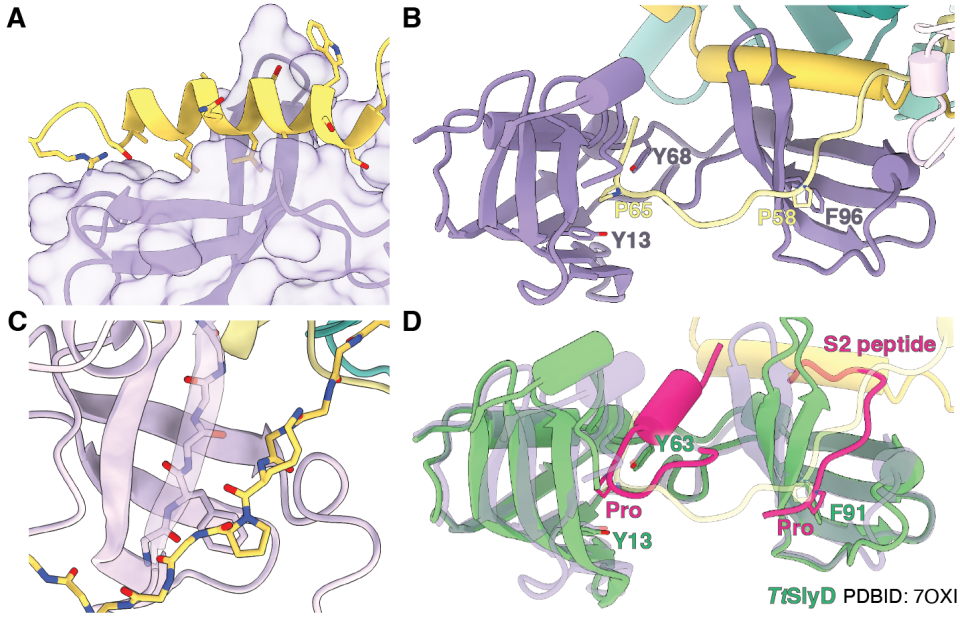


Figure 2.16: **Interactions between protein E and SlyD. A,** The amphipathic helix of protein E (yellow) as a cartoon with side chains contacting SlyD as sticks. SlyD (purple) is shown as transparent accessible surface and cartoon. **B,** Full view of SlyD bound to two protein E molecules shown in different shades of yellow. **C,** As in **A** highlighting the β-augmentation of the extended C-terminal protein E. **D,** Structures of SlyD from the YES_{ID21} complex and *Thermus thermophilus* SlyD (PDB:7OXI) (green) aligned to the IF domains. The two S2 peptides bound to the *Tt*SlyD are shown in pink.

The lack of contacts between SlyD and MraY suggest the soluble domain of protein E alone could form a complex with SlyD. We co-expressed *Ec*SlyD with N-terminal truncations of protein E from either ID21 (residues 33-76) or the shortest isoform α3

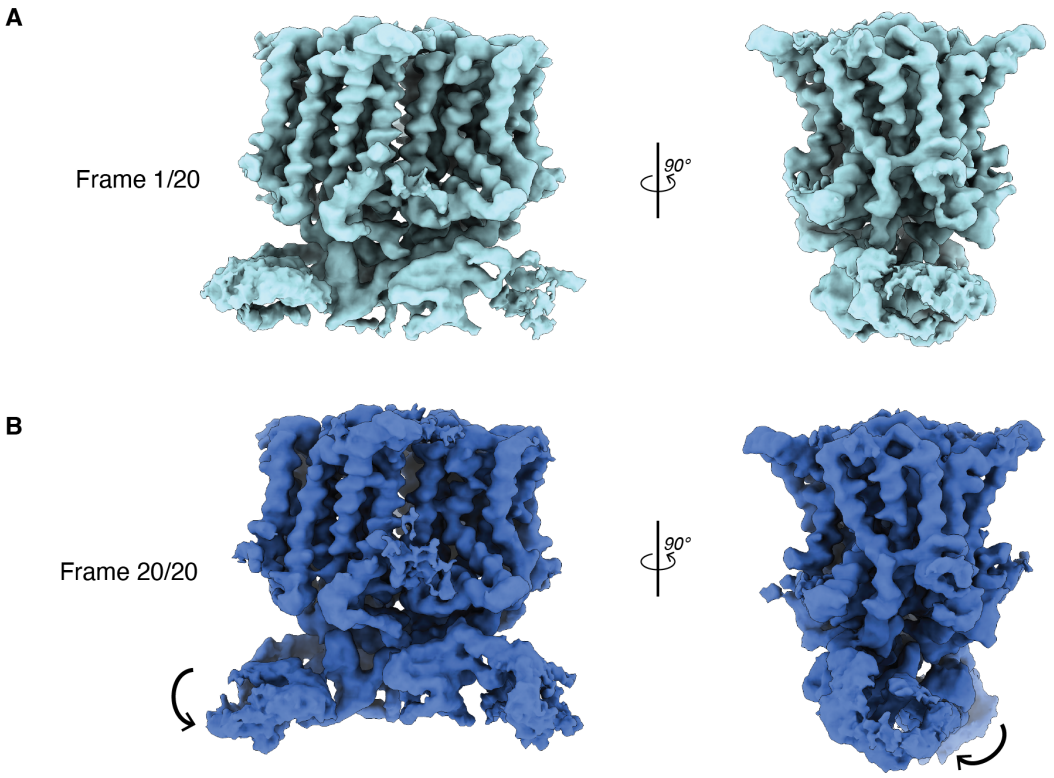


Figure 2.17: **Structural variability in the YES complex.** 3D-variability analysis of the YES complex. Movement is visualized by highlighting the first (**A**) and last frame (**B**).

(residues 33-75). Both form stable complexes that could be purified by an affinity tag on the soluble domain (Fig. 2.2C). These observations point to a high-affinity interaction between SlyD and the C-terminal domain of protein E.

protein E is unstable in the absence of SlyD and rapidly degraded (Bernhardt et al., 2002; Roof et al., 1994). While it has been speculated that prolyl-isomerization was central to the lysis mechanism, evidence points to the contrary. Non-proline mutations in protein E can rescue lysis in a Δ*slyD* background, as well as complete replacement of the cytoplasmic domain of protein E with a globular protein (Buckley & Hayashi, 1986; Roof et al., 1994; Roof & Young, 1995; Tanaka & Clemons Jr, 2012). To explore this we performed our lysis assay with several SlyD variants in the Δ*slyD* strain (Fig. 2.2A). As before, protein $E_{\Phi X174}$ was unable to lyse in the absence of SlyD, however this could be rescued by expression of either *Ec*SlyD or *Ec*$SlyD_{154}$. *Tt*SlyD has high structural homology to *Ec*SlyD maintaining all the same features (Fig. 2.16D) and it was also able to rescue lysis in the Δ*slyD* strain (Fig. 2.2A). We generated a *Ec*SlyD Y68K mutant that dramatically reduces prolyl-isomerase

activity (Ikura & Ito, 2007) and this too could rescue lysis. Finally, we purified a complex of protein E using the Epos rescue mutants (R3H & L19F) in the Δ*slyD* strain; however, this complex was very unstable and showed significant aggregation by SEC (Fig. 2.2C) (Roof & Young, 1995) . This all supports that the primary role of SlyD is not prolyl-isomerization, but to protect protein E and stabilize the YES complex. This does not obviate a role for proline binding in complex assembly. In fact, the two prolines in the soluble domain have a lysis phenotype which may indicate an importance of the FKBP domain in the initial assembly of the YES complex (Witte et al., 1997).

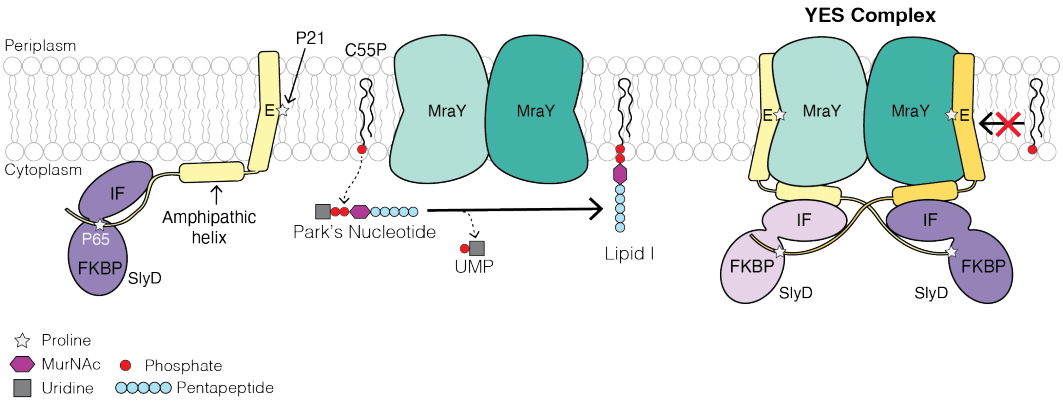


Figure 2.18: **Model for inhibition by protein E.** protein E (yellow) would be targeted to the membrane through the normal secretory pathway. Upon insertion, it must remain stable long enough to find and bind to MraY (cyan). During this time SlyD (purple) likely binds to the C-terminus of protein E through proline residues recognized by the SlyD FKBP domain. In the absence of protein E, MraY catalyzes the transfer of a phospho-MurNAc-pentapeptide from Park's Nucleotide onto lipid C55P, producing Lipid I. The protein E-SlyD complex eventually arrives at MraY and forms a stable dimer of heterotrimers. The binding of protein E prevents access of C55P to the active site of MraY. Symbols used throughout the scheme are defined in the bottom left. Stars highlight key proline residues P21 and P65 (E_{ID21}). N-acetylmuramic acid is depicted as a flattened hexagon. Uridine is shown as a square, phosphates as red circles, and the pentapeptide as five blue circles.

2.3 Discussion

We have determined two structures of an inhibited complex of the peptidoglycan biosynthesis enzyme MraY bound to the viral protein E and a native chaperone SlyD. The structures reveal detailed interactions in the YES complex and provide a clear mechanism for inhibition of MraY by protein E from the historically important ΦX174 and the entire *Bullavirinae* subfamily. The mechanism for inhibition is remarkably simple. Protein E binds in the active site cleft preventing access of the lipid substrate and blocking conformational changes needed for catalysis (Fig. 2.18).

The YES complex structure provides a template for interpreting all of the previous results for lysis by ΦX174. Notable among these are that all of the functional mutations in protein E and MraY map to the interface between the two proteins. Protein E exploits the C55P binding site by complementing the active site cleft in MraY. SlyD binds to the cytoplasmic domain of protein E to stabilize the inhibited MraY/protein E complex. As we demonstrated, the requirement for SlyD is not species dependent, with *Tt*SlyD supplementation in a Δ*slyD* background being

sufficient to restore lysis. SlyD in this case does not serve directly in the mechanism of inhibition of MraY, and can be functionally replaced.

While there remain many important functional questions about MraY, the *Ec*MraY structure by cryoEM provides new insight and allows access to previous features that were disordered in the crystal structures. The NTH is an example of this, and this unique structural feature merits further study on its own. The resolving of loop 1-2 and the inhibited conformation of loop 9-10 provide additional mechanistic insight. The many lipid densities in our map support the importance of these in the function of MraY including supporting evidence for a C55P binding site near the dimer interface (Oluwole et al., 2022). While providing no further resolution to the identity of the molecule in the periplasmic cavity, the density supports that it is an ordered hydrophobic molecule and identification of this molecule will likely have important functional implications.

One of the most remarkable features of protein E is its simplicity. During the evolution of a phage, a lysis mechanism likely evolves late, as phage that have their lysis genes removed can still propagate, albeit much less efficiently (K. R. Chamakura & Young, 2020). This can be seen in action in ssRNA phages whose hosts vary due to their receptor being plasmid-borne pili which can transfer across species and, therefore, require new lysis genes in new hosts (Kannoly et al., 2012; Rumnieks & Tars, 2012). For these phage, new lysis genes rapidly develop and can overprint across the entire genome (K. R. Chamakura et al., 2020). For protein E, it clearly evolved late likely due to the introduction of a ribosome binding site (Pavesi, 2021). This new peptide would be inefficient, but constrained against improvement by being embedded in an essential gene. A feature of protein E is that it is rapidly degraded, likely due to its disordered C-terminus, and SlyD binding would protect this region from the degradation machinery. We know that simple mutations, such as L19F, can improve the efficiency of lysis. All of this points to the fact that protein E can likely be dramatically improved to become more efficient at inhibiting MraY.

The YES complex structure sheds light on the distinct ΦX174 lysis mechanism first observed over 50 years ago. While this mode of SGL inhibition of a membrane protein is novel, evidence supports that there will be many more examples. SGL proteins provide a unique route towards developing tools towards killing bacteria. While direct use of hydrophobic peptides as drugs is not likely, one can imagine several ways in which these peptides can immediately be used towards the development of therapies. The first is that protein E binding to *Ec*MraY identifies additional MraY

pockets that can be exploited by small molecules and provides a new template for structure based drug design. Next, with the need for tools towards killing synthetic bacteria, the small SGL peptides will be potent programmable genetic kill switches. One can imagine that a lysis operon encoding several of these in parallel can be designed to turn on in response to an external signal. Finally, the most obvious utility is in the optimization of phage therapeutics. These genetically efficient SGLs will be important tools for optimizing phage potency. The use of phage for medical therapies, while known for a hundred years, has only recently become possible in the West (Schooley et al., 2017). The desperate need to combat the dangerous rise of antibacterial resistance will include phages. SGL genes, such as protein E from ΦX174, will play an important role.

2.4 Methods

Co-Expression of *Ec*MraY, protein E and *Ec*SlyD

Δ*slyD* BL21(DE3) competent cells were co-transformed with pET22b-$SlyD_{154}$ and either pRSFDuet-*Ec*MraY-E_{ID21} or pRSFDuet-*Ec*MraY-$E_{\Phi X174}$ and plated in LB-agar containing 35 μg/ml Kanamycin and 100 μg/mL Ampicillin. Our pET22b-$SlyD_{154}$ construct expresses *E. coli* SlyD, modified by the removal of the flexible C-terminus. The pRSFDuet-*Ec*MraY-E_{ID21} plasmid contains the ID21 isoform of protein E, along with a wild-type *Ec*MraY to prevent cell lysis from the overexpression of protein E. Cells were grown in 2xYT media at 37°C, 225 r.p.m., and induced at an OD_{600} of 0.9 with 0.4mM IPTG at 18°C overnight. The culture was harvested by centrifugation for 10 minutes at 9,000*xg*, 4°C then frozen or used immediately for purification.

Purification of the YES complex

The cells were resuspended in lysis buffer (20mM Tris-HCl pH 7.5, 300 mM NaCl, 10% Glycerol, 5mM βME, 0.1mM PMSF, 0.1mM Benzamidine) and homogenized using a M-110 L microfluidizer (Microfluidics). The lysate was cleared by a 20 minute centrifugation at a speed of 22,000*xg*. The supernatant was then centrifugated at 167,424*xg* and the resulting membrane pellet was then solubilized in the extraction buffer (10 mM HEPES pH 7.5, 300 mM NaCl, 5% Glycerol, 5mM β mercaptoethanol (BME), 0.1mM phenylmethylsulfonyl fluoride, 0.1mM benzamidine, 10 mM imidazole and 1% dodecyl 4-O-α-D-glucopyranosyl-β-D-glucopyranoside (DDM)) After allowing for extraction for 1.5 hours at 4°C, the solution was centrifuged at 167,424*xg* for 30 minutes and the remaining lysate was mixed with 1mL NiNTA resin (Qiagen, Alameda, CA) then nutated at 4°C for two hours. This solution was loaded onto a gravity column and then washed with five column volumes of wash buffer (10 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, 5mM BME, & 0.03% DDM) with 10mM imidazole followed by five column volumes of wash buffer with 30 mM imidazole. The YES complex was eluted in 20mL of wash buffer containing 200 mM imidazole. The final purification step was SEC (Superdex 200 5/150 GL, Milipore Sigma) in 10mM HEPES pH 7.5, 75 mM NaCl, 5% Glycerol, 5mM BME and 0.03% DDM. Fractions were assessed by SDS-PAGE and directly used for cryo-EM sample preparation.

Co-expression of *Ec*MraY and protein E in various SlyD backgrounds

The pRSFDuet-E-EcMraY and pRSFDuet-Epos-EcMraY expression vectors were transformed into BL21-Star cells (Novagen). Similarly, the pRSFDuet-E(C-term)-$SlyD_{154}$ was transformed into SlyD-knockout cells. The cultures were grown at 37 °C to an OD_{600} 0.8 and induced with 1 mM IPTG. Induced cultures were grown for 3 hours followed by harvesting by centrifugation at 9,000*xg* for 20 min. Cell pellets were resuspended in lysis buffer and lysed by sonication. The lysate was then cleared by centrifugation at 22,000*xg*, followed by a second centrifugation at 234,78*xg* for 1 hour to isolate the membrane fraction. The complex was extracted in 20 mM Tris-HCl pH 7.5, 300 mM NaCl, 10% Glycerol, 10 mM Imidazole, and 1% n-Decyl-β-Maltoside (DM) and incubated at 4 °C for 1.5 hours. The debris was cleared by centrifugation at 234,788*xg* for 30 min. The sample was incubated with 1 mL NiNTA resin for 1 hour, followed by a wash with 50 column volumes lysis buffer with 30mM Imidazole. The protein E complexes were similarly eluted in 300mM Imidazole. The elutions were concentrated and further purified by size exclusion chromatography (Superdex 200 5/150 GL, Milipore Sigma).

Lysis assays of WT protein E ΦX174 and ID21

LEMO DE3 competent cells were transformed with a pRSF-Duet vector either empty, with protein $E_{\Phi X174}$, or protein E_{ID21}. Cultures were grown to an OD_{600} of 0.2 and inoculated into a Corning 96-well Clear Flat Bottom plates in 100μL triplicate aliquots and induced as described previously. Cultures were incubated at 37C with orbital shaking at 220rpm using an Infinite M Nano+ (Tecan, Switzerland). Readings were taken in 5 minute intervals for 90 minutes.

Lysis assay for protein E constructs

LEMO DE3 competent cells (New England Biolabs, MA, USA) were transformed with a pRSF-Duet vector either empty, with C-terminally FLAG tagged protein $E_{\Phi X174}$ variants (WT, P21A, K46A). The lysis assays were performed in triplicates as previously described (Tanaka & Clemons Jr, 2012). Absorbance readings were recorded in 5 minute intervals for 1 hour and 30 minutes. Manual readings were taken using a Biowave Cell Density Meter CO8000. The values were plotted using GraphPad Prism version 9.1.1 for macOS.

Lysis assays based on SlyD variants

Δ*slyD* (Roof et al., 1994) cells were transformed with either a control empty pRSF-Duet vector or pRSF-Duet-Protein$_{\Phi X174}$ and either pET22b-*Ec*SlyD, pET22b-SlyD$_{154}$, pET22b-*Ec*SlyD Y68K, or pET22b-*Thermus thermophilus* SlyD. Cultures were grown in 2xYT media at 37°C and induced with 0.4mM IPTG once at an OD_{600} of 0.2. Absorbance measurements were manually recorded in 5 minute intervals for 70 minutes. Similarly, Δ*slyD* cells were transformed with either a control empty pRSF-Duet vector or pRSF-Duet-Protein$_{ID21}$ either alone, with pET22b-*Ec*SlyD, or with pET22b-*Ec*SlyD$_{154}$ and induced with 0.4mM IPTG. Readings were recorded using an Infinite M Nano+ plate reader as described above.

Sample preparation for CryoEM

The YES complex was diluted to 5.0 mg/mL in 10 mM HEPES ph 7.5, 75 mM NaCl, 5% Glycerol, 5mM βME and 0.03% DDM, supplemented with 1mM *E. coli* total lipid extract (Avanti Polar Lipids, 100600P). Quantifoil holey carbon films R1.2/1.3 300 Mesh, Copper (Quantifoil, Micro Tools GmbH) grids were glow discharged with a 2 minute 20Å plasma current using a Pelco easiGlow, Emitech K100X. Grids were prepared using a Vitrobot (FEI Vitrobot Mark v4 x2, Mark v3) by appling 3μL of sample onto the grid followed by a 3.5 second blot using a +8 blot force and plunge frozen into liquid ethane.

Data acquisition and analysis

The grids were imaged in a 300 kV cryo-TEM microscope equipped with a Gatan K3 6k x 4k direct electron detector and a Gatan Energy Filter (slit width 20eV) in super-resolution mode using Serial EM. Data-sets were collected at a 105 k magnification with a pixel size of 0.416 Å/pixel. Movies with 40 frames were recorded with a total expousre dose of 60 $e^-/Å^2$ and a defocus range of -1.0 to -2.0 μm. A total of 12,070 movies were recorded. Movies were normalized by gain reference and motion corrected using the patch motion correction built in function in cryosparc (v3.3.2) with a twofold bin that resulted in a pixel size of 0.832 Å/pixel. The contrast transfer function (CTF) was estimated using CTFFIND4 (Rohou & Grigorieff, 2015). Micrographs were manually curated, and low-quality images were removed for further analyses. A total of 2,462,335 particles were obtained followed by the generation of 6 ab-initio models. Out of the 6 models, two models are selected for classification into "good" and "trash" volumes. All of the particles were then sorted in these two volumes through heterogeneous refinement

using particles extracted with a 4x bin, which produced 6,589,696 good particles. Heterogeneous refinement was used in an iterative manner to sort the particles into the 5 volumes (4 good and 1 trash). The 1,151,777 good particles were used for non-uniform homogeneous refinement to generate a higher resolution volume. The particles were then extracted with a 3x bin and sorted into 4 iterations of the higher resolution volume and 1 trash volume. Iterative rounds of heterogeneous refinement at 3x bin produced 935,754 particles. Particles were then extracted in a 2x bin and heterogeneously refined into either high or low resolution volumes. At this point, discerning features in the soluble region of the model were used to select the most complete volumes. The volumes were individually refined through non-uniform refinement and the particles that composed the volumes with most complete and highest resolution were used. A total of 122,452 particles were used for the most complete model obtained upon non-uniform refinement. The half-maps were then used for post-processing through DeepEMhancer (Sanchez-Garcia et al., 2021) with the high-resolution model selected for our most complete density map. Post-processing through DeepEMhancer removed the micelle and improved the features on the soluble portions of the map, however the lipid densities were also removed. The lipid densities described in this work are those observed before post-processing. For further characterization of the transmembrane region of the map, the particles were further refined through heterogeneous refinement using 3 good volumes with C2 symmetry imposed. Particles within the two highest resolution volumes and the highest resolution volume were used for a final non-uniform refinement with C1 symmetry, yielding higher resolution in the transmembrane domains of MraY and protein E. The local resolution of both complete and transmembrane maps was performed on cryosparc (v3.3.2).

Model building

The model for *Ec*MraY was built using PDBID:4J72 as a starting model, and PDBID:2K8I for *Ec*SlyD using phenix.dock. Protein E was modeled de-novo using Coot 0.8.9.2. The structure was refined using phenix.real space refinement and validated with PHENIX-1.19.2. RMSDs were calculated using ChimeraX Matchmaker chain alignment. Structure figures were made using ChimeraX and sequence alignments using Jalview (Pettersen et al., 2021; Waterhouse et al., 2009).

Acknowledgements

We are particularly grateful to Ryland Young and Thomas Bernhardt for providing inspiration and feedback during the course of this project. We thank them both and Doug Rees for critical feedback on the manuscript. We further thank R. Young for providing the Δ*slyD* strain. (Cryo)Electron microscopy was done in the Beckman Institute Resource Center for Transmission Electron Microscopy at Caltech. We are grateful to Songye Chen for help with data collection and processing. This project was funded by grants to WMC from NIGMS (R01GM114611 & 5DP1GM105385) and from The G. Harold and Leila Y. Mathers Foundation.

Table 2.1: **YES Complex - Cryo-EM data collection, refinement, and validation statistics**

	YES_{ID21} (EMDB-29641) (PDB 8G01)	**$YES_{\Phi X174}$** (EMDB-29642) (PDB 8G02)
Data collection and processing		
Microscope	FEI Titan Krios	FEI Titan Krios
Magnification	105,000	105,000
Voltage (kV)	300	300
Electron exposure (e^{-}/Å^2)	60	60
Defocus range (μm)	-1 to -2.5	-1 to -2.5
Pixel size (Å)	0.832	0.832
Symmetry imposed	C1	C1
Initial particle images (no.)	11,700,795	1,516,368
Final particle images (no.)	122,452	155,270
Map resolution (Å)	3.47	3.6
FSC threshold	0.143	0.143
Refinement		
Software	PHENIX 1.19.2	PHENIX 1.19.2
Initial model used (PDB code)	4J72, 2K8I	8G01
Resolution of unmasked reconstructions (Å, FSC=0.5)	3.72	3.8
Resolution of masked reconstructions (Å, FSC=0.5)	3.72	3.79
Correlation coefficient (CCmask)	0.71	0.69
Map sharpening B factor (Å^2)	142.4	148.3
Model composition		
Atoms (Hydrogens)	18153 (9214)	18198 (9232)
Protein residues	1148	1148
Ligands	0	0
B factors (Å^2)		
Protein (min/max/mean)	60.55/203.29/93.87	48.57/311.57/93.33
Ligand	-	-
R.m.s. deviations		
Bond lengths (Å)	0.002 (0)	0.003 (0)
Bond angles (°)	0.436(0)	0.636(2)
Validation		
MolProbity score	1.12	1.53
Clashscore	3.31	10.33
Poor rotamers (%)	0	0.1
Ramachandran plot		
Favored (%)	99.03	98.15
Allowed (%)	0.97	1.85
Disallowed (%)	0	0

References

Al-Dabbagh, B., Henry, X., Ghachi, M. E., Auger, G., Blanot, D., Parquet, C., Mengin-Lecreulx, D., & Bouhss, A. (2008). Active site mapping of mray, a member of the polyprenyl-phosphate n-acetylhexosamine 1-phosphate transferase superfamily, catalyzing the first membrane step of peptidoglycan biosynthesis. *Biochemistry*, *47*(34), 8919–8928.

Barrell, B. G., Air, G., & Hutchison, C. (1976). Overlapping genes in bacteriophage Φx174. *Nature*, *264*(5581), 34–41.

Benbow, R., Hutchison III, C., Fabricant, J., & Sinsheimer, R. (1971). Genetic map of bacteriophage Φx174. *J. Virol.*, *7*(5), 549–558.

Bernhardt, T. G., Roof, W. D., & Young, R. (2000). Genetic evidence that the bacteriophage Φx174 lysis protein inhibits cell wall synthesis. *Proc. Nat. Acad. Sci.*, *97*(8), 4297–4302.

Bernhardt, T. G., Roof, W. D., & Young, R. (2002). The escherichia coli fkbp-type ppiase slyd is required for the stabilization of the e lysis protein of bacteriophage φx174. *Molecular microbiology*, *45*(1), 99–108.

Bernhardt, T. G., Wang, I.-N., Struck, D. K., & Young, R. (2001). A protein antibiotic in the phage qβ virion: Diversity in lysis targets. *Science*, *292*(5525), 2326–2329.

Bouhss, A., Trunkfield, A. E., Bugg, T. D., & Mengin-Lecreulx, D. (2007). The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol. Rev.*, *32*(2), 208–233.

Bradley, D., Dewar, C. A., & Robertson, D. (1969). Structural changes in *Escherichia coli* infected with a Φx174 type bacteriophage. *J. Gen. Virol.*, *5*(1), 113–121.

Buckley, K. J., & Hayashi, M. (1986). Lytic activity localized to membrane-spanning region of ϕx174 e protein. *Molecular and General Genetics MGG*, *204*(1), 120–125.

Bugg, T. D., Braddick, D., Dowson, C. G., & Roper, D. I. (2011). Bacterial cell wall assembly: Still an attractive antibacterial target. *Trends Biotech.*, *29*(4), 167–173.

Chamakura, K., & Young, R. (2019). Phage single-gene lysis: Finding the weak spot in the bacterial cell wall. *J. Biol. Chem.*, *294*(10), 3350–3358.

Chamakura, K. R., Sham, L.-T., Davis, R. M., Min, L., Cho, H., Ruiz, N., Bernhardt, T. G., & Young, R. (2017). A viral protein antibiotic inhibits lipid ii flippase activity. *Nature Micro.*, *2*(11), 1480–1484.

Chamakura, K. R., Tran, J. S., O'Leary, C., Lisciandro, H. G., Antillon, S. F., Garza, K. D., Tran, E., Min, L., & Young, R. (2020). Rapid *de novo* evolution of lysis genes in single-stranded rna phages. *Nat. Comm.*, *11*(1), 1–11.

Chamakura, K. R., & Young, R. (2020). Single-gene lysis in the metagenomic era. *Curr. Opin. Microbiol.*, *56*, 109–117.

Cherwa, J. E., & Fane, B. A. (2011). Microviridae: Microviruses and gokushoviruses. *eLS*.

Chourasia, M., Sastry, G. M., & Sastry, G. N. (2011). Aromatic–aromatic interactions database, a2id: An analysis of aromatic π-networks in proteins. *Int. J Biologic. Macromol.*, *48*(4), 540–552.

Chung, B. C., Mashalidis, E. H., Tanino, T., Kim, M., Matsuda, A., Hong, J., Ichikawa, S., & Lee, S.-Y. (2016). Structural insights into inhibition of lipid i production in bacterial cell wall synthesis. *Nature*, *533*(7604), 557–560.

Chung, B. C., Zhao, J., Gillespie, R. A., Kwon, D.-Y., Guan, Z., Hong, J., Zhou, P., & Lee, S.-Y. (2013). Crystal structure of mray, an essential membrane enzyme for bacterial cell wall synthesis. *Science*, *341*(6149), 1012–1016.

Denhardt, D. T., & Sinsheimer, R. L. (1965). The process of infection with bacteriophage Φx174: Iii. phage maturation and lysis after synchronized infection. *J. Mol. Biol.*, *12*(3), 641–646.

Dokland, T., Bernal, R. A., Burch, A., Pletnev, S., Fane, B. A., & Rossmann, M. G. (1999). The role of scaffolding proteins in the assembly of the small, single-stranded dna virus Φx174. *J. Mol. Biol.*, *288*(4), 595–608.

Duan, Y., Young, R., & Schnabl, B. (2022). Bacteriophages and their potential for treatment of gastrointestinal diseases. *Nat. Rev. Gastroenterol. Hepatol.*, *19*(2), 135–144.

Egan, A. J., Errington, J., & Vollmer, W. (2020). Regulation of peptidoglycan synthesis and remodelling. *Nat. Rev. Microbiol.*, *18*(8), 446–460.

Ferriol-González, C., & Domingo-Calap, P. (2021). Phage therapy in livestock and companion animals. *Antibiotics*, *10*(5), 559.

Fiddes, J. C., & Godson, G. N. (1979). Evolution of the three overlapping gene systems in g4 and Φx174. *J. Mol. Biol.*, *133*(1), 19–43.

Haas, K. N., & Blanchard, J. L. (2020). Reclassification of the *Clostridium clostridioforme* and *Clostridium sphenoides* clades as *Enterocloster* gen. nov. and *Lacrimispora* gen. nov., including reclassification of 15 taxa. *Int. J. System. Evolut. Microbiol.*, *70*(1), 23–34.

Hakulinen, J. K., Hering, J., Brändén, G., Chen, H., Snijder, A., Ek, M., & Johansson, P. (2017). Mray–antibiotic complex reveals details of tunicamycin mode of action. *Nat. Chem. Biol.*, *13*(3), 265–267.

Henrich, B., Lubitz, W., & Plapp, R. (1982). Lysis of escherichia coli by induction of cloned Φx174 genes. *Mol. Gen. Genet.*, *185*(3), 493–497.

Herelle, F. (1948). The bacteriophage. *Atomes*, *3*(33), 399–403.

Holtappels, D., Fortuna, K., Lavigne, R., & Wagemans, J. (2021). The future of phage biocontrol in integrated plant protection for sustainable crop production. *Curr. Opin. Biotech.*, *68*, 60–71.

Hutchison III, C. A., & Sinsheimer, R. L. (1963). Kinetics of bacteriophage release by single cells of Φx174-infected *E. coli*. *J. Mol. Biol.*, *7*(2), 206–208.

Hutchison III, C. A., & Sinsheimer, R. L. (1966). The process of infection with bacteriophage Φx174: X. mutations in a Φx lysis gene. *J. Mol. Biol.*, *18*(3), 429–IN2.

Ikura, T., & Ito, N. (2007). Requirements for peptidyl-prolyl isomerization activity: A comprehensive mutational analysis of the substrate-binding cavity of fk506-binding protein 12. *Protein Sci.*, *16*(12), 2618–2625.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žıdek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with alphafold. *Nature*, *596*(7873), 583–589.

Kannoly, S., Shao, Y., & Wang, I.-N. (2012). Rethinking the evolution of single-stranded rna (ssrna) bacteriophages based on genomic sequences and characterizations of two r-plasmid-dependent ssrna phages, c-1 and hgal1. *J. Bact.*, *194*(18), 5073–5079.

Kim, M.-S., Park, E.-J., Roh, S. W., & Bae, J.-W. (2011). Diversity and abundance of single-stranded dna viruses in human feces. *App. Environ. Microbiol.*, *77*(22), 8062–8070.

Kortright, K. E., Chan, B. K., Koff, J. L., & Turner, P. E. (2019). Phage therapy: A renewed approach to combat antibiotic-resistant bacteria. *Cell Host Microbe*, *25*(2), 219–232.

Lee, A. S., & Sinsheimer, R. L. (1974). A cleavage map of bacteriophage Φx174 genome. *Proc. Nat. Acad. Sci.*, *71*(7), 2882–2886.

Lehrman, M. A. (1994). A family of udp-glcnac/murnac: Polyisoprenol-p glcnac/murnac-1-p transferases. *Glycobiology*, *4*(6), 768–771.

Löw, C., Neumann, P., Tidow, H., Weininger, U., Haupt, C., Friedrich-Epler, B., Scholz, C., Stubbs, M. T., & Balbach, J. (2010). Crystal structure determination and functional characterization of the metallochaperone slyd from *Thermus thermophilus*. *J. Mol. Biol.*, *398*(3), 375–390.

Lubitz, W., Halfmann, G., & Plapp, R. (1984). Lysis of *Escherichia coli* after infection with Φx174 depends on the regulation of the cellular autolytic system. *Microbiology*, *130*(5), 1079–1087.

Luria, S. (1962). Genetics of bacteriophage. *Ann. Rev. Microbiol.*, *16*(1), 205–240.

Madeira, F., Pearce, M., Tivey, A. R., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from embl-ebi in 2022. *Nucleic Acids Res.*, *50*(W1), W276–W279.

Maratea, D., Young, K., & Young, R. (1985). Deletion and fusion analysis of the phage Φx174 lysis gene e. *Gene*, *40*(1), 39–46.

Martino, L., He, Y., Hands-Taylor, K. L., Valentine, E. R., Kelly, G., Giancola, C., & Conte, M. R. (2009). The interaction of the *Escherichia coli* protein slyd with nickel ions illuminates the mechanism of regulation of its peptidyl-prolyl isomerase activity. *FEBS J.*, *276*(16), 4529–4544.

Mashalidis, E. H., Kaeser, B., Terasawa, Y., Katsuyama, A., Kwon, D.-Y., Lee, K., Hong, J., Ichikawa, S., & Lee, S.-Y. (2019). Chemical logic of mray inhibition by antibacterial nucleoside natural products. *Nat. Comm.*, *10*(1), 1–12.

Matsuzaki, S., Uchiyama, J., Takemura-Uchiyama, I., & Daibata, M. (2014). Perspective: The age of the phage. *Nature*, *509*(7498), S9–S9.

Mezhyrova, J., Martin, J., Peetz, O., Dötsch, V., Morgner, N., Ma, Y., & Bernhard, F. (2021). Membrane insertion mechanism and molecular assembly of the bacteriophage lysis toxin Φx174-e. *FEBS J.*, *288*(10), 3300–3316.

Mokhonov, V. V., Vasilenko, E. A., Gorshkova, E. N., Astrakhantseva, I. V., Novikov, D. V., & Novikov, V. V. (2018). Slyd-deficient escherichia coli strains: A highway to contaminant-free protein extraction. *Biochem. Biophys. Res. Comm.*, *499*(4), 967–972.

Oluwole, A. O., Corey, R. A., Brown, C. M., Hernández-Rocamora, V. M., Stansfeld, P. J., Vollmer, W., Bolla, J. R., & Robinson, C. V. (2022). Peptidoglycan biosynthesis is driven by lipid transfer along enzyme-substrate affinity gradients. *Nat. Comm.*, *13*(1), 1–12.

Pavesi, A. (2021). Origin, evolution and stability of overlapping genes in viruses: A systematic review. *Genes*, *12*(6), 809.

Pazicky, S., Werle, A.-L. A., Lei, J., Löw, C., & Weininger, U. (2022). Impact of distant peptide substrate residues on enzymatic activity of slyd. *Cell. Mol. Life Sci.*, *79*(3), 1–18.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). Ucsf chimerax: Structure visualization for researchers, educators, and developers. *Protein Sci.*, *30*(1), 70–82.

Price, N. P., & Momany, F. A. (2005). Modeling bacterial udp-hexnac: Polyprenol-p hexnac-1-p transferases. *Glycobiology*, *15*(9), 29R–42R.

Quistgaard, E. M., Weininger, U., Ural-Blimke, Y., Modig, K., Nordlund, P., Akke, M., & Löw, C. (2016). Molecular insights into substrate recognition and catalytic mechanism of the chaperone and fkbp peptidyl-prolyl isomerase slyd. *BMC Biol.*, *14*(1), 1–25.

Rohou, A., & Grigorieff, N. (2015). Ctffind4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.*, *192*(2), 216–221.

Roof, W. D., Horne, S. M., Young, K. D., & Young, R. (1994). *slyD*, a host gene required for Φx174 lysis, is related to the fk506-binding protein family of peptidyl-prolyl cis-trans-isomerases. *J. Biol. Chem.*, *269*(4), 2902–2910.

Roof, W. D., & Young, R. (1995). ΦX174 lysis requires slyd, a host gene which is related to the fkbp family of peptidyl-prolyl cis-trans isomerases. *FEMS Microbiol. Rev.*, *17*(1-2), 213–218.

Rumnieks, J., & Tars, K. (2012). Diversity of pili-specific bacteriophages: Genome sequence of incm plasmid-dependent rna phage m. *BMC Micro.*, *12*(1), 1–8.

Saladi, S. M., Maggiolo, A. O., Radford, K., & Clemons, W. M. (2020). Structural biologists, let's mind our colors. *bioRxiv*.

Sanchez-Garcia, R., Gomez-Blanco, J., Cuervo, A., Carazo, J. M., Sorzano, C. O. S., & Vargas, J. (2021). Deepemhancer: A deep learning solution for cryo-em volume post-processing. *Comm. Biol.*, *4*(1), 1–8.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C., Slocombe, P. M., & Smith, M. (1977). Nucleotide sequence of bacteriophage Φx174 dna. *Nature*, *265*(5596), 687–695.

Scholz, C., Eckert, B., Hagn, F., Schaarschmidt, P., Balbach, J., & Schmid, F. X. (2006). Slyd proteins from different species exhibit high prolyl isomerase and chaperone activities. *Biochemistry*, *45*(1), 20–33.

Schooley, R. T., Biswas, B., Gill, J. J., Hernandez-Morales, A., Lancaster, J., Lessor, L., Barr, J. J., Reed, S. L., Rohwer, F., Benler, S., et al. (2017). Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant *Acinetobacter baumannii* infection. *Antimicrob. Agents Chemo.*, *61*(10).

Sertic, V., & Boulgakov, N. (1935). Classification et identification des typhi-phages. *C. R. Seances Soc. Biol. Fil.*, *119*, 1270–1272.

Sinsheimer, R. L. (1959). A single-stranded deoxyribonucleic acid from bacteriophage Φx174. *J. Mol. Biol.*, *1*(1), 43–IN6.

Smith, H. O., Hutchison, C. A., Pfannkoch, C., & Venter, J. C. (2003). Generating a synthetic genome by whole genome assembly: ΦX174 bacteriophage from synthetic oligonucleotides. *Proc. Nat. Acad. Sci.*, *100*(26), 15440–15445.

Steindorf, D., & Schneider, D. (2017). *In vivo* selection of heterotypically interacting transmembrane helices: Complementary helix surfaces, rather than conserved interaction motifs, drive formation of transmembrane hetero-dimers. *Biochim. Biophys. Acta-Biomemb.*, *1859*(2), 245–256.

Steinegger, M., & Söding, J. (2018). Clustering huge protein sequence sets in linear time. *Nat. Comm.*, *9*(1), 1–8.

Sun, L., Young, L. N., Zhang, X., Boudko, S. P., Fokine, A., Zbornik, E., Roznowski, A. P., Molineux, I. J., Rossmann, M. G., & Fane, B. A. (2014). Icosahedral bacteriophage Φx174 forms a tail for dna transport during infection. *Nature*, *505*(7483), 432–435.

Tanaka, S., & Clemons Jr, W. M. (2012). Minimal requirements for inhibition of mray by lysis protein e from bacteriophage Φx174. *Molecular microbiology*, *85*(5), 975–985.

Teo, A. C., & Roper, D. I. (2015). Core steps of membrane-bound peptidoglycan biosynthesis: Recent advances, insight and opportunities. *Antibiotics*, *4*(4), 495–520.

Tessman, I. (1959). Some unusual properties of the nucleic acid in bacteriophages s13 and Φx174. *Virology*, *7*(3), 263–275.

Vollmer, W., Blanot, D., & De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.*, *32*(2), 149–167.

von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J.*, *5*(11), 3021–3027.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191.

Weininger, U., Haupt, C., Schweimer, K., Graubner, W., Kovermann, M., Brüser, T., Scholz, C., Schaarschmidt, P., Zoldak, G., Schmid, F. X., et al. (2009). Nmr solution structure of slyd from *Escherichia coli*: Spatial separation of prolyl isomerase and chaperone function. *J. Mol. Biol.*, *387*(2), 295–305.

Winn, M., Goss, R. J., Kimura, K.-i., & Bugg, T. D. (2010). Antimicrobial nucleoside antibiotics targeting cell wall assembly: Recent advances in structure–function studies and nucleoside biosynthesis. *Natural Prod. Rep.*, *27*(2), 279–304.

Witte, A., Schrot, G., Schön, P., & Lubitz, W. (1997). Proline 21, a residue within the α-helical domain of Φx174 lysis protein e, is required for its function in escherichia coli. *Molecular microbiology*, *26*(2), 337–346.

Witte, A., Wanner, G., Bläsi, U., Halfmann, G., Szostak, M., & Lubitz, W. (1990). Endogenous transmembrane tunnel formation mediated by Φx174 lysis protein e. *J. Bact.*, *172*(7), 4109–4114.

Young, K. D., & Young, R. (1982). Lytic action of cloned Φx174 gene e. *J. Virol.*, *44*(3), 993–1002.

Zhang, J. W., Butland, G., Greenblatt, J. F., Emili, A., & Zamble, D. B. (2005). A role for slyd in the *Escherichia coli* hydrogenase biosynthetic pathway. *J. Biol. Chem.*, *280*(6), 4360–4366.

Zheng, Y., Struck, D. K., Bernhardt, T. G., & Young, R. (2008). Genetic analysis of mray inhibition by the Φx174 protein e. *Genetics*, *180*(3), 1459–1466.

Zheng, Y., Struck, D. K., & Young, R. (2009). Purification and functional characterization of Φx174 lysis protein e. *Biochemistry*, *48*(22), 4999–5006.

Chapter 3

INTERACTIONS BETWEEN MRAY AND ITS LIPIDIC ENVIRONMENT

Adapted from:

Marmont, L. S., Orta, A. K., Baileeves, B. W., Sychantha, D., Fernández-Galliano, A., Li, Y. E., Greene, N. G., Corey, R. A., Stansfeld, P. J., Clemons Jr, W. M., et al. (2024). Synthesis of lipid-linked precursors of the bacterial cell wall is governed by a feedback control mechanism in *Pseudomonas aeruginosa*. *Nature Microbiology*, 1–13. https://doi.org/10.1038/s41564-024-01603-2.

Orta, A. K., Riera, N., Li, Y. E., Tanaka, S., Yun, H. G., Klaic, L., & Clemons Jr, W. M. (2023). The mechanism of the phage-encoded protein antibiotic from Φx174. *Science*, *381*(6654), eadg9091. https://doi.org/10.1126/science.adg9091.

ABSTRACT

In bacterial cells, the synthesis of surface glycans, such as the peptidoglycan layer, involves single units that are linked to a lipid carrier. The distribution of this carrier among competing biosynthetic pathways is a critical, yet not fully understood aspect of bacterial metabolism. In this study, we focus on the structural and dynamic aspects of MraY, which catalyzes the initial step in forming the lipid-linked peptidoglycan precursor. The cryoEM structure of *E. coli* MraY derived from its native environment brought the observation of several lipids, likely including the lipid carrier substrate. Additionally, we identify several hyperactive MraY mutants that cause the overproduction of lipid II. These variants are located within a previously uncharacterized cavity at the periplasmic side of the MraY dimer interface, distal to the active site. Through detailed structural analysis and molecular dynamics simulations, we propose that this cavity acts as a novel binding site for periplasmic lipid II. The interaction at this site likely leads to allosteric inhibition of MraY, serving as a feedback mechanism that regulates lipid carrier availability in peptidoglycan biosynthesis. This structural insight not only elucidates a new regulatory mechanism but also opens pathways for targeting bacterial cell wall synthesis in drug-resistant strains.

3.1 Introduction

The integral membrane enzyme MraY plays an essential role in the synthesis of the bacterial cell wall. MraY catalyzes the attachment of Park's Nucleotide to the lipid carrier undecaprenyl phosphate (C55P), yielding the peptidoglycan precursor lipid I. C55P is central to the biosynthesis of cell wall polymers in both Gram-positive and Gram-negative bacteria (reviewed in (Bouhss et al., 2008)). Within this pathway, MraY catalyzes the transfer of the MurNAc-pentapeptide to C55P, which undergoes an attachment of a GlcNAc molecule and is flipped onto the periplasm for building the peptidoglycan mesh. Following its role in precursor transport, the sugar pentapeptide moiety of lipid II is removed, yielding undecaprenyl pyrophosphate (C55PP). The C55PP lipid is then dephosphorylated and flipped onto the cytoplasmic side of the bilayer for recycling (Fig. 3.1). Beyond peptidoglycan biosynthesis, C55P also serves as a carrier molecule for synthesis of the outermembrane polysaccharide O-antigen, and lipoteichoic acid (Mancuso & Chiu, 1982; Samuel & Reeves, 2003). The amount of C55P in the membrane is rate limiting and essential for cell morphology and survival, meaning the distribution of this carrier must be tightly regulated between pathways (Jorgenson & Young, 2016; Lehrman, 1991).

Lipid Dependencies of MraY

Studies utilizing cell-free systems have demonstrated that *E. coli* MraY, along with other Gram-negative homologs, requires anionic lipids with phosphoglycerol head-groups to retain enzymatic activity (Henrich et al., 2016). This contrasts with the *B. subtilis* MraY homolog, which shows no such lipid selectivity (Henrich et al., 2016). Recent insights from mass spectrometry data suggested that the inactivity in non-anionic lipids is caused by impaired dimerization, opening the possibility that lipids play a role in the correct interactions between the MraY subunits (Henrich et al., 2017). However, the specific structural interactions that result in dimerization of MraY have yet to be defined. A mechanistic understanding of MraY protein-lipid interactions could pave the way for the development of novel antibiotic strategies targeting the lipid composition of pathogenic bacteria. Interestingly, other proteins that are involved in peptidoglycan biosynthesis and re-organization -notably FtsA and MinD- have been shown to directly interact with anionic phospholipids (Mileykovskaya & Dowhan, 2005).

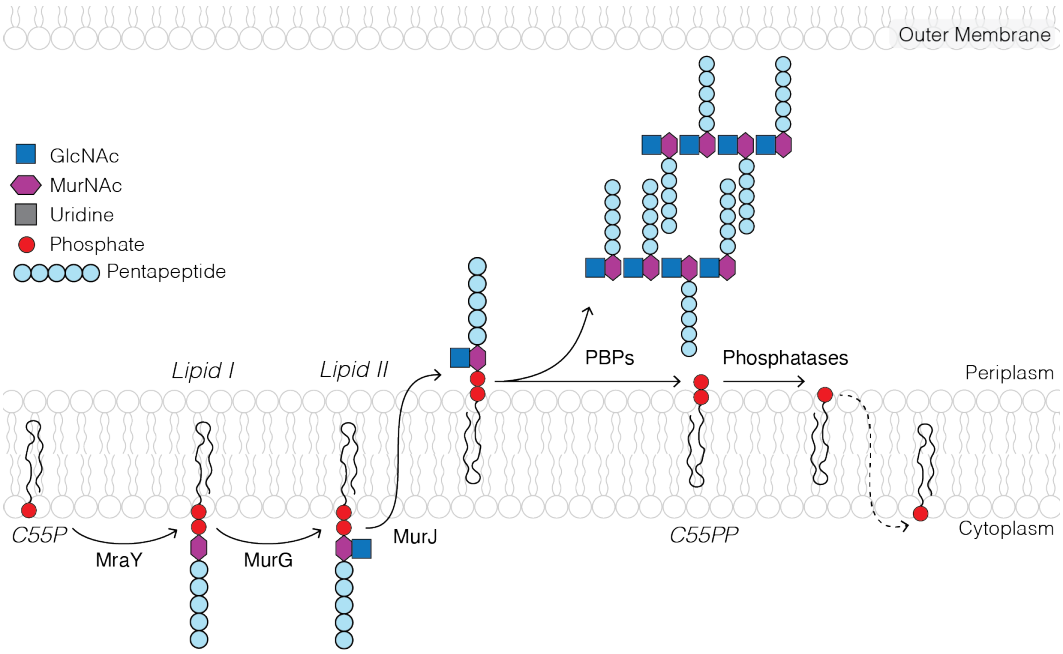


Figure 3.1: **C55P recycling model** Cytoplasmic C55P is first used as a carrier lipid by MraY to attach a phospho-MurNAc-pentapeptide. MurG follows by the attachment of a GlcNAc sugar and the product, lipid II, is flipped onto the periplasmic side of the bilayer by MurJ. PBPs then catalyze the cross-linking and cleavage of the carried molecule leaving C55PP. Phosphatases dephosphorylate the C55PP lipid and the lipid then gets flipped onto the cytoplasm for further cycles.

New Insights from Mass Spectrometry and Molecular Dynamics

Besides the importance of an anionic phospholipid environment, recent work has demonstrated MraY is often bound to the lipid substrate undecaprenyl phosphate (C55P) or to the product lipid I. Native mass spectrometry experiments demonstrated that endogenous C55P or lipid I binds to MraY dimers in a 4:2 stoichiometric ratio of C55P lipid to MraY molecules (Oluwole et al., 2022). Molecular dynamics simulations of MraY in a bilayer with enriched C55P, lipid I, and lipid II identified a putative lipid-binding region in MraY, which is thought to be located near the cytoplasmic dimer interface and in a region distinct from the active site (Oluwole et al., 2022). This study also showed that alanine substitutions at the identified charged residues cause a dramatic reduction in the dimeric MraY population. This suggests a structural basis for lipid-induced dimerization, highlighting an important aspect of MraY stability in its lipid environment.

The MraY, Protein E, SlyD (YES) complex density we obtained (Chapter 2) allowed us to assay the structure of MraY in a more native lipidic environment. Several obser-

vations regarding the lipids surrounding MraY were consistent with the aforementioned recent publications. Additionally, we sought to investigate MraY mutants that allowed the cell to survive in the presence of conditionally lethal defective cell wall synthases, class A penicillin-binding proteins, through the overproduction of lipid II. Altogether these studies support a feedback inhibition of MraY by accumulated lipid II through periplasmic interactions.

3.2 Results

Lipids bound to the YES complex

The YES complex was solubilized from its native environment and, for the YES_{ID21} complex, the final detergent solution was supplemented with *E. coli* lipids. We observe lipid densities around the membrane surface of MraY (Orta et al., 2023) (Fig. 3.2A & B). Previous work supports a functional role for anionic phospholipids with MraY including stabilizing the dimer (Henrich et al., 2016; Henrich et al., 2017; Oluwole et al., 2022). We observed a substantial density of lipids near the dimer interface with characteristics consistent with glycerophospholipids, although these densities could not be definitively assigned (Fig. 3.2A & B). In agreement with previous MraY structures (Chung et al., 2013; Hakulinen et al., 2017), we observed a hydrophobic periplasmic cavity within the MraY dimer that contains unexplained density (Fig. 3.2C). Although this density has clear structure, we are unable to fit typical *E. coli* phospholipids.

Recently, a study using native mass spectrometry and molecular dynamics identified a binding site for C55P at the dimer interface with the phosphate head-group predicted to bind to R341 in *Ec*MraY (R340 in *Aa*MraY) (Oluwole et al., 2022). In our structure, we observed a long density tube at the dimer interface (Fig. 3.2A, purple) that ends at R341 and is consistent with C55P (Fig. 3.2D). The role of this putative binding site remains to be elucidated, and merits further studies.

A structural comparison of MraY dimers

Published MraY structures have shown significant variability in the distance between the monomers. This is especially noticeable when the structures are aligned to a single MraY protomer. To compare the movement between dimers, we performed a structural alignment to the periplasmic half of the MraY transmembrane helices of a single protomer. This alignment was between the MraY of *Aquifex aeolicus* without bound ligands, the AlphaFold predicted structure, and the protein E-inhibited *Ec*MraY (Fig. 3.3). The distance between the same residue (Asn223 in *Ec*MraY) in each MraY monomer was calculated. The shortest distance was observed in the unbound *Aa*MraY structure, with a distance of approximately 11.4Å. The AlphaFold 2.0 structures varied between 14.7Å and 16.2Å at the same residue. The *Ec*MraY structure was similar to the AlphaFold 2.0 predicted dimers, with a distance of 15.7Å. Therefore, the MraY molecules in the unbound conformation (PDB:4J72) have a shorter distance between dimers at the periplasmic face.

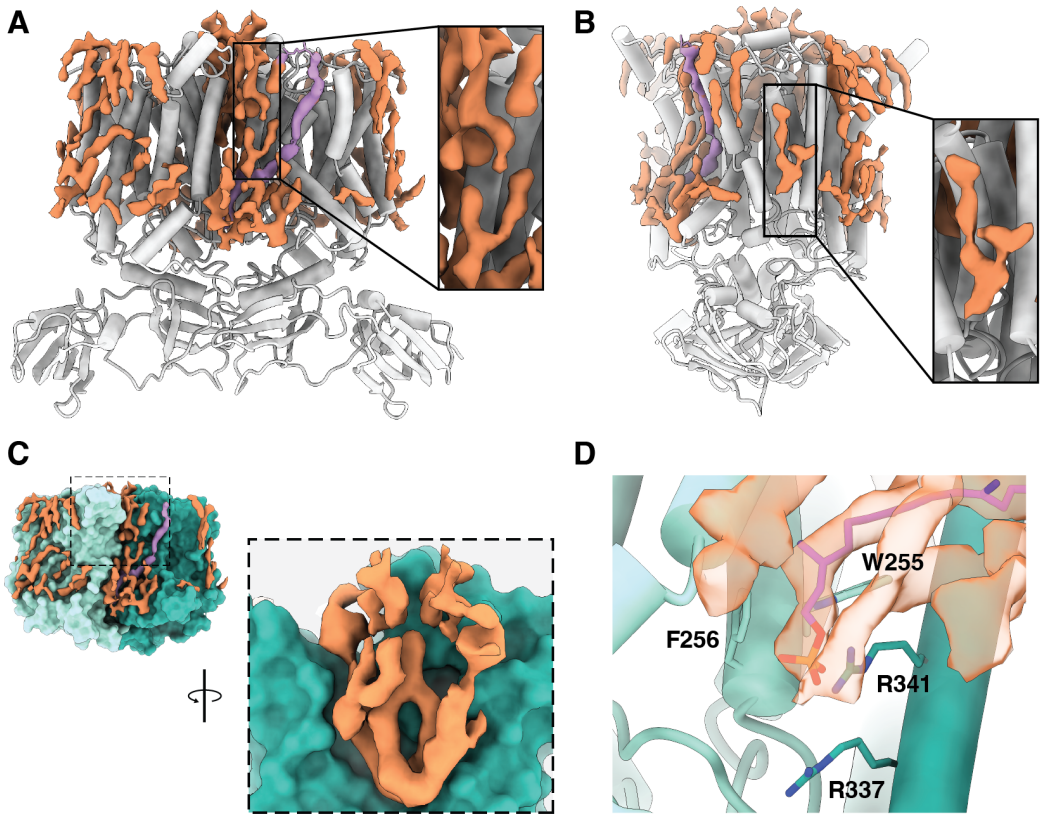


Figure 3.2: **Likely lipid densities bound in the YES complex**
A, Cartoon representation of the $YES_{(ID21)}$ complex (gray) with likely lipid densities in orange. Purple density is possible C55P. Inset shows possible phospholipid densities. **B,** As in (**A**) with a 90° rotation viewed towards the active site cleft. **C,** Accessible surface of *Ec*MraY (cyan) and unmodeled densities (orange/purple) that are likely a mix of lipids and detergent. The inset is a view of the periplasmic cavity viewed from a removed monomer. **D,** Putative pocket (Oluwole et al., 2022) for binding the phosphate of C55P. The putative binding residues and the modeled C55P (purple) are shown as sticks.

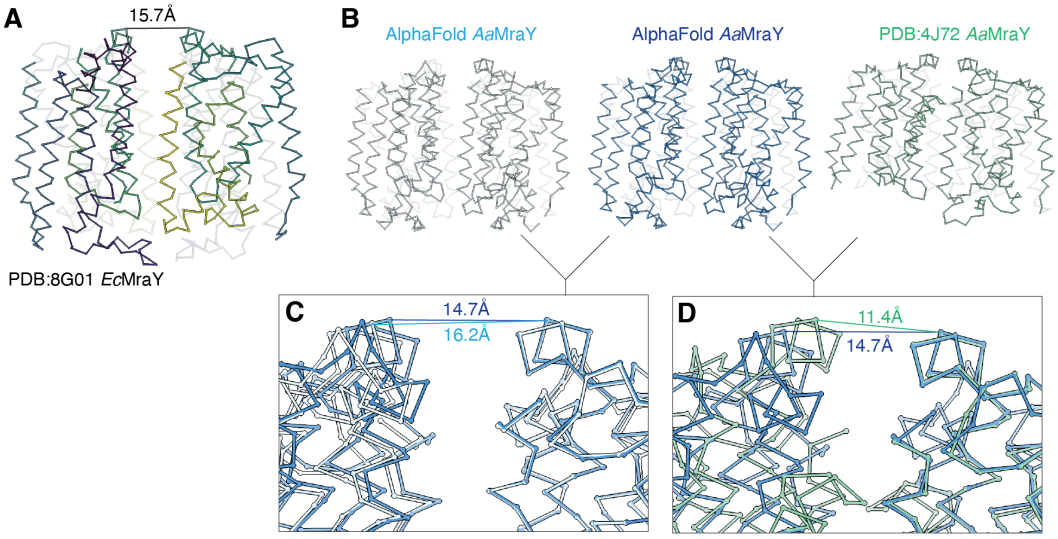
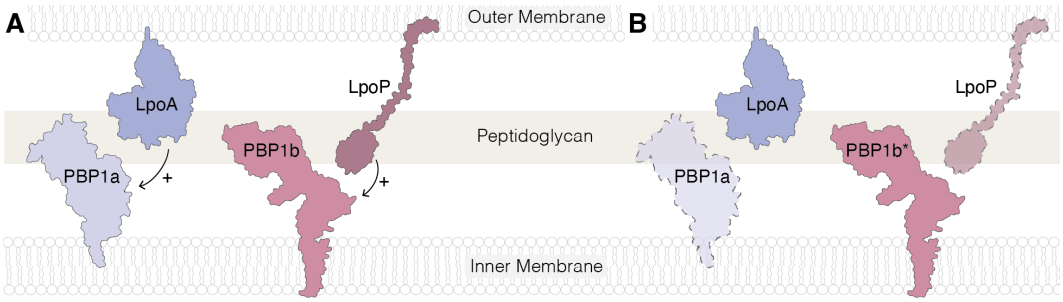


Figure 3.3: **Structural alignment of MraY dimers**
Structural alignment between *Ec*MraY, *Aa*MraY, and the predicted AlphaFold stuctures. **A,** Structure of the α-carbon atoms of *Ec*MraY within the YES complex (PDB:8G01) (sticks, Viridis). The distance between residue R341 to the opposing protomer at the same residue is 15.7 Å. **B,** Stick models of the α-carbon atoms of *Aa*MraY from AlphaFold 2.0 (light blue and navy blue) and the apo-crystal structure (PDB:4J72, green). The AlphaFold 2.0 models shown were selected by the longest (light blue) and shortest (navy blue) distance between protomers. **C,** Structural alignment between a single protomer of both AlphaFold 2.0 predictions shown in (**B**). The distance between residues R340 are 16.2 Åand 14.7 Årespectively. **D,** As in (**C**), single protomer structural alignment between the shortest AlphaFold 2.0 predicted distance and the crystal structure (PDB:4J72, green). The distance between residues R340 in the apo-structure is 11.4Å.

Periplasmic regulation of MraY

Penicillin-binding proteins (PBPs) play a crucial role in the synthesis of the bacterial cell wall and their activity is located in the periplasm. PBPs catalyze the final steps of peptidoglycan biosynthesis, cross-linking peptidoglycan strands, and maintaining the structural integrity of the cell wall (reviewed in Sauvage et al., 2008). *Pseudomonas aeruginosa* uses two class A penicillin-binding proteins (aPBPs), PBP1a and PBP1b, encoded by *ponA* and *ponB* respectively. These aPBPs are activated by outer membrane lipoproteins LpoA (PBP1a) and LpoP (PBP1b), encoded by *lpoA* and *lpoP* respectively (Fig.3.4A) (Greene et al., 2018; Paradis-Bleau et al., 2010). The Bernhardt group developed a *P. aeruginosa* double knockdown Δ*ponB* Δ*lpoA* strain that produces only one aPBP, PBP1a, without its activator LpoA (Fig.3.4B) (Greene et al., 2018). A spontaneous mutation in MraY (T23P) was found that



C

Res	*Pa*	*Ec*
14	Lys	Ser
21	Tyr	Tyr
22	Leu	Leu
23	Thr	Thr
31	Leu	Leu
160	Leu	Leu
164	Met	Phe
166	Lys	Lys
217	Tyr	Trp
220	Gly	Gly

Res	*Pa*	*Ec*
224	Phe	Phe
225	Ala	Ala
227	Tyr	Tyr
228	Leu	Leu
248	Gly	Gly
339	Ile	Ile
347	Val	Leu
354	Leu	Leu
358	Lys	Lys

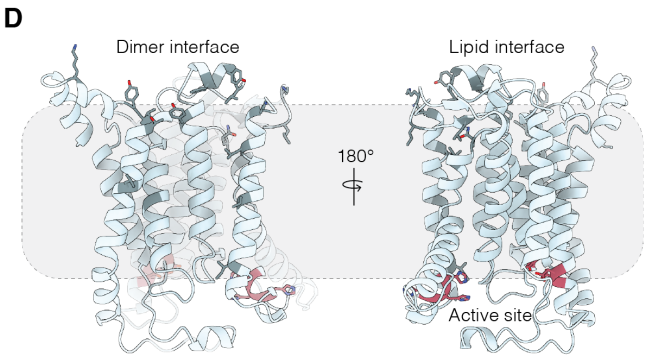


Figure 3.4: ***Pseudomonas aeruginosa* experimental scheme.**
A, Activation scheme of PBP1a (light purple) by LpoA (dark purple) and of PBP1b (pink) by LpoP (maroon). Inner membrane, outer membrane and peptidoglycan are shown for reference. **B,** As in (**A**), Dotted and opaque molecules represent the gene knockouts in the *P. aeruginosa* strain used in this study. An inactive PBP1b is denoted by an asterisk (*). **C,** Hyperactive MraY mutants identified in this study. The residue number corresponding to both species is shown on the left of the tables as 'Res'. The corresponding amino acid at this position is shown for *P. aeruginosa* (*Pa*) and *E. coli* (*Ec*). **D,** AlphaFold 2.0 predicted structure of *Pa*MraY shown from the dimer interface (left) or towards the active site (right). Residues identified in this study are highlighted in a darker hue. Essential residues in the active site are labeled in red.

rescued survival in minimal media. Suprisingly, the mutation T23P localizes to the cavity on the periplasmic side of the MraY dimer far from the active site.

To understand the functional implications of the T23P mutation in MraY of *Pseudomonas aeruginosa*, a series of biochemical assays were implemented. These assays were designed to quantitatively evaluate the mutation's impact on the enzyme's ability to synthesize lipid II. By comparing the enzymatic activity of the T23P mutant with that of wild-type MraY, it was discovered that this mutant T23P produced a higher lipid II output than in wild-type cells. Given the uncontrolled production of lipid II, it was hypothesized MraY would sequester any available C55P, which hinders its usage for synthesis of surface glycans such as O-antigen. This

was shown to be the case as *P. aeruginosa* cells with a T23P MraY mutation which had substantially less O-antigen than the wild type. Altogether these biochemical assays showed hyperactivity of MraY caused by a mutation in the periplasmic face of the protein. Further assays identified twenty-one additional mutants that rescued the cell in the *P. aeruginosa* Δ*ponB* Δ*lpoA* strain (Fig. 3.4C & 3.5). Interestingly, none of these mutants were found at the active site. Instead, they were all located at the MraY dimer interface and the periplasmic side of the protein (Fig. 3.4D).

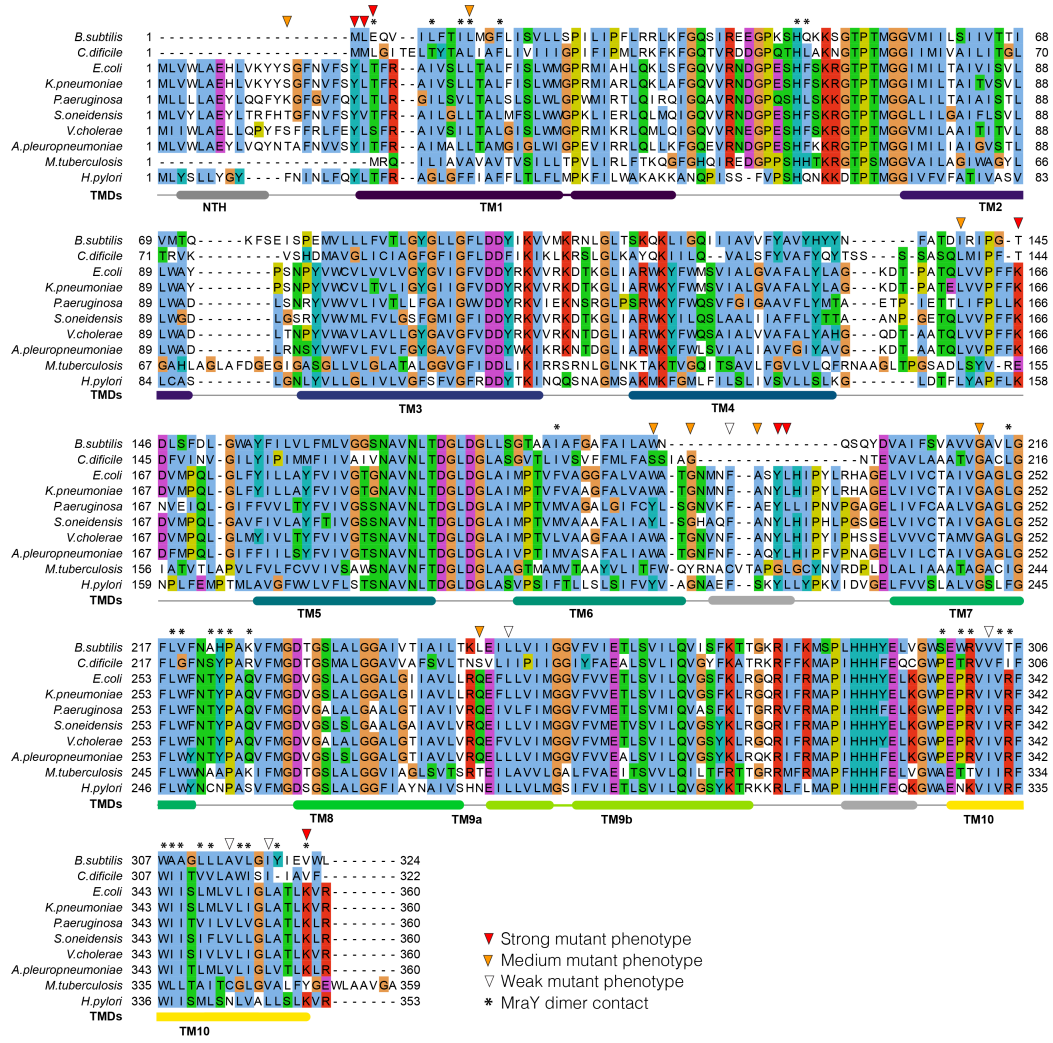


Figure 3.5: **Multiple sequence alignment of representative MraY species** Structure based multiple sequence alignment using Promals3D with ClustalW coloring for residues. Secondary structure based *Ec*MraY from the YES complex structure is shown below the sequences, TMDs are colored as in viridis. Upside-down triangles (▽) above the sequences denote mutants identified in this study colored by weak (white), medium (orange), and strong (red) phenotype. Asterisks (*) denote the dimer interface of MraY.

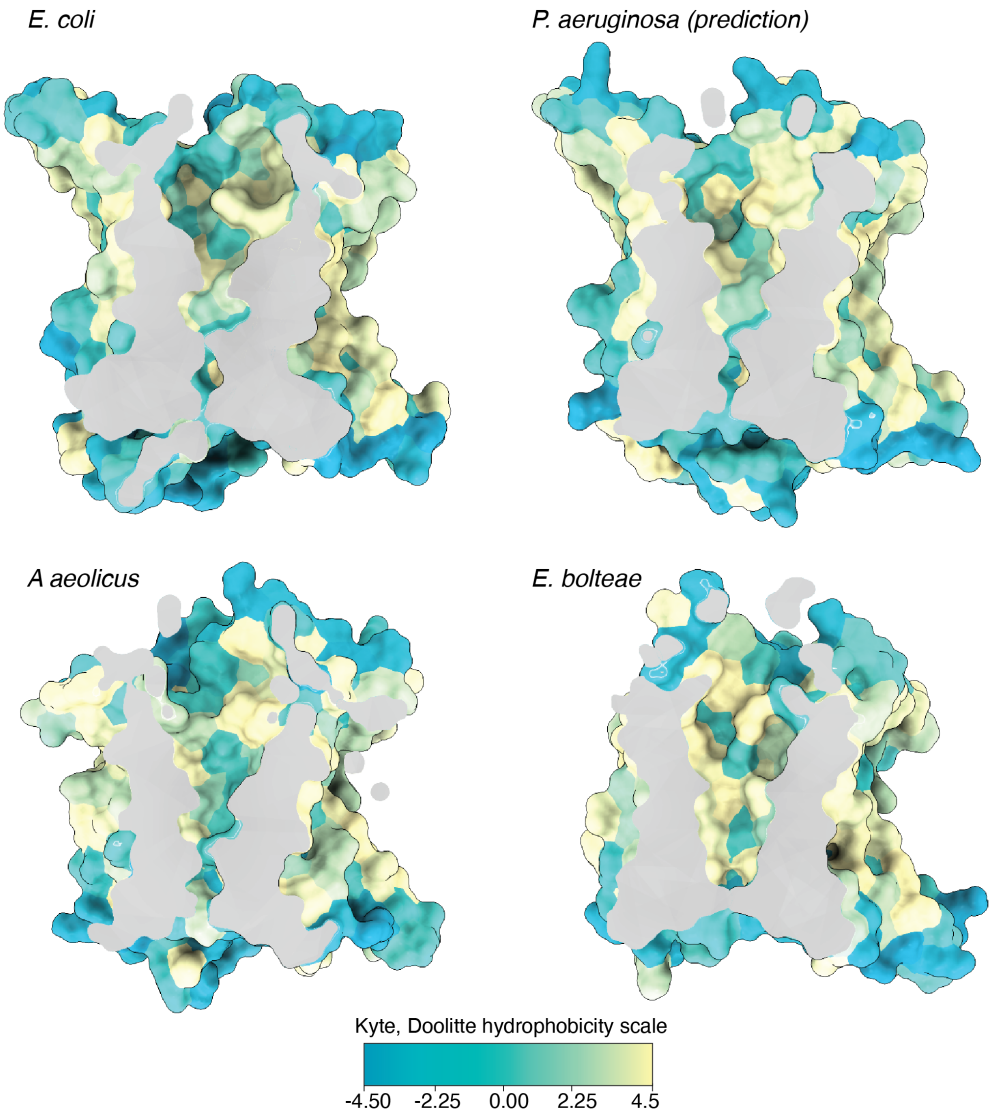


Figure 3.6: **Slab of MraY dimers colored by hydrophobicity**
Surface model slabbed at the dimer interface of the structures of MraY from *E. coli* (PDB:8G01), *A. aeolicus* (PDB:4J72), *E. boltae* (PDB:5JNQ), and the AlphaFold 2.0 prediction of *P. aeruginosa* MraY. Models are colored using the Kyte, Dootlittle hydrophobicity scale (Kyte & Doolittle, 1982).

Both the *A. aeolicus* and *E. boltae* MraY structures revealed the presence of a cavity located at the dimer interface that is lined by hydrophobic residues (Chung et al., 2013; Hakulinen et al., 2017). This hydrophobic cavity is a conserved characteristic of the enzyme (Fig. 3.6) and it was suggested that the unidentified electron density within it could accommodate one or more lipid molecules. Although it has been speculated to be C55P (Chung et al., 2013), the identity of the lipid has

remained unclear. A recent study identified lipid molecules that co-purify with MraY, including C55P, lipid I, and lipid II (Oluwole et al., 2022). Thus, MraY may bind a lipid molecule within the dimer interface near the residues that we have implicated in controlling the activity of the enzyme.

Clues to the potential identity of the lipid bound at the MraY dimer interface came from structural analysis of *Ec*MraY in complex with a phage-encoded inhibitor (protein E) and the *E. coli* chaperone SlyD (YES complex) (Orta et al., 2023). We leveraged the benefits of purifying *Ec*MraY in the presence of SlyD and protein E to obtain the structure of *Ec*MraY(T23P) within the same complex (Fig.3.9 and Table 3.1). In both cases, electron density was observed at the MraY dimer interface. Focused refinement of MraY alone in the *Ec*MraY(T23P) complex substantially improved the potential lipid density at the MraY dimer interface (Fig.3.9).

Similar to previous *A. aeolicus* and *E. boltae* MraY structures, this electron density fills the hydrophobic cavity found at the MraY dimer interface. However, we uniquely observed this electron density that extends into the periplasmic space above MraY molecules, where the environment is more hydrophilic (Fig.3.6). Although the density could not be conclusively modeled, the electron density is consistent with a large head group such as the disaccharide-pentapeptide found on lipid II. The structure of *Ec*MraY (T23P) reveals a hydrogen bonding network between residues Y21 in one protomer and Y227, and K358 in the second protomer (Fig.3.7). The proline appears to stabilize a down conformation of residue Y21, which is further enhanced by the tyrosine and lysine in the second molecule.

We wondered whether the gap between the MraY protomers was large enough to allow lipid entry into the cavity. The Stansfeld group applied molecular dynamics simulations for a deeper analysis of MraY in the lipid bilayer. The structure of *Ec*MraY from the YES complex (PDB:8G01) was used to perform coarse-grained molecular dynamic simulations. The simulations consisted of MraY in a bilayer in the presence of lipids C55P, C55PP, lipid I, or lipid II.

The coarse-grained simulations demonstrated that lipid I and lipid II spontaneously would enter the periplasmic cavity with the highest occupancy time, while C55P and C55PP occasionally enter but do not stably remain in the cavity. The phospholipids from the bilayer did not enter the cavity during the simulations. These experiments identified residues in which MraY putatively make contact with lipid II in nearly 100% of the experiments. Amongst these residues were Y21, T23, W217, F224, Y227, and K358, identified in the previously described mutagenesis screening.

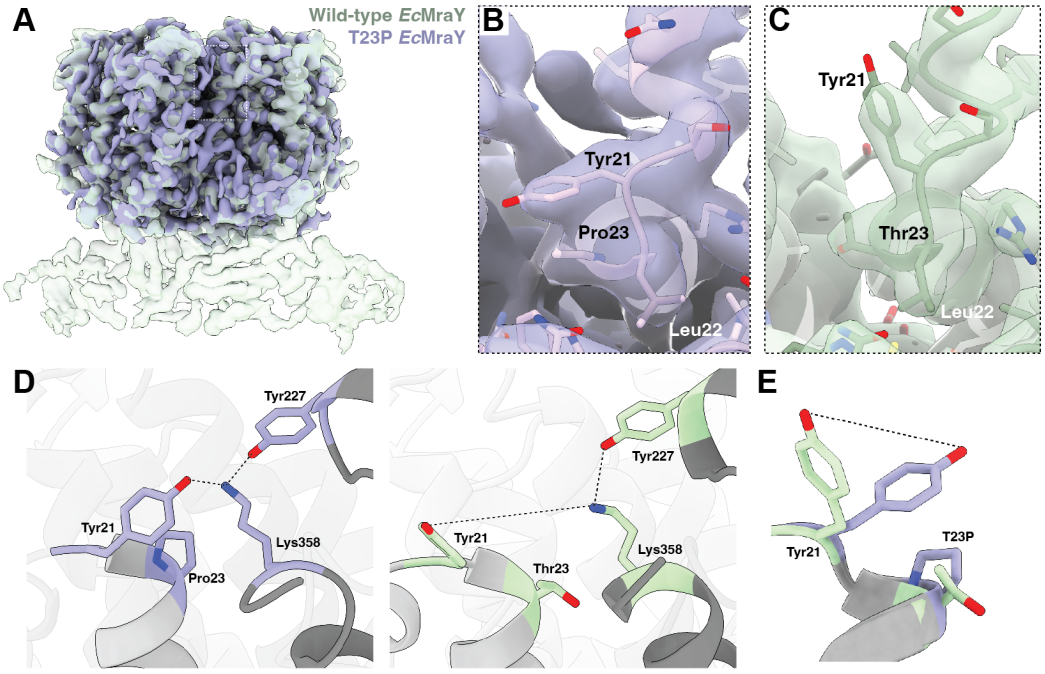


Figure 3.7: **Structural differences between the wild-type and the T23P MraY mutant**
A, Overlay of densities of MraY(WT) (EMD-29641)(green) and MraY(T23P) (purple) viewed in the plane of the membrane. **B,** Enlarged view of the densities around the T23P mutant. Residues are shown in stick representation. Residues 21-23 are labeled for reference. **C,** As in (**B**) for the wild-type complex. **D,** Hydrogen bonding network observed in MraY(T23P) (left, purple) compared to WT (right, green) at the mutagenesis site involving Y21, Y227, and K358. **E,** Similar to (**D**), overlay of the two models highlighting the conformational differences of residue Y21.

Residues Y21, L22, T23, W217, G220, F224, A225, Y227, L228, G238, L347, L354, K358 are neighboring the electron density in the periplasmic cavity of MraY.

The simulations showed that C55P-linked lipids, such as lipid I and lipid II, are capable of entering the periplasmic cavity and have long occupancy periods before exiting the cavity. Lipid I is not a periplasmic lipid, which suggests lipid II is the likely electron density in the periplasmic cavity. This preferential binding was evident in the higher binding constants observed for lipid II. Lipid II simulations showed the carbon tails curling into the cavity, with the sugar and amino acid moieties positioned towards the MraY periplasmic loops. In these simulations, two lipids would occupy the cavity simultaneously. We observed similarities between these putative lipid bound molecules and the electron density, including the curled carbon tail, density that could account for two lipids, and a density above the periplasmic

loops potentially accounting for the disaccharide-pentapeptide (Fig.3.8).

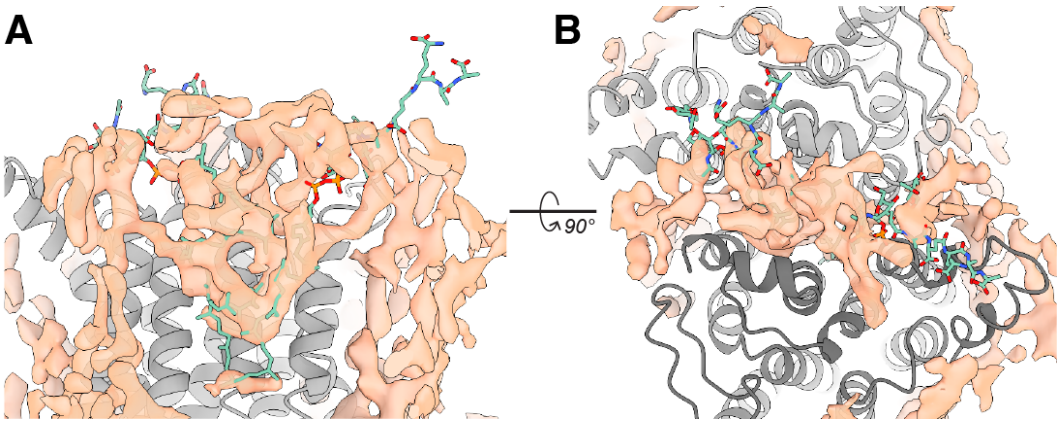


Figure 3.8: **Lipid fitting into densities from Molecular Dynamics experiments** **A,** Electron density after the subtraction of density accounted for by MraY (orange). View of the dimer interface of a single MraY molecule (gray, ribbon). One of the orientations of the two lipid II molecules simulated by molecular dynamics (green, stick) is fitted for reference into the density. **B,** Top view of model shown in (**A**). The second MraY molecule is shown in a darker shade of gray.

3.3 Discussion

In this study, we have explored the interactions between MraY and specific lipid molecules within the YES complex, particularly focusing on the structural and functional implications of these interactions. Our findings highlight the crucial role of lipid environments in stabilizing and regulating the MraY enzyme, a key component in bacterial cell wall biosynthesis.

The identification of lipid densities around the membrane surface of MraY, especially at the dimer interface, points to a direct role of lipids in modulating the MraY activity. Further insights were gained from the structural alignment between different MraY configurations, revealing significant variability in the distance between monomers across different species and conditions. These variations could potentially explain the differential lipid binding patterns observed, suggesting that the enzyme's conformational flexibility might be a mechanism for adapting to various lipid environments.

Outstanding questions remain regarding the role C55P takes in bacterial physiology, and its potential as a target for antibiotics. Further studies are required to investigate the mechanisms through which C55P gets distributed to different pathways. The putative C55P density observed in our data merits further study, it is likely that

lipid binding at this site is biochemically relevant. Further structural and mass spectrometry studies with crosslinking lipids to MraY can conclusively identify this lipid.

The presence of an occupied periplasmic hydrophobic cavity within the MraY dimer, consistently observed across different studies, is likely to be a critical site for lipid interaction. The biochemical analyses presented in this study detail MraY mutants capable of suppressing an aPBP deficiency in *E. coli* and *P. aeruginosa*. We infer this is through the saturation of lipid II bypassing the requirement for activators LpoA and LpoB (LpoP in *E. coli*). These hyperactive MraY mutants are in the periplasmic interface, distant from the active site in the cytoplasmic interface. As demonstrated by mass spectrometry, these mutants produce large amounts of lipid II. Additionally, the structural analysis of MraY provides new evidence of a periplasmic lipid II binding site, particularly through the identification of a long curled density at the dimer interface.

The application of molecular dynamics simulations has further elucidated the interaction dynamics within the MraY lipid bilayer, revealing that lipid II can spontaneously enter the periplasmic cavity and remain there for extended periods. This suggests affinity for lipid II over other phospholipids, C55P, and C55PP. In summary, our study provides compelling evidence that specific lipid interactions are not merely supportive of MraY stability but are integral to its functional regulation. We propose a feedback regulation mechanism where overproduction of lipid II leads to binding MraY at the periplasmic cavity, inactivating MraY. These interactions highlight potential targets for antibacterial strategies, especially through the design of molecules that can alter lipid interactions at critical sites on MraY. Further research into the detailed mechanisms of these interactions will be crucial in leveraging this knowledge towards the development of novel antibacterial therapies.

3.4 Methods

Expression and Purification of the YES Complex

The YES complex was expressed as described previously (Orta et al., 2023). Briefly, Δ*slyD* BL21(DE3) competent cells were transformed with pET22b-SlyD1–154 and pRSFDuet*Ec*MraY-E_{ID21}, and plated in LB agar containing 35 μg ml^{-1} kanamycin and 100 μg ml^{-1} ampicillin. The culture was grown in 2xYT media at 37°C and 225 r.p.m., and induced at an OD_{600} of 0.9 with 0.4 mM IPTG at 18°C overnight. The culture was collected by centrifugation at 9,000 × g for 10 min at 4°C, followed by flash freezing.

The cells were lysed using an M-110L microfluidizer (Microfluidics) in 20 mM Tris-HCl pH 7.5, 300 mM NaCl, 10% glycerol, 5 mM β-ME, 0.1 mM phenylmethyl sulfonyl fluoride, and 0.1 mM benzamidine. The lysate was cleared by a 20-min centrifugation at 22,000 × g. The membrane was isolated by ultracentrifugation at 167,424 × g and solubilized in 10 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 5 mM β-ME, 0.1 mM phenylmethyl sulfonyl fluoride, 0.1 mM benzamidine, 10 mM imidazole, and 1% DDM. The extract was cleared by ultracentrifugation and then nutated with 1 ml NiNTA resin (Qiagen) at 4°C for 2 hours. The resin was washed with 5 CVs of wash buffer (10 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, 5 mM β-ME and 0.03% DDM) with 10 mM imidazole and eluted in 20 ml of wash buffer containing 200 mM imidazole. The eluent was further purified by size exclusion chromatography (Superdex 200 5/150 GL, Millipore Sigma) in 10 mM HEPES pH 7.5, 75 mM NaCl, 5% glycerol, 5 mM β-ME and 0.03% DDM. Fractions were assessed by SDS–PAGE, concentrated, and directly used for cryo-EM sample preparation.

Sample Preparation for Cryo-EM

The protein sample was diluted to a concentration of 5 mg ml^{-1} in 10 mM HEPES pH 7.5, 75 mM NaCl, 2% glycerol, 5 mM β-ME, 0.03% DDM, and 1 mM *E. coli* total lipid extract (Avanti Polar Lipids, 100600P). Quantifoil holey carbon films R1.2/1.3 300 mesh copper grids (Quantifoil, Micro Tools) were glow discharged with a 2 min 20 Å plasma current using a Pelco easiGlow, Emitech K100X. Grids were prepared using a Vitrobot system (FEI Vitrobot Mark v4 x2, Mark v.3) by applying 3 μl of 5 mg ml^{-1} YES (T23P) complex onto the grid, followed by a 3.5s blot using a +8-blot force and plunge frozen into liquid ethane.

Data Acquisition and Analysis

Datasets were collected at ×105,000 magnification with a pixel size of 0.416 Å pixel^{-1} using a 300 kV cryo-TEM Krios microscope equipped with a Gatan K3 6k × 4k direct electron detector and a Gatan energy filter (slit width 20 eV) in super-resolution mode using Serial EM. Movies with 40 frames were recorded with a total exposure dose of 60 e^{-} Å^{-2} and a defocus range of −1.0 to −2.5 μm. A total of 7,083 movies were gain reference and motion corrected using the patch motion correction built-in function in cryosparc (v.3.3.2) with a 2-fold bin that resulted in a pixel size of 0.832 Å pixel^{-1} (Punjani et al., 2017). The contrast transfer function (CTF) was estimated using CTFFIND4 (Rohou & Grigorieff, 2015). A total of 3,885,223 particles were obtained by template picker using PDB 8G01 (Orta et al., 2023) as reference. Four ab-initio models were obtained using 500,000 particles, from which the best and worst volumes were used to sort 4x binned particles through heterogeneous refinement.

Iterative rounds of heterogeneous and non-uniform refinement were performed before re-extracting particles using a 2x bin. This process was continued and the resulting particles were re-extracted using a 1.3x bin. After several rounds of heterogeneous and non-uniform refinement, 575,243 particles were extracted without binning and used to create a map through non-uniform refinement. Using the MraY model from PDB 8G01, a mask covering only the density encompassing MraY was created using ChimeraX (Pettersen et al., 2021). Density outside of this mask was removed using particle subtraction, followed by ab-initio modelling. The best fitting map was then used for further refinement using global CTF, heterogeneous, and non-uniform refinement. The final map with a 3.8 Å resolution was composed by 287,765 particles and sharpened using the autosharpen module in PHENIX-1.19.2. The data processing pipeline is summarized in Figure 3.9. The data collection, refinement, and validation statistics can be found in Table 3.1.

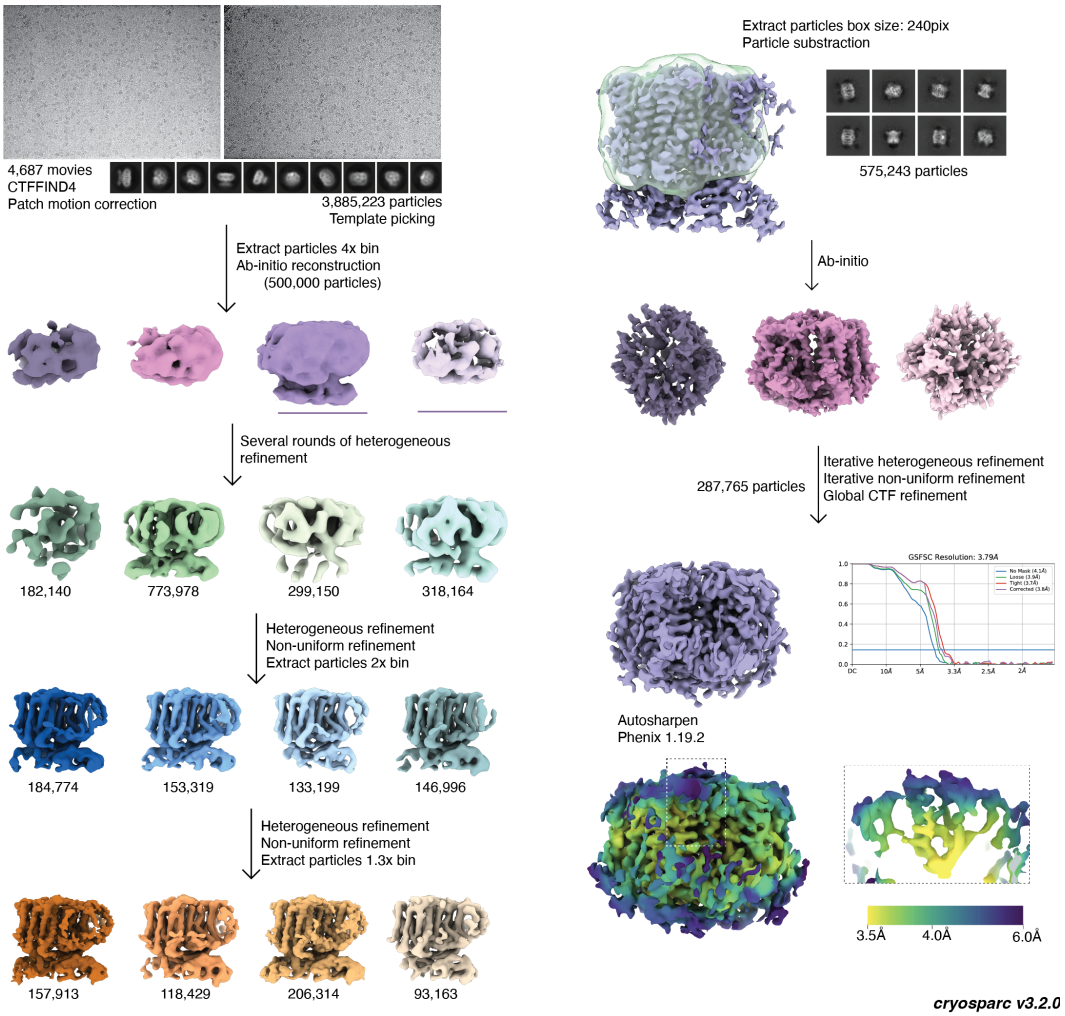


Figure 3.9: **CryoEM data processing of *Ec*MraY(T23P)**
Processing was done using cryosparc v3.2.0. The order of processing follows the arrows. The number of particles after each step is shown. Local resolution is shown in a Viridis color scheme ranging from 3.5Å(yellow) to 6.0Å(purple) resolution. Overall GSFSC resolution was determined on cryosparc.

Table 3.1: ***Ec*MraY - Cryo-EM data collection, refinement, and validation statistics**

	***Ec*MraY** (EMDB:41373) (PDB: 8TLU)
Data collection and processing	
Microscope	FEI Titan Krios
Magnification	105,000
Voltage (kV)	300
Electron exposure (e^{-}/Å^2)	60
Defocus range (μm)	-1 to -2.5
Pixel size (Å)	0.832
Symmetry imposed	C1
Initial particle images (no.)	3,885,223
Final particle images (no.)	287,765
Map resolution (Å), FSC threshold	3.8, 0.143
Refinement	
Software	PHENIX 1.19.2
Initial model used (PDB code)	8G01
Resolution of unmasked reconstructions (Å, FSC=0.5)	4.1
Resolution of masked reconstructions (Å, FSC=0.5)	3.8
Correlation coefficient (CCmask)	0.76
Model composition	
Atoms (Hydrogens)	11180 (5703)
Protein residues	698
Ligands	0
B factors (Å^2)	
Protein (min/max/mean)	20.34/79.77/41.69
Ligand	-
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.536
Validation	
MolProbity score	1.43
Clashscore	7.87
Poor rotamers (%)	0
Ramachandran plot	
Favored (%)	98.40
Allowed (%)	1.60
Disallowed (%)	0

References

Bouhss, A., Trunkfield, A. E., Bugg, T. D., & Mengin-Lecreulx, D. (2008). The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiology Reviews*, *32*(2), 208–233. https://doi.org/10.1111/j.1574-6976.2007.00089.x

Chung, B. C., Zhao, J., Gillespie, R. A., Kwon, D.-Y., Guan, Z., Hong, J., Zhou, P., & Lee, S.-Y. (2013). Crystal structure of mray, an essential membrane enzyme for bacterial cell wall synthesis. *Science*, *341*(6149), 1012–1016.

Greene, N. G., Fumeaux, C., & Bernhardt, T. G. (2018). Conserved mechanism of cell-wall synthase regulation revealed by the identification of a new PBP activator in Pseudomonas aeruginosa [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *115*(12), 3150–3155. https://doi.org/10.1073/pnas.1717925115

Hakulinen, J. K., Hering, J., Brändén, G., Chen, H., Snijder, A., Ek, M., & Johansson, P. (2017). Mray–antibiotic complex reveals details of tunicamycin mode of action. *Nat. Chem. Biol.*, *13*(3), 265–267.

Henrich, E., Ma, Y., Engels, I., Münch, D., Otten, C., Schneider, T., Henrichfreise, B., Sahl, H.-G., Dötsch, V., & Bernhard, F. (2016). Lipid requirements for the enzymatic activity of mray translocases and in vitro reconstitution of the lipid ii synthesis pathway. *J. Biol.l Chem.*, *291*(5), 2535–2546.

Henrich, E., Peetz, O., Hein, C., Laguerre, A., Hoffmann, B., Hoffmann, J., Dötsch, V., Bernhard, F., & Morgner, N. (2017). Analyzing native membrane protein assembly in nanodiscs by combined non-covalent mass spectrometry and synthetic biology. *eLife*, *6*, e20954.

Jorgenson, M. A., & Young, K. D. (2016). Interrupting Biosynthesis of O Antigen or the Lipopolysaccharide Core Produces Morphological Defects in Escherichia coli by Sequestering Undecaprenyl Phosphate [Publisher: American Society for Microbiology]. *Journal of Bacteriology*, *198*(22), 3070–3079. https://doi.org/10.1128/jb.00550-16

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, *157*(1), 105–132. https://doi.org/10.1016/0022-2836(82)90515-0

Lehrman, M. A. (1991). Biosynthesis of N-acetylglucosamine-P-P-dolichol, the committed step of asparagine-linked oligosaccharide assembly. *Glycobiology*, *1*(6), 553–562. https://doi.org/10.1093/glycob/1.6.553

Mancuso, D. J., & Chiu, T. H. (1982). Biosynthesis of glucosyl monophosphoryl undecaprenol and its role in lipoteichoic acid biosynthesis [Publisher: American Society for Microbiology]. *Journal of Bacteriology*, *152*(2), 616–625. https://doi.org/10.1128/jb.152.2.616-625.1982

Mileykovskaya, E., & Dowhan, W. (2005). Role of membrane lipids in bacterial division-site selection. *Current Opinion in Microbiology*, *8*(2), 135–142. https://doi.org/10.1016/j.mib.2005.02.012

Oluwole, A. O., Corey, R. A., Brown, C. M., Hernández-Rocamora, V. M., Stansfeld, P. J., Vollmer, W., Bolla, J. R., & Robinson, C. V. (2022). Peptidoglycan biosynthesis is driven by lipid transfer along enzyme-substrate affinity gradients. *Nat. Comm.*, *13*(1), 1–12.

Orta, A. K., Riera, N., Li, Y. E., Tanaka, S., Yun, H. G., Klaic, L., & Clemons Jr, W. M. (2023). The mechanism of the phage-encoded protein antibiotic from Φx174. *Science*, *381*(6654), eadg9091. https://doi.org/10.1126/science.adg9091.

Paradis-Bleau, C., Markovski, M., Uehara, T., Lupoli, T. J., Walker, S., Kahne, D. E., & Bernhardt, T. G. (2010). Lipoprotein Cofactors Located in the Outer Membrane Activate Bacterial Cell Wall Polymerases. *Cell*, *143*(7), 1110–1120. https://doi.org/10.1016/j.cell.2010.11.037

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). Ucsf chimerax: Structure visualization for researchers, educators, and developers. *Protein Sci.*, *30*(1), 70–82.

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination [Publisher: Nature Publishing Group]. *Nature Methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

Rohou, A., & Grigorieff, N. (2015). Ctffind4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.*, *192*(2), 216–221.

Samuel, G., & Reeves, P. (2003). Biosynthesis of O-antigens: Genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydrate Research*, *338*(23), 2503–2519. https://doi.org/10.1016/j.carres.2003.07.009

Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A., & Charlier, P. (2008). The penicillin-binding proteins: Structure and role in peptidoglycan biosynthesis. *FEMS Microbiology Reviews*, *32*(2), 234–258. https://doi.org/10.1111/j.1574-6976.2008.00105.x

Chapter 4

PROTEIN E ENGINEERING

ABSTRACT

The study presented here explores the adaptability and efficacy of ΦX174 bacteriophage protein E as a broad-spectrum MraY inhibitor across different bacterial species, including Gram-positive. As a proof of concept, we select for protein E variants that inhibit the Gram-positive *Bacillus subtilis* MraY. Utilizing deep mutational scanning, we identify a double mutant of protein E that includes a residue previously characterized as essential for lysis. In this work, we demonstrate that truncations at the N-terminus of protein E and a C-terminal fusion to lysozyme rescues the ability of protein E to lyse *Bs*MraY-expressing *E. coli*. The study highlights the role of specific residues that are critical for the protein E mechanism of action and its potential expansion as a generalized antibacterial agent.

4.1 Introduction

The lytic activity of protein E makes it an attractive tool for the development of novel antibacterial agents. It has been observed that expression of recombinant protein E can cause host lysis as soon as 20 minutes after induction in actively growing bacteria (Bernhardt et al., 2000).

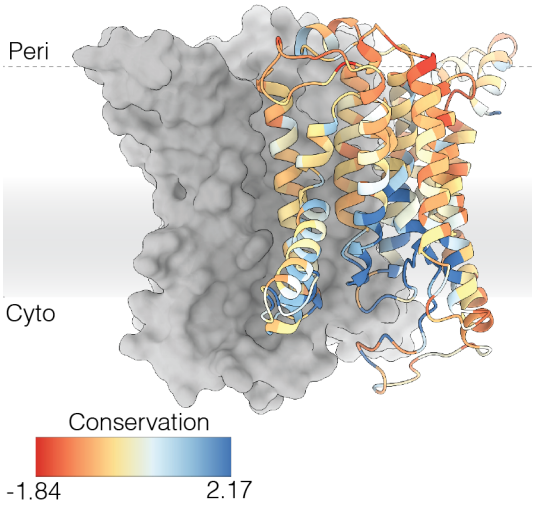


Figure 4.1: ***Ec*MraY conservation of the substrate binding regions**. MraY dimer (PDB:8G01) with one protomer shown as a surface model (gray) and the second in ribbon colored by conservation using a representative species alignment. Lower conservation is shown in red, higher conservation in blue. The periplasm (Peri) and cytoplasm (Cyto) are annotated for reference.

Protein E-mediated lysis has been demonstrated in all tested Gram-negative species to date (Cai et al., 2013; Marchart et al., 2003; Muhammad et al., 2019; Simon et al., 2011; Szostak et al., 1996). In contrast, Gram-positive species are thought to be resistant to lysis induced by protein E. For instance, heterologous expression of protein E in the Gram-positive *Staphylococcus carnosus* failed to produce host lysis (Halfmann et al., 1993). Moreover, it has been established that replacing the endogenous MraY in *E. coli* with its homolog from the Gram-positive *Bacillus subtilis* prevents protein E-mediated lysis (Zheng et al., 2008). This finding suggests that intrinsic differences in the MraY sequences between Gram-negative and Gram-positive bacteria may determine their susceptibility to inhibition by protein E.

The first goal in the design of a generalized protein E is to optimize a sequence that can bind to MraY homologs from Gram-positive bacteria. Although the soluble substrate binding site is conserved, significant sequence and structural variability is observed in the periplasmic-facing loops connecting the transmembrane helices, which includes the lipid binding site (Fig. 4.1). Intriguingly, most mutations in MraY that confer resistance to protein E are localized to this region and some are already present in Gram-positive homologs. For example, the phenylalanine at

position 288 in *E. coli* is replaced by leucine in both *E. bolteae* and *C. difficile* and by isoleucine in *B. subtilis*—all Gram-positive species. Additionally, proline at position 170 in *E. coli* is typically replaced by bulkier residues in Gram-positive species. These substitutions at F288 and P170 might partially account for the disrupted binding of protein E to MraY in these homologs.

Another requirement for protein E-mediated lysis is the presence of the chaperone SlyD, which is absent in Gram-positive bacteria. It was previously established that the proline residue at position 21 of protein E is essential for lysis (Tanaka & Clemons Jr, 2012; Witte et al., 1997). Initially, several hypotheses were proposed to explain this requirement. One such hypothesis suggested that the prolyl-isomerase chaperone SlyD facilitated the isomerization of P21, which was thought crucial for membrane insertion. However, this hypothesis was disproved when ΦX174 protein E mutants L19F and R3H, called Epos, were found to bypass the requirement for SlyD (Roof et al., 1994). These mutants occur naturally in the G4 phage within the Bullavirinae species and are silent mutations in *gene D*, which encodes a procapsid scaffolding protein. From the YES structure (Chapter 2), it was suggested that the L19F mutant allows for π-stacking with residue F182 in *Ec*MraY. It has been noted that phenylalanine residues contribute to the stabilization of membrane proteins (Steindorf & Schneider, 2017). The R3H mutant is believed to facilitate membrane insertion in accordance with the positive-inside rule (vonHeijne, 1989). These findings, coupled with the YES structure, support the notion that the role of SlyD is primarily for protein stability.

Further studies on protein E have refined our understanding of the minimal regions required for its inhibitory action. It has been demonstrated that truncating the first ten N-terminal residues does not substantially affect the lytic activity (Tanaka & Clemons Jr, 2012). Interestingly, truncation of the C-terminus appears to compromise protein E stability; however, this stability can be restored by fusion to soluble globular proteins such as maltose-binding protein (MBP) and lysozyme (Buckley & Hayashi, 1986; Maratea et al., 1985). For instance, a truncated version of protein E, comprising only the first 29 amino acids, exhibits inhibitory properties when fused to MBP (Tanaka & Clemons Jr, 2012). The shortest length tested that was viable for lysis included the N-terminal ten residue truncation and a lysozyme fused at position 29, which contained only residues 11-28 of protein E from ΦX174 (Tanaka & Clemons Jr, 2012).

Bacillus subtilis serves as a model organism for Gram-positive species, widely

used in molecular biology due to its non-pathogenic nature and well-documented genetic accessibility. As previously described, protein E appears unable to inhibit *Bs*MraY, likely due to low affinity (Zheng et al., 2008). In this study, we extend our exploration into engineering protein E through various strategies, including mutagenesis, deletions, and protein fusions. Our objective is to develop protein E into a versatile MraY inhibitor by specifically targeting the *Bacillus subtilis* MraY homolog.

4.2 Results

To pursue the mutagenesis of protein E towards a generalized MraY inhibitor, deep mutational scanning was applied to assay a broad range of mutants in a high-throughput manner. Deep mutational scanning has emerged as a transformative technique in the field of protein engineering, providing a comprehensive approach to understanding protein function and enhancing protein design. This high-throughput method involves creating a library of protein variants covering every possible amino acid substitution at regions of interest in the protein sequence. These variants are then expressed in a suitable system, in this case *E. coli*, and subjected to a selection process from which each mutant is enriched at different ratios.

Library Screening

The Epos mutants evolved in a Δ*slyD* background and are beneficial mutations for protein E while being silent in protein D (Roof et al., 1994). Despite being one of the few allowed mutations, L19F appears to greatly enhance the affinity of protein E for MraY, which is surprising as it is not found broadly in ΦX174 variants. We used deep mutational scanning to search for mutations that select for protein E-induced cell death, without the constraints of being embedded in protein D. We developed a BL21 *E. coli* strain where the wild-type *E. coli mraY* gene was replaced with *B. subtilis mraY*. This was accomplished using the INTEGRATE system, which allows for the insertion of transposable elements by guide RNA-assisted targeting (Vo et al., 2021). This method yielded a new *E. coli* strain that mostly relies on *Bs*MraY for survival. Of note, PCR experiments showed that 100% insertion efficiency was not attainable, and therefore *Ec*MraY is still present at low quantities in this cell population. This cell strain BL21-Gold Δ*mraY* +$mraY^{Bs}$, is hereby referred to as BL21-*BsmraY*.

We targeted the site-saturation library to assay residues around the essential proline at position 21 (Tanaka & Clemons Jr, 2012). This library limited the codon variability to only G and T bases at the wobble position, reducing the number of codons generated at each position from 64 to 32, while still including every amino acid mutant (NNK) (Acevedo-Rocha et al., 2015). Two libraries were generated, NNK2 and NNK4, with 2 (F19,P21) and 4 (F19, P21, S22, and L23) residues targeted for mutagenesis, respectively. Overall, the NNK2 library can generate 400 unique protein variants and the NNK4 library can generate 160,000 unique protein variants. In this study we developed a survival assay, through which protein E variants that are non-lytic remain in the cells, while lytic variants lyse the host and get released in

the media (Fig. 4.2A). By Next Generation Sequencing of the resulting supernatant, we observed that a double mutant (L19V P21S) of protein E was the most prevalent variant in a wild-type *E. coli* background. The proline residue at position 21 has been shown to be essential for lysis (Tanaka & Clemons Jr, 2012; Witte et al., 1997). It is intriguing that a mutation at position 21 is possible without the loss of function and that this is accompanied by a second mutation at position 19. Lysis assays corroborating the efficiency of this mutant were performed in *E. coli* BL21, with a cell lysis profile comparable to the wild-type protein E (Fig. 4.2C). This result merits further research.

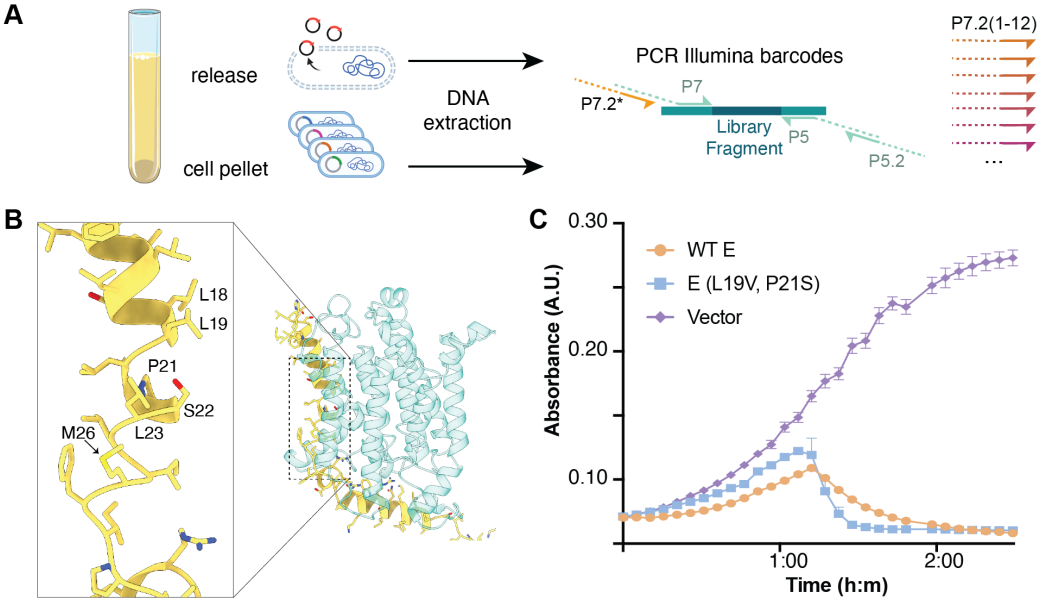


Figure 4.2: **Protein E mutant selection, isolation and characterization.** **A,** Schematic of the protocol followed for the isolation of lytic mutants. A cell culture with a NNK E library is transformed onto cells where the cells are pelleted and the release supernatant is isolated. The DNA is extracted and labeled for analysis. Primer P7.2 is used as a label with a different sequence per sample for analysis purposes. **B,** Structure of protein E from the YES complex. Inset shows the residues targeted in this mutagenesis screen. **C,** Cell growth as a function of time of *E. coli* upon induction of the protein E mutant L19V, P21S.

Truncation Tests

The majority of the MraY mutants that confer resistance to protein E were at the interface with the N-terminus of protein E (Fig. 4.3A). We hypothesized that an N-terminally truncated protein E might bind and inhibit *B. subtilis* MraY.

The substitution of *E. coli* MraY with the *B. subtilis* homolog conferred cells with the

ability to withstand the recombinant expression of wild-type protein E (Fig. 4.3B). We then used truncated forms of protein E in IPTG-inducible expression vectors to test for cell lysis in the BL21-*BsmraY* strain. Truncations of the first 9 (EΔN9) and 10 (EΔN10) residues (retaining the N-terminal methionine) were tested and the cell growth profiles upon IPTG induction are shown in Figure 4.3B. Interestingly, mutants EΔN9 and EΔN10 caused a defect in growth. This is distinct from the effect in wild-type BL21 *E. coli* where EΔN10 lysed with a similar profile to wild-type protein E.

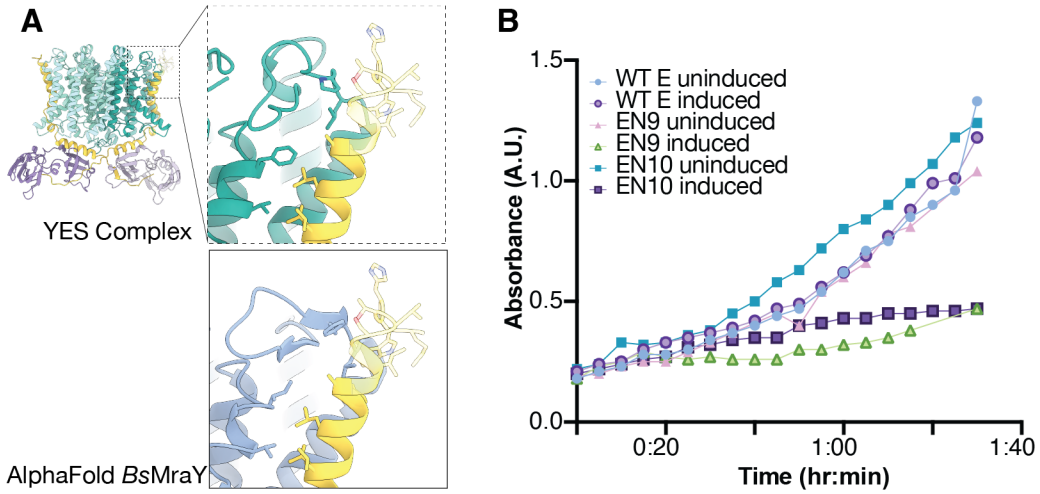


Figure 4.3: **N-terminal truncations of protein E**
A, Structure of the YES complex (PBD:8G01) with inset showing the protein E resistant mutant positions in *Ec*MraY (top). Predicted structure of *Bs*MraY structurally aligned to the YES complex (bottom) showing putative position of MraY relative to protein E. Residues shown in *Bs*MraY are the those corresponding to the protein E resistant mutants in *E. coli*. **B,** Lysis assay measuring cell density of BL21-Gold Δ*mraY* +*mraY*^{Bs} as a function of time after induction of WT protein E (circles), truncations of the first 9 residues of protein E (triangles), truncations of the first 10 residues of protein E (squares).

E-lysozyme

Protein E requires the host chaperone SlyD for stabilization. This dependency on SlyD for wild-type protein E may be another barrier for function in Gram-positive bacteria. *slyD* is only found in Gram-negative species and it is unlikely that protein E will be stabilized by a different chaperone posing another barrier for function in Gram-positive bacteria. To address this, we turned to a chimera of protein E with an inactive lysozyme appended after the transmembrane domain (Tanaka & Clemons Jr, 2012). We further hypothesized that the L19F mutant would also be

stabilizing for the interaction with *Bs*MraY.

We expressed protein E(L19F)-lysozyme in BL21-*BsmraY* and observed the growth profile upon induction. The expression of protein E(L19F)-lysozyme led to a lytic phenotype (Fig. 4.4A). Due to this observation, we wondered if a stable complex similar to YES could be isolated. For this, we generated a pRSFDuet vector containing protein E(L19F)-lysozyme-His6 in MCS1 and *Bs*MraY-StrepTag in MCS2. Cells were lysed, the membrane was pelleted, and protein was extracted in DDM. After NiNTA pull-down, we were unable to conclusively confirm the complex had been purified. Analysis of these results awaits further study.

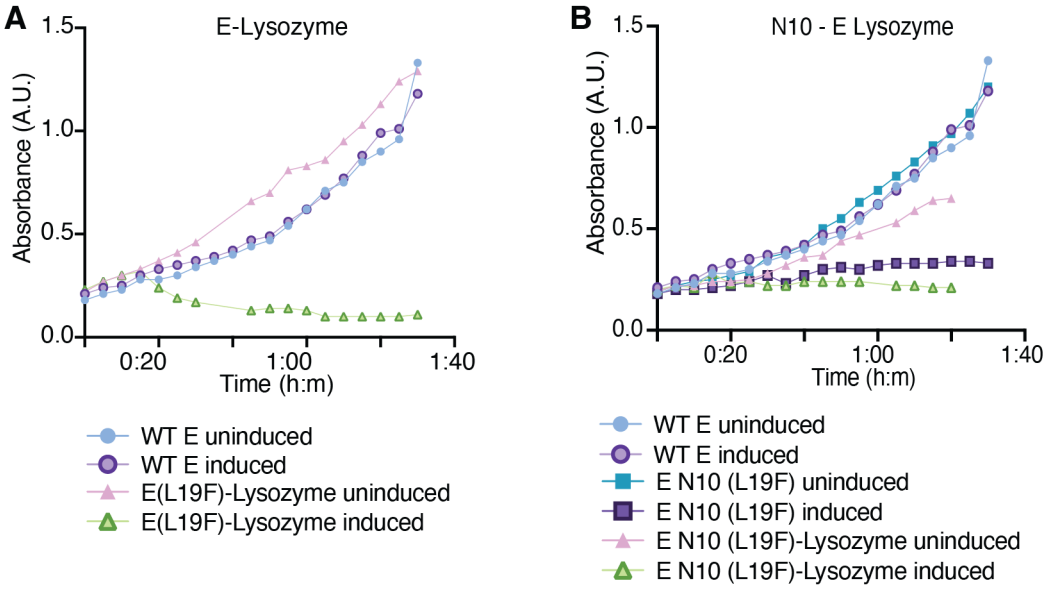


Figure 4.4: **Truncated E and fusion lysis assays**
A, Lysis assay measuring cell density of BL21-Gold Δ*mraY* +*mraY*^{Bs}as a function of time after induction of WT E (circles) and E(L19F)-Lysozyme (triangles). **B,** As in (**A**), cell density plot against time upon induction of truncations of the first 10 residues of E(L19F) (squares), E(L19F)-Lysozyme (triangles), and WT E (circles).

4.3 Discussion

The exploration of ΦX174 protein E as a broad-spectrum antibacterial agent, particularly its application beyond its native context, is an important step towards development of protein E as a tool. Our findings extend the understanding of protein E interactions with MraY, underscoring the critical role of specific amino acids in its mechanism of action. The mutagenesis studies focusing on the L19F and R3H mutations elucidate how subtle alterations can drastically enhance the capability of protein E to interact with and disrupt the function of MraY in Gram-negative bacteria. This increase likely reflects a stabilization of protein-protein interactions in the membrane.

Efforts to extend the application of protein E to Gram-positive species have faced challenges, likely due to the inherent sequence differences in MraY releative to Gram-negative species. The inability of protein E to effectively inhibit *Bs*MraY suggests a possible differences in the lipid substrate binding. These insights highlight the necessity for a targeted approach when designing broad-spectrum antibacterials. The data from the *Bs*MraY experiments, including the use of IPTG-inducible expression systems and mutagenesis libraries, reveal that while the adaptation of protein E to new targets is feasible, it requires precise modifications.

The integration of a lysozyme fusion strategy represents a simple approach to overcoming the stability issues associated with truncated forms of protein E. This strategy bypasses the requirement for SlyD . The preliminary results from the E(L19F)-lysozyme construct are promising, indicating that such chimeric proteins could maintain lytic activity beyond *E. coli* MraY.

In conclusion, while the extension of the protein E lytic activity to Gram-positive bacteria remains a complex challenge, these studies illustrate conditions in which protein E can interact with *Bs*MraY. Continued exploration of protein engineering, coupled with deep mutational scanning and structural analysis will allow overcoming barriers posed by diverse bacterial sequences.

4.4 Methods

Transposition of B. subtilis MraY into E. coli

The cargo block in pSPIN (Addgene, pSL1425) was substituted with $mraY^{Bs}$ (strain 168) . The pSPIN-BsMraY plasmid was then used to transform *E. coli* BL21 Gold. The cells were recovered for 1 hour at 37°C with shaking at 225rpm and plated onto 50ug/mL streptomycin selection plates. The colonies were then restreaked onto streptomycin plates supplemented with 0.1mM IPTG and allowed to grow at room temperature for 36 hours. Successful replacement of $mraY^{Ec}$ with $mraY^{Bs}$ was confirmed through colony PCR. Colonies were grown in 5mL of LB overnight at 37°C with shaking at 225rpm. Glycerol stocks were then prepared to a final concentration of 25% glycerol, flash frozen, and stored at -80°C for future use.

Plasmid library design and construction

An oligo-library was designed using error-prone nucleotides, oPools (IDT), of the nucleotide sequence corresponding to the transmembrane domain residues 2 through 29 of protein E from ΦX174. Primers were designed to contain the wild-type nucleotide sequence with a randomization of each codon at all three bases using "N", which assigns a canonical nucleotide at random. This was repeated with every codon accounting for residues 2 through 28. Additionally, every primer contained a 5' NcoI cut site and a 3' BsmbI cut site. Base pairs 87-273 were amplified with an additional 5' BsmbI cut site and a 3' SacI cut site. The primers were used to amplify the transmembrane domain of protein E using Q5 DNA polymerase (NEB). The amplified product was digested with 1 unit of restriction enzyme for 2 hours at 55°C. The fragments were then ligated using KLD (NEB) and incubating at 55°C for 1 hour. The product was then used to transform BL21-DE3 competent cells in triplicate. The cells were allowed to recover for 1 hour with shaking at 225rpm at 37°C in SOC media and plated in LB plates supplemented with 35 mg/L Kanamycin. 18,000 colonies were established as the minimal requirement for this transformation as it accounts for the number of variants required to encompass all mutations. The colonies were pooled together into 50 mL of LB media supplemented with 35mg/mL Kanamycin and incubated overnight with shaking at 225rpm at 37°C. The cells were then harvested and DNA extracted by midi-prep (Qiagen). Three libraries were generated from the same PCR product for statistical analysis. This method was repeated to generate the NNK libraries 2 (F19,P21) and 4 (F19, P21, S22 and L23) using Epos as template. The libraries were confirmed by the isolation of a single colony in LB+Kan plates, miniprep plasmid purification, and Primordium full plasmid sequencing.

Survival assay

BL21-Gold and BL21-Gold$\Delta mraY$ $+mraY^{Bs}$ were transformed with 200ng of the transmembrane library and grown directly in 40mL LB + Kanamycin to an OD600 of 0.5. The culture was then split into three 10mL culture tubes, one of which was harvested proceeded by DNA extraction of the pre-induction cells. One sample was then induced with 0.4mM IPTG while the other was grown under the same conditions as a control. After a 4-hour induction, the cells from both induced and uninduced samples were harvested at 6000rpm 4°C followed by mini-prep DNA extraction. This protocol was repeated with the replicate transmembrane libraries.

Twelve samples were obtained and the genes encoding for protein E within each sample were labeled with Illumina Index primers using NEBNext Multiplex Oligos for Illumina as follows:

Index	Cell strain	IPTG
1	BL21-Gold $\Delta mraY +mraY^{Bs}$	+
2	BL21-Gold $\Delta mraY +mraY^{Bs}$	-
3	BL21-Gold $\Delta mraY +mraY^{Bs}$	+
4	BL21-Gold $\Delta mraY +mraY^{Bs}$	-
5	BL21-Gold $\Delta mraY +mraY^{Bs}$	+
6	BL21-Gold $\Delta mraY +mraY^{Bs}$	-
7	BL21-Gold	+
8	BL21-Gold	-
9	BL21-Gold	+
10	BL21-Gold	-
11	BL21-Gold	+
12	BL21-Gold	-

The libraries were then sequenced using a 50 base-pair paired end flow cell with one million reads per sample using an Illumina HiSeq2500 high throughput sequencer at the Millard and Muriel Jacobs Genetics and Genomics Laboratory - Caltech. The results were multiplexed based on the index sequence. The adapter sequences were removed and filtered on a minimal length of 75 base pairs using Fastp through the Galaxy server (Chen et al., 2018; The Galaxy Community et al., 2022). The filtered sequences were then analyzed as translations and the amino acid count was then compared to the wild-type sequence and plots were generated using Enrich2 (Fowler & Fields, 2014).

Mutant screening

The highest count mutant in the wild-type *E. coli* assay was tested for lytic efficiency

by a lysis assay. The plasmid was isolated by plating and single colony selection followed by Primordium sequencing, which was then used to transform BL21 Gold competent cells. The cultures were grown in 100uL aliquots using a plate reader Infinite M Nano+ (Tecan, Switzerland) with shaking at 220rpm incubated at 37°C. Once the cell density reached early to mid log phase, the cultures were induced with 0.5mM IPTG and cell density was measured in 5 minute intervals.

References

Acevedo-Rocha, C. G., Reetz, M. T., & Nov, Y. (2015). Economical analysis of saturation mutagenesis experiments [Number: 1 Publisher: Nature Publishing Group]. *Scientific Reports*, *5*(1), 10654. https://doi.org/10.1038/srep10654

Bernhardt, T. G., Roof, W. D., & Young, R. (2000). Genetic evidence that the bacteriophage Φx174 lysis protein inhibits cell wall synthesis. *Proc. Nat. Acad. Sci.*, *97*(8), 4297–4302.

Buckley, K. J., & Hayashi, M. (1986). Lytic activity localized to membrane-spanning region of ϕx174 e protein. *Molecular and General Genetics MGG*, *204*(1), 120–125.

Cai, K., Zhang, Y., Yang, B., & Chen, S. (2013). *Yersinia enterocolitica* ghost with msbb mutation provides protection and reduces proinflammatory cytokines in mice. *Vaccine*, *31*(2), 334–340.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Fowler, D. M., & Fields, S. (2014). Deep mutational scanning: A new style of protein science. *Nature Methods*, *11*(8), 801–807. https://doi.org/10.1038/nmeth.3027

Halfmann, G., Götz, F., & Lubitz, W. (1993). Expression of bacteriophage Φx174 lysis gene e in *Staphylococcus carnosus* tm300. *FEMS Microbiol. Lett.*, *108*(2), 139–143.

Maratea, D., Young, K., & Young, R. (1985). Deletion and fusion analysis of the phage Φx174 lysis gene e. *Gene*, *40*(1), 39–46.

Marchart, J., Rehagen, M., Dropmann, G., Szostak, M., Alldinger, S., Lechleitner, S., Schlapp, T., Resch, S., & Lubitz, W. (2003). Protective immunity against pasteurellosis in cattle, induced by *Pasteurella haemolytica* ghosts. *Vaccine*, *21*(13-14), 1415–1422.

Muhammad, A., Kassmannhuber, J., Rauscher, M., Falcon, A. A., Wheeler, D. W., Zhang, A. A., Lubitz, P., & Lubitz, W. (2019). Subcutaneous immunization of dogs with *Bordetella bronchiseptica* bacterial ghost vaccine. *Front. Immunol.*, *10*, 1377.

Roof, W. D., Horne, S. M., Young, K. D., & Young, R. (1994). *slyD*, a host gene required for Φx174 lysis, is related to the fk506-binding protein family of peptidyl-prolyl cis-trans-isomerases. *J. Biol. Chem.*, *269*(4), 2902–2910.

Simon, R., Tennant, S. M., Wang, J. Y., Schmidlein, P. J., Lees, A., Ernst, R. K., Pasetti, M. F., Galen, J. E., & Levine, M. M. (2011). *Salmonella enterica* serovar enteritidis core o polysaccharide conjugated to h: G, m flagellin as a

candidate vaccine for protection against invasive infection with *S. enteritidis*. *Infec. Immun.*, *79*(10), 4240–4249.

Steindorf, D., & Schneider, D. (2017). *In vivo* selection of heterotypically interacting transmembrane helices: Complementary helix surfaces, rather than conserved interaction motifs, drive formation of transmembrane hetero-dimers. *Biochim. Biophys. Acta-Biomemb.*, *1859*(2), 245–256.

Szostak, M. P., Hensel, A., Eko, F. O., Klein, R., Auer, T., Mader, H., Haslberger, A., Bunka, S., Wanner, G., & Lubitz, W. (1996). Bacterial ghosts: Non-living candidate vaccines. *J. Biotech.*, *44*(1-3), 161–170.

Tanaka, S., & Clemons Jr, W. M. (2012). Minimal requirements for inhibition of mray by lysis protein e from bacteriophage Φx174. *Molecular microbiology*, *85*(5), 975–985.

The Galaxy Community, Afgan, E., Nekrutenko, A., Grüning, B. A., Blankenberg, D., Goecks, J., Schatz, M. C., Ostrovsky, A. E., Mahmoud, A., Lonie, A. J., Syme, A., Fouilloux, A., Bretaudeau, A., Nekrutenko, A., Kumar, A., Eschenlauer, A. C., DeSanto, A. D., Guerler, A., Serrano-Solano, B., ... Briggs, P. J. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Research*, *50*(W1), W345–W351. https://doi.org/10.1093/nar/gkac247

Vo, P. L. H., Ronda, C., Klompe, S. E., Chen, E. E., Acree, C., Wang, H. H., & Sternberg, S. H. (2021). CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering [Publisher: Nature Publishing Group]. *Nature Biotechnology*, *39*(4), 480–489. https://doi.org/10.1038/s41587-020-00745-y

vonHeijne, G. (1989). Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues [Publisher: Nature Publishing Group]. *Nature*, *341*(6241), 456–458. https://doi.org/10.1038/341456a0

Witte, A., Schrot, G., Schön, P., & Lubitz, W. (1997). Proline 21, a residue within the α-helical domain of Φx174 lysis protein e, is required for its function in escherichia coli. *Molecular microbiology*, *26*(2), 337–346.

Zheng, Y., Struck, D. K., Bernhardt, T. G., & Young, R. (2008). Genetic analysis of mray inhibition by the Φx174 protein e. *Genetics*, *180*(3), 1459–1466.

Chapter 5

CONCLUDING REMARKS

The recently elucidated structures of *E. coli* MraY in complex with viral protein E and the chaperone SlyD offer insights into the inhibition of peptidoglycan biosynthesis, a critical aspect of bacterial cell wall synthesis. These structures, protein E from ΦX174 and *Bullavirinae* relative ID21, provide a detailed view of the YES complex, highlighting the simplicity of the protein E mechanism of inhibition. Protein E binding within the active site cleft of MraY both obstructs the lipid substrate access and precludes active-site residues from conformational changes.

The role of SlyD is crucial but indirect; it stabilizes the inhibited complex by binding to the cytoplasmic domain of protein E, enhancing the overall structural stability of the complex, without being involved directly in the inhibition mechanism itself. The intial binding of SlyD to protein E likely is initiated by the presence of proline residues in the unstructured C-terminal domain. SlyD not only binds the hydrophobic face of the cytoplasmic helix of protein E, but also orders the C-terminal tail of the opposite protein E molecule likely protecting it from degradation. Thus, protein E brings together two *E. coli* proteins, MraY and SlyD, that would otherwise not interact. This is likely a coincidence of evolution and the constraints of small phage genomes against evolution.

The structure enhances our understanding of MraY inhibition by the phage derived protein E and also serves as a template for exploring novel antibacterial strategies. The structural studies have uncovered previously unresolved features of *Ec*MraY, providing insights into essential loops and lipid densities that underscore the importance of the C55P binding site at the dimer interface. These findings are instrumental in advancing our knowledge of bacterial cell wall synthesis.

The feedback regulation of MraY provides insights into the distribution of the shared C55P lipid carrier among various glycan biogenesis pathways. The enzyme's ability to catalyze the formation of lipid-linked peptidoglycan precursors marks a pivotal step in bacterial cell wall synthesis. In this work we show substantial evidence that the regulatory mechanism of MraY is mediated by feedback inhibition involving periplasmic lipid II, a derivative of the product lipid I, which accumulates under specific cellular conditions. This feedback inhibition ensures that MraY activity is adjusted in response to the availability and demand for lipid II, thereby balancing its production. The observation that the MraY(T23P) mutation can increase lipid II levels and compensate for deficiencies in peptidoglycan biosynthesis underlines the enzyme's role in managing cellular resource allocation.

Further supporting this model, biochemical assays with purified MraY variants

illustrate that the enzyme's regulatory control is intrinsic and not dependent on additional proteins. This finding is critical because it shows that MraY can autonomously adjust its activity based on the lipid II levels within the cell, which acts as a feedback signal. Molecular dynamics simulations and structural analyses suggest that the regulatory effect is exerted through specific binding interactions at the MraY dimer interface. These interactions likely affect the enzyme's activity by altering the structural dynamics necessary for its catalytic function. This regulatory mechanism underscores the importance of MraY in peptidoglycan biosynthesis but also positions it as a potential target for novel antibacterial strategies. Historically, targeting the active site of MraY has proven to be difficult as potential molecules must cross two membranes in diderms. Having a periplasmic regulatory site allows for new approaches towards targeting MraY to disrupt bacterial cell wall synthesis selectively in a more accessible manner.

Outstanding questions behind the regulatory mechanisms of MraY and the inhibition by protein E remain. Future work should aim at further detailing the lipidic environment and requirements for MraY. The putative C55P binding site near the dimer interface presented by Oluwole et al. is consistent with our electron density should to be further studied. It is unknown whether binding at this site in our maps is physiologically relevant, or an artifact of purification. Additionally, we underscore the importance of the periplasmic lipid binding site for regulation of Gram-negative MraY. This regulation was consistently observed in both *P. aeruginosa* and *E. coli*. An interesting remaing question is whether the Gram-positive MraY homologs also are regulated this way.

Our investigations into the interactions between protein E and the enzyme MraY have identified critical amino acids that play key roles in the mechanism of inhibition. However, expanding the scope of the protein E antimicrobial activity to include Gram-positive bacteria presents significant challenges, primarily due to sequence variations in MraY across different bacterial classes. The ineffectiveness of wild-type protein E against *Bs*MraY highlights these challenges, pointing to differences in the the interfaces that may influence lipid-substrate binding. This underscores the need for a highly targeted approach in the development of broad-spectrum antibacterials. Moreover, the adoption of lysozyme fusion strategies to improve the stability and efficacy of protein E variants, like the E(L19F)-Lysozyme constructs, shows promise in inhibition of Gram-positive MraY. These approaches in protein engineering and the encouraging preliminary results mark significant strides towards

expanding the utility of phage-derived lytic proteins, potentially paving the way for novel treatments against a diverse array of bacterial pathogens.