Structural and Biochemical Studies of Enzymes in Bacterial Glycobiology

Thesis by

Hyun Gi Yun

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2019

(Defended May 22, 2019)

Hyun Gi Yun
ORCID: 0000–0002–3508–5791

ACKNOWLEDGEMENTS

I would like to first thank my advisor, Professor Bil Clemons, for giving me this opportunity of working in his lab and for being patient with me as I slowly learned and grew to become an independent researcher. Over the past years, his encouragement and advice helped me get through challenging times. I really appreciate that he has always been supportive on my decisions and choices.

I also thank my committee members, Professor Linda Hsieh-Wilson, Professor Shu-ou Shan, and Professor Dianne Newman, for adding different perspectives to my research projects during annual meetings. Their advice and comments let me think about some specific topics more carefully and deeply. Also, all the interactions I had with these successful woman scientists motivated me to work harder.

I was lucky to collaborate with Professor Michio Kurosu at the University of Tennessee Health Science Center and Professor Jeremiah Johnson at the University of Tennessee, Knoxville. Their expertise in synthetic chemistry and microbiology, respectively, greatly expanded my understanding of the bacterial proteins I studied.

I have great appreciation toward former members of the Clemons Lab. Dr. Kyoung-Soon Jang was my mentor when I joined the lab. He not only taught me all the basics to perform experiments, but also motivated me with interesting research ideas, including the Gne project. For the MraY project I next worked on, resources I collected from Dr. Shiho Tanaka, Dr. Lada Klaić, and Dr. Eugene Chun were a tremendous help for me in making progress. Dr. Austin Rice went on a trip to the Advanced Photon Source (APS) with me to shoot X-ray on crystals we harvested together. I really appreciate that he came with me, even though he was ready to start a new job at that time.

In addition, I was close friends with Dr. Ku-feng (Geoffrey) Lin and Dr. Jee-Young (Amanda) Mock. Especially in my first few years in the lab, they taught me techniques I was not familiar with and suggested ideas and tips when I was having issues with experiments. I also thank Dr. Harry Gristick, Dr. Axel Muller, and Dr. Stephen Marshall for giving me helpful comments and feedback on my projects.

I am grateful that I have got to work with current members of the Clemons Lab. I received great help from and had fruitful discussions with Dr. Nadia Riera, Cody Gillman, and Alex Barbato in Team Glycobio. It was also fun to work with Shyam Saladi, Michelle Fry, and Dr. Aye Myat Thinn.

I also enjoyed working with two BMB graduate rotation students, Patrick Almhjell and Anna (Karen) Orta, and volunteers, Raymond Liu, Ayesha Tariq, and Kendal Tapia. I would like to thank them for their enthusiasm and hard work.

I met amazing people outside the lab, who helped me to perform a variety of experiments throughout my Ph.D. program. I would like to thank Dr. Jens Kaiser and Pavle Nikolovski

in the Molecular Observatory for allowing me to use all the instruments and resources related to X-ray crystallography. I appreciate Dr. Andrey Malyutin and Dr. Alasadair McDowall for spending hours and days to train me on electron microscopes and providing resources at the Caltech Cryo-EM facility. I also thank Dr. Nathan Dalleska in the Environmental Analysis Center (EAC), Mona Shahgholi at Caltech CCE Multiuser Mass Spectrometry Laboratory, Dr. Annie Moradian in the Proteome Exploration Laboratory (PEL), Rees Lab, Chan Lab, and Shan Lab for letting me to use their instruments.

I am extremely lucky to meet and build close relationships with many friends throughout my Ph.D. years: Dr. Seo Jung Han, Dr. Sarah Ynnmi Lee, Dong-wook Kim, Yon-il Jung, Dr. Chris Roh, Dr. Ja Yeon Kim, Dr. Yu-Hsien Hwang Fu, Chengcheng Fan, Jae Ho Lee, Dr. Hyunju Cho, Dr. Un Seng Chio. Because of their presence and support, all these years were more fun than stressful.

Coincidently, my uncle, Professor Jinbum Choi, had a Sabbatical Year at the division of Geological and Planetary Sciences at Caltech in my second year. During that year, both my uncle and aunt, Namsuk Sung, cooked Korean food for me all the time and that reminded me of my parents back home in South Korea even more. It is hard to describe my feelings toward my parents these days. I think mostly sadness because I have been away from home for a long time and I cannot visit them as often as before. However, I am very grateful that they stay healthy, are positive about the future, and believe in me as always. Also, I am very proud of my older brother, who finished a Ph.D. program last year.

Finally, I would not have come this far without Kyu Hyun Lee.

ABSTRACT

The speed that bacterial pathogens gain resistance to antibiotics is alarming. Designing new antibacterial agents is urgent, but it requires understanding their bacterial targets at the molecular level to achieve high specificity and potency. In this thesis, I discuss the structural and biochemical investigations of three potential protein targets for antibiotics. The first is a UDP-Glc/GlcNAc 4-epimerase, called Gne, from the human pathogen *Campylobacter jejuni*. This enzyme is the sole source of *N*-acetylgalactosamine (GalNAc) in *C. jejuni*, which is a common component in three major glycoconjugates decorating the cell surface and is critical for pathogenesis. The second target protein is an integral membrane protein, called MraY, which catalyzes the transfer of phospho-*N*-acetylmuramyl (MurNAc) pentapeptide to a lipid carrier, undecaprenyl phosphate (C_{55}-P), producing Lipid I in the peptidoglycan biosynthesis pathway. In the following step, a peripheral protein called MurG catalyzes transferring *N*-acetylglucosamine (GlcNAc) to Lipid I and produces Lipid II, which provides the first building block of the peptidoglycan layer. Peptidoglycan is uniquely bacterial, with MraY and MurG both being essential for cell viability; therefore, they are attractive targets for the development of antibacterial agents and work toward their structures is presented. Finally, MraY from *Escherichia coli* is the target for the lysis protein E from phage ΦX174. Efforts toward elucidating the *Ec*MraY-E complex structure are demonstrated here. In total, this thesis provides important data toward a full mechanistic understanding of these important antibacterial targets.

PUBLISHED CONTENT AND CONTRIBUTIONS

Mitachi, K., **Yun, H.G.**, Kurosu, S.M., Eslamimehr, S., Lemieux, M.R., Klaić, L., Clemons, W.M., and Kurosu, M. (2018). Novel FR-900493 Analogues That Inhibit the Outgrowth of *Clostridium difficile* Spores. *ACS Omega* 3 (2), pp. 1726–1739. doi: 10.1021/acsomega.7b01740.

Contributions: H.G.Y provided purified proteins (*Hy*MraY and *Mj*AglH) for the experiments and participated in revising the manuscript.

Yun, H.G., Jang, K., Tanaka, S. and Clemons, W.M. The Structure of the UDP-Glc/GlcNAc 4-Epimerase from the Human Pathogen *Campylobacter jejuni*. **[In Preparation]**

Contributions: H.G.Y refined the X-ray diffraction data set, performed all the biochemical experiments, made figures, and wrote the manuscript.

Yun, H.G. and Clemons, W.M. Structure of MurG from *Hydrogenivirga sp*. in Complex with UDP-*N*-acetylglucosamine. **[In Preparation]**

Contributions: H.G.Y designed and performed all the experiments, made figures, and wrote the manuscript.

Klaić, L.*, **Yun, H.G.***, Saladi, S.*, Kurosu, M. and Clemons, W.M. Insight into the Catalytic Mechanism of the Bacterial Phosphotransferase MraY. **[In Preparation]**

Contributions: H.G.Y made figures and wrote the manuscript.

TABLE OF CONTENTS

Chapter 5

Insight into the catalytic mechanism of the bacterial phosphotransferase MraY

Chapter 6

Concluding Remarks

LIST OF FIGURES AND TABLES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 1

Introduction

Discovery of penicillin in 1928 by Dr. Alexander Fleming saved a large number of patients suffering from a variety of infectious diseases caused by bacterial pathogens and opened the world to the antibiotic age. Since then, a variety of antibiotics were discovered and used in clinics, but the pathogens that were targeted quickly developed resistance (CDC, 2013). For example, streptomycin was discovered in 1944 to treat tuberculosis, but a *Mycobacterium tuberculosis* strain that is resistant to the antibiotic appeared already during patient treatment (Davies & Davies, 2010). Emergence and subsequent spread of antibiotic resistance occurs so rapid that as a society we are running out of effective antibiotics to treat patients resulting in the increased prevalence of multi-drug resistant strains.

Antibiotic resistance is considered a pandemic. At least two million people are infected with antibiotic-resistant species resulting in 23,000 deaths each year in the United States (Frieden, 2013). The result of humans' overuse and misuse of antibiotics for therapeutic and prophylactic purposes and the unregulated disposal to the environment, bacterial species face constant selection pressure and they are winning the race against antibiotic development. If urgent action is not taken, such as tightly regulating the use of antibiotics globally, returning to the pre-antibiotic era will be inevitable. Extended periods of hospitalization due to longer treatment of infections will be an enormous financial burden on health care systems worldwide and we will lose lives to infectious diseases that were once treatable.

What is clear is that there should be constant and increased efforts in the search for new antibiotic targets and new antibacterial agents with complete understanding of their modes of action and molecular mechanisms. During my Ph.D. training, I focused on the structural and biochemical investigation of promising antibiotic protein targets in human pathogens. The goal is that this work can lead to development of specific and potent antibacterial agents. I have primarily focused on the three human pathogens *Campylobacter jejuni*, *Escherichia coli*, and *Mycobacterium tuberculosis*. In these organisms, my goal was to understand enzymes central to bacterial glycobiology.

Chapter 2

The structure of the UDP-Glc/GlcNAc 4-epimerase from the human pathogen *Campylobacter jejuni*

ABSTRACT

Worldwide, the food-born pathogen *Campylobacter jejuni* is the leading bacterial source of human gastroenteritis. *C. jejuni* produces a variety of diverse cell-surface carbohydrates that are essential for pathogenicity. A critical component of these oligo- and polysaccharides is the sugar *N*-acetylgalactosamine (GalNAc). The sole source of this sugar is the epimerization of UDP-*N*-acetylglucosamine (GlcNAc), a reaction catalyzed by the enzyme UDP-GlcNAc 4-epimerase (Gne). This enzyme is unique among known bacterial epimerases in that it also catalyzes the equivalent reaction with the non-*N*-acetylated sugars. Understanding how *Cj*Gne catalyzes these various interconversions is critical to designing novel inhibitors of this enzyme. Here, to further the mechanistic understanding we present a 2.0Å structure of *Cj*Gne with its NAD^+ co-factor bound. Based on novel features found in the structure we perform a variety of biochemical studies to probe the mechanism and compare these results to another structurally characterized bifunctional epimerase from humans GalE. We also show that ebselen, previously identified for inhibition of *Hs*GalE, is active against *Cj*Gne, suggesting a route for antibiotic development.

INTRODUCTION

Campylobacter jejuni, a microaerophilic pathogen, is a commensal of chickens and other avians and the most prevalent member of the *Campylobacter spp*., which are the leading causes of bacterial gastroenteritis worldwide (Kaakoush, Castaño-Rodríguez, Mitchell & Man, 2015; World Health Organization (WHO), 2017). It can be associated with post-infectious sequelae include Guillain-Barrè syndrome (Nachamkin, Allos & Ho, 1998), bacteremia (Fernández-Cruz et al., 2010), and reactive arthritis (Pope, Krizova, Garg, Thiessen-Philbrook & Ouimet, 2007). The *C. jejuni* glycome contains a number of surface-accessible carbohydrate structures required for interactions with the various hosts that include capsular polysaccharide (CPS), lipooligosaccharide (LOS), and N- and O-linked glycans (Abdi et al., 2012; Bacon et al., 2001; Day, Semchenko & Korolik, 2012; Ben N. Fry et al., 1998; Benjamin N. Fry et al., 2000; Guerry et al., 2002; Linton et al., 2005a). Variability in the nature of the surface glycans challenges the development of anti-*Campylobacter* therapies (Parkhill et al., 2000; Szymanski et al., 2003).

All of the exposed currently identified glycans in *C. jejuni* contain a GalNAc residue (Bernatchez et al., 2005; Mahdavi et al., 2014; Szymanski et al., 2003; N. M. Young et al., 2002), which is only produced by the epimerization of GlcNAc by the enzyme encoded by the gene *gne* (UDP-GlcNAc 4-epimerase). It was recently reannotated from *galE* (UDP-Gal 4-epimerase) due to the discovery of the UDP-GlcNAc epimerization activity (Bernatchez et al., 2005) (**Fig. 2.1.A**). In the *C. jejuni* genome (strain NCTC11168), this gene is located between the gene clusters of the N-linked glycosylation and LOS biosynthesis pathways (Karlyshev, Ketley & Wren, 2005). The gene in *C. jejuni* was first believed to be involved

in the LPS biosynthesis pathway (Ben N. Fry et al., 1998), but later a functional role in N-linked glycosylation was also determined (Linton et al., 2005b; Szymanski, Ruijin, Ewing, Trust & Guerry, 1999). Furthermore, experiments with an insertional mutant of *gne* indicated that it is responsible for providing GalNAc residues to the three major cell-surface glycoconjugates (Bernatchez et al., 2005). The bifunctional protein *Cj*Gne represents a potential therapeutic target as most oligosaccharides in *C. jejuni*, which contain a GalNAc residue, are required for pathogenesis (K. T. Young, Davis & DiRita, 2007). A mechanistic understanding is key to developing therapeutics that target this unusual epimerase.

UDP-hexose 4-epimerases, including *Cj*Gne, belong to a protein family of short-chain dehydrogenases/reductases (SDRs) (Kavanagh, Jörnvall, Persson & Oppermann, 2008). SDR enzymes possess diverse substrate specificity (Persson & Kallberg, 2013), as seen in the UDP-hexose 4-epimerases that have undergone detailed characterization (Beerens, Soetaert & Desmet, 2015). Structures of UDP-hexose 4-epimerases are available from all three domains of life: GalE from *Escherichia coli* (*Ec*GalE) (Thoden, Frey & Holden, 1996, 2002; Thoden & Holden, 1998), GalE from *Trypanosoma brucei* (*Tb*GalE) (Shaw et al., 2003), GalE from *Pyrobaculum calidifontis* (*Pc*GalE) (Sakuraba, Kawai, Yoneda & Ohshima, 2011), GalE from *Homo sapiens* (*Hs*GalE) (Thoden, Wohlers, Fridovich-Keil & Holden, 2000, 2001), GalE from *Thermotoga maritima* (*Tm*GalE) (Shin et al., 2015), WbpP from *Pseudomonas aeruginosa* (*Pa*WbpP) (Ishiyama, Creuzenet, Lam & Berghuis, 2004), and WbgU from *Plesiomonas shigelloids* (*Ps*WbgU) (Bhatt et al., 2011). A classification scheme has been proposed for substrate preference of the UDP-hexose 4-epimerases (Ishiyama et al., 2004). The scheme categorizes the epimerases in three different groups

depending on the side chain size of six key active site residues. Group 1 epimerases, such as *Ec*GalE and *Tb*GalE, interconvert only the non-acetylated moieties, UDP-Glc and UDP-Gal. Group 3 enzymes prefer epimerizing acetylated moieties and the examples are *Pa*WbpP and *Ps*WbgU. Group 2 members, which include *Cj*Gne and *Hs*GalE, catalyze interconversion of both the non-acetylated and acetylated UDP-hexoses. All of the UDP-hexose 4-epimerases structurally characterized so far are functional either in the Leloir pathway for galactose metabolism or the LPS O-antigen biosynthesis pathway. No epimerase known to act on multiple pathways has yet been a subject of structural studies.

Toward a mechanistic understanding of this bifunctional multi-pathway enzyme, here we provide a 2.0Å crystal structure of NAD^+-bound *Cj*Gne. Structural characteristics of *Cj*Gne that are common or distinct to its homologs are discussed. Based on the structural features and the epimerization results of the wild type and its mutants, we propose some critical residues of *Cj*Gne that are catalytically or structurally important for shaping its substrate-binding site and determining substrate specificity.

RESULTS

The overall architecture of CjGne bound with NAD^+

The gene for *Cj*Gne was cloned into an expression vector, induced in *E. coli*, purified by chromatography, and crystallized via conditions initially obtained through standard screens. Crystals diffracted to 2.0Å and a complete data set was collected. Data were processed using standard tools and the structure was solved by molecular replacement using *Ba*GalE (PDB entry: 2C20). The fully refined model contained a dimer of *Cj*Gne with an R-factor of 19.5% and an R_{free} of 22.5%. A complete atomic model containing residues 2–328 was attained for each with an RMSD of 0.38 Å between the two copies. Crystallographic statistics are found in **Table 2.1**. The two copies in the asymmetric unit are related by a two-fold rotational axis (**Fig. 2.1.B**). Consistent with this as a dimer interface, in solution the enzyme purified as a dimer as detected by SEC-MALLS (**Fig. 2.1.C**). For clarity, subunit A (indicated in **Fig. 2.1.B**) will be used in figures when onlyone is present.

The general structure of *Cj*Gne is consistent with that seen for related epimerases with a Glycosyltransferase A fold. Each subunit of *Cj*Gne contains two domains and one NAD^+ cofactor (**Fig. 2.1.D**). The domain that contains the N-terminus (M1-Y174, I230-H260, I295-D312) is composed of a central, twisted, parallel, seven-stranded, β-sheet flanked on each side by four α-helices. This domain resembles a 'Rossmann-fold' motif where NAD^+ binds that is commonly present in SDR enzymes (Lesk, 1995) and will be referred to as the NAD-binding domain (NBD). The substrate-binding domain (SBD) containing mostly C-terminal residues (F175-F229, G261-L294, D313-C328) is composed of three α-helices and two

parallel β-strands. Previous structural studies of related epimerases demonstrate that the UDP-hexose substrate resides in the cleft between the two domains with the UDP moiety primarily contacting SBD (Thoden, Frey, et al., 2002).

The conserved NAD-binding domain (NBD)

Both NAD^+ and NADH have been previously resolved in crystal structures of related epimerases (Thoden et al., 1996). The only structural difference between NAD^+ and NADH is in the nicotinamide ring (**Fig. 2.2.A**) and NAD^+ would have the nicotinamide ring that is planar. At this resolution, analyzing the $2F_o$-F_c map around the nicotinamide ring did not conclusively demonstrate whether or not the ring is planar. (**Fig. 2.2.B**). We instead compared F_o-F_c maps after refining the data with the model of NAD^+ and NADH. When NADH was used in refinement an additional positive density in the F_o-F_c map appeared in the nicotinamide ring; no density was seen for NAD^+ (**Fig. 2.2.B**). The unaccounted-for electron density with NADH is consistent with the additional electrons found in the nicotinamide ring of NAD^+, which we conclude is the major state in our crystal form. The nicotinamide ring adopts the *syn* conformation with respect to the ribose as seen in some of the *Ec*GalE (PDB ID: 1NAI, 1XEL, 1LRL) and all of the *Hs*GalE structures (PDB ID: 1EK5, 1EK6, 1HZJ).

The interaction with NAD^+ seen here is characteristic of related NAD-binding domains. The details were first described in the structure of GalE from *E. coli* where a series of conserved residues line the binding pocket (PDB ID: 1NAI) (Bauer, Rayment, Frey & Holden, 1992).

In *Cj*Gne, Ser34, and Lys35 forms one more and one less hydrogen bond than the corresponding residues in *Ec*GalE, respectively (**Fig. 2.2.C**). In *Ec*GalE, an additional residue Lys84 stabilizes the pyrophosphate group of NAD^+ via two hydrogen bonds, while the corresponding residue Ile82 in *Cj*Gne makes none (**Fig. 2.2.C**). Overall, seven residues (Asp31, Asn21, Ser34, Lys35, Asp56, Leu57, Asn97) are in hydrogen-bonding distance with the adenosine of NAD^+. Two residues (Tyr11, Ile12) and four residues (Phe78, Tyr146, Lys150, Tyr174) are interacting through hydrogen bonds with the pyrophosphate group and nicotinamide/ribose of NAD^+, respectively.

Apo-substrate-binding pocket is expanded

In a comparison to the other bifunctional epimerase *Hs*GalE, our apo-*Cj*Gne structure aligns best with the only available apo-GalE 'resting enzyme' from human (PDB ID: 1EK5) (Thoden et al., 2000) with an RMSD of 1.292 Å. When the SBDs of *Cj*Gne and *Hs*GalE are aligned, the RMSD value is 1.224 Å, close to the overall RMSD. In order to find differences in the substrate-binding sites, *Cj*Gne was aligned only with the SBD of *Hs*GalE bound to UDP-GlcNAc (PDB ID: 1HZJ) (**Fig. 2.3.A**). A careful look at the superposition revealed that while R300 and D303 of *Hs*GalE stabilize UDP-GlcNAc through some hydrogen bonds (**Fig. 2.3.B**), the corresponding residues (R287 and D290) of *Cj*Gne make no contact with UDP-GlcNAc (**Fig. 2.3.C**). Instead, D290 forms a hydrogen bond with R287, which in turn interacts with the backbone of Y190 and P191 (**Fig. 2.3.CD**). Interestingly, Y190 and P191 are part of the shifted loop in *Cj*Gne that will be described in the next section.

A shifted loop in the CjGne structure

Our crystal structure of *Cj*Gne revealed a surprising feature not found in any other structures of its related enzymes. In *Cj*Gne, one loop region (174-YFNAGACMDYTLGQRYPKATL-195) is shifted toward the NBD obstructing the substrate-binding pocket (**Fig. 2.4.A)**. In the crystal, this shift is supported by the presence of a disulfide bond between C181 and the C-terminal residue, C328 (**Fig. 2.4.A**). Both of these cysteine residues are not found in related epimerases, except for closely related *Campylobacter spp*; yet notably they always occur as a pair (**Fig. 2.4.BC**). We presume disulfide bond formation results in stabilization of the shifted loop. *Cj*Gne localizes in the cytoplasm and, under typical media conditions, this would be a reducing environment incompatible with a disulfide bond. During purification fresh DTT was added to all buffers (pH 7.5) to maintain a reducing environment. However, the presence of the disulfide suggested this was insufficient. DTT is reported to be readily oxidized above pH 7.5 (Han & Han, 1994) and in the presence of metal ions like Ni^{2+} contaminants from Ni column (Getz, Xiao, Chakrabarty, Cooke & Selvin, 1999). TCEP was added as an alternative to the final protein solution and incubated overnight before setting up crystal trays. It was also added to cryo-protecting conditions when harvesting crystals. Although TCEP is known to be stable at both acidic and basic conditions (Han & Han, 1994) and active with some metal contaminants, the crystals acquired always had the disulfide bond. This suggests that the formation of the disulfide is favored by the protein.

The internal cysteine forming the disulfide bond is important for activity

We first examined the *Cj*Gne-catalyzed percent conversion of the four substrates (UDP-Glc, UDP-Gal, UDP-GlcNAc, and UDP-GalNAc) at equilibrium using capillary electrophoresis

(CE). After 24 hours, when the sample reaction started from either UDP-Glc or UDP-Gal, the ratio for the integral areas reached 23:75 (UDP-Glc:UDP-Gal) at equilibrium irrespective of the starting substrate (**Fig. 2.5.AB** and **Table 2.2**). Similarly, in the presence of either UDP-GlcNAc or UDP-GalNAc as a substrate, the reactions reached a similar ratio of 28:71 (UDP-GlcNAc:UDP-GalNAc) at equilibrium (**Fig. 2.5.AB** and **Table 2.2**). With this assay, we confirmed the bifunctional activity of *Cj*Gne in interconverting similar amounts of non-acetylated and acetylated substrates.

The unexpected disulfide bond brings into question whether there may be biological importance to these residues. To examine the roles of C181 and C328, we expressed and purified various cysteine mutants (C181A, C181S, C328S, ΔC328, and C181S/ ΔC328) and analyzed their percent conversion after 24 hours (**Fig. 2.5.C** and **Table 2.2**). All of the C181 mutants (C181A, C181S, and C181S/ΔC328) resulted in significant loss of activity, less than 20% of wild type conversion, irrespective of the starting UDP-hexose. For the C-terminal cysteine, the C328S mutant lost more than 50% of the wild-type conversion. Yet surprisingly, the deletion of C-terminal C328 retained the full activity of the wild type regardless of the substrate added.

The similar HsGalE cysteine is not required for activity

In the multiple sequence alignment of *Cj*Gne and its homologs (**Fig. 2.4.E**), there is no epimerase that has a cysteine residue at the corresponding position to C181 of *Cj*Gne. Instead, the human epimerase among the homologs is the only example that has a cysteine residue (C196) within the corresponding loop region (*Hs* Y185-N207) to the previously

mentioned loop (*Cj* Y174-L195) in *Cj*Gne. We expressed and purified the C196A and C196S mutants of *Hs*GalE and investigated their epimerization activity at equilibrium in the reaction with each of the sugar substrates. Again, we confirmed the bifunctional nature of the enzyme. In this case, both mutants had no significant loss in activity relative to the wild type (**Fig. 2.5.C** and **Table 2.2**).

The CjGne internal cysteine affects the thermal stability of the enzyme

Thermal denaturation of the wild-type *Cj*Gne and its cysteine mutants was monitored through melting curves and calculated melting temperature (T_m) (**Fig. 2.5.D** and **Table 2.3**). The T_m values across the four UDP-hexose substrates within an enzyme stayed the same or had a difference less than 2°C. All the cysteine mutants of *Cj*Gne had the T_m values that are lower than the wild type by 2°C or more. The single and double mutants containing C181A decreased the T_m values the most by 8°C. Protein unfolding was also monitored for the *Hs*GalE C196 mutants (**Fig. 2.5.D** and **Table 2.3**). Unlike those of the *Cj*Gne mutants, the T_m values were retained regardless of the substrates.

Critical residues of CjGne for substrate binding and specificity

We performed a structural alignment between *Cj*Gne and *Hs*GalE bound with UDP-GlcNAc (PDB ID: 1HZJ) (**Fig. 2.6.A**). From the alignment, we identified five residues of *Cj*Gne that likely make polar contacts with the UDP-GlcNAc (T122, N176, P185, K192, and T194). Among them, both T122, either a serine or threonine, and N176 are conserved among the homologous sequences (**Fig. 2.4.E**). In order to examine their functional roles within the

predicted active site, we introduced mutations that retain the size, but lose polarity of the side chain: T122V and N176L (**Fig. 2.6.A**).

Based on the classification scheme for UDP-hexose 4-epimerases (Ishiyama et al., 2004), the T122 and N176 residues are two of the six key residues in *Cj*Gne, determining substrate specificity. Another key residue, L294, was also subject to mutation varying the size of hydrophobic side chains (valine, methionine, and tyrosine) to investigate its role in substrate specificity (**Fig. 2.6.A**). The corresponding residue of L294 in *Ec*GalE (Y299) was the first example that showed a single-residue mutation at this position to cysteine altered substrate specificity (Thoden, Henderson, Fridovich-Keil & Holden, 2002).

The equilibrium assay was performed with each of the mutants. For T122V and N176L, there was nearly a complete loss of conversion for all of the four UDP-sugar substrates (**Fig. 2.6.B**). This supports that T122 and N176 have catalytic roles that are essential for the epimerization activity of *Cj*Gne. T122 of *Cj*Gne is structurally aligned to S124 (*Ec*GalE), which is suggested to mediate catalysis by being hydrogen-bonded to Y149 (*Ec*GalE), an active site base, and to the 4'-hydroxyl group of the glucosyl ring of UDP-Glc (Yijeng Liu et al., 1997). Also, in *Hs*GalE, the corresponding residue, S132 interacts with the 4'-OH group of the hexose (**Fig. 2.6.A**). In *Cj*Gne, T122 and Y146 (Y149 in *Ec*GalE) are not in hydrogen-bonding distance.

N179 of *Ec*GalE (N176 in *Cj*Gne) has been suggested as one of the amino acid residues that contact the hexose portion of UDP-Glc/Gal within a hydrogen-bonding distance and thereby

important for binding of the substrate in the *E. coli* (Thoden et al., 1996). Similarly, N176 (*Cj*GalE) forms a hydrogen bond with 6'-hydroxyl group of UDP-GlcNAc bound to *Hs*GalE, but through the amino group. Alternatively, the amino group of N187 (*Hs*GalE) interacts directly with the two phosphoryl oxygen atoms of the β-phosphate group of the nucleotide (**Fig. 2.6.A**). With substrate bound, N176 may re-orient in *Cj*Gne to make this same interaction.

The epimerization assay was also performed for the L294V, L294M, and L294Y mutants to investigate if the size of the L294 side chain alters substrate specificity of *Cj*Gne (**Fig. 2.6.B**). None of the mutants retained activities at a comparable level to those of the wild type with all four substrates. However, interconversion of non-acetylated substrates was least disrupted with L294V among the mutants, whereas L294M was most favorable for activity than the other two mutants with acetylated substrates.

Ebselen inhibiting HsGalE is a potential inhibitor of CjGne

Few studies have reported inhibitors of GalE, and none of them are promising for further drug development. One literature reported ebselen, an organoselenium compound, inhibits *Hs*GalE with nanomolar IC_{50} (0.014 μM) (Urbaniak et al., 2006). Since *Cj*Gne possesses a bifunctional activity just like *Hs*GalE, we predicted that the inhibitors that are potent to *Hs*GalE would also exert inhibitory effect on *Cj*Gne. An inhibition assay of *Cj*Gne using capillary electrophoresis was performed in the presence of ebselen (**Fig. 2.7**). Indeed, the inhibition was almost complete when ebselen was incubated with the enzyme and either

UDP-Gal or UDP-GalNAc, while the results were slightly less dramatic in the reactions with either UDP-Glc or UDP-GlcNAc (13% and 22% of the wild-type activity was retained, respectively).

DISCUSSION

In this study, we provide the first structural and biochemical characterization of *Cj*Gne, the bifunctional UDP-Glc/GlcNAc 4-epimerase which catalyzes the production of GalNAc, a critical component of the major surface polysaccharides in *C. jejuni.* This functionally separates Gne from other UDP-Glc/GlcNAc 4-epimerases that either work in the Leloir pathway for galactose metabolism or are part of the LPS biosynthesis pathway.

A surprising result is the presence of a structural disulfide bond in *Cj*Gne. Expressed heterologously in the *E. coli* system, one expects cysteines to be reduced in the cytosol (Hatahet, Boyd & Beckwith, 2014). Other exceptions have been reported where structural disulfides in cytosolic proteins of thermophilic archaeal species protect them from denaturation at high temperature (Jorda & Yeates, 2011). *C. jejuni* is a moderate thermophilic species thriving at 37–42°C (Hofreuter, 2014) with a T_m for *Cj*Gne at 57°C (**Fig. 2.5.D** and **Table 2.3**). The presence of a structural disulfide bond might help *Cj*Gne tolerate growing at high temperature.

The cysteine at position 181 is not conserved (**Fig. 2.4.E**), yet replacing C181 with either alanine or serine led to a loss of activity. In the structure, a water mediated network connects the phosphate groups of the substrate to the backbone of the C181 residue. These are unlikely to be disrupted by the simple mutations. When the other half of the disulfide pair, C328, was replaced with a serine, less than 50% wild-type conversion was retained. Surprisingly, deletion of the C-terminal cysteine (C328) was fully active. In total, we can conclude that the disulfide bond is not required for catalysis. However, the cysteines are important for full

activity. Cysteines often make hydrophobic contacts and the effects from the loss of the spatial hydrophobic environment may provide rationale for the reduction in activity upon mutation to serine. Supporting this, all the cysteine mutants of *Cj*Gne reduced the overall stability of the protein (**Fig. 2.5.D** and **Table 2.3**). This suggests general structural roles that could affect catalysis.

A major feature of our *Cj*Gne structure is the shifted loop region (Y174-L195) toward the active site. The presence of the disulfide suggests that this may fix the conformation of the loop. The sequence of this region is conserved across other *Campylobacter* species and the electron density shows the conformation is well ordered. In this conformation, the protein would be expected to be inhibited, perhaps a response to the complicated environments experienced by *C. jejuni*.

While the cysteines play a functional role in *C. jejuni*, a related cysteine in the functional homolog *Hs*GalE was not required for full activity. In the structure of the human, there is no shift of the corresponding *Cj*Gne loop (Y174-L195). This lends credence to the importance of these residues in the *C. jejuni* enzyme.

The mechanism of this class of epimerases has been elucidated over time; however, there remain some details to be resolved. We chose candidate residues that are predicted to be critical in substrate binding or specificity from the structural and sequence alignments with the human epimerase in complex with NADH and UDP-GlcNAc (PDB entry: 1HZJ). Both of the T122V and N176L mutants killed almost all of the activity irrespective of the substrate

tested. This result is indicative of catalytic importance of the residues. The highly conserved serine residues of the other homologs aligned with T122 of *Cj*Gne have been shown to play a role in mediating the electron transfer during the catalytic mechanism. The asparagine residue corresponding to N176 in *Cj*Gne is also highly conserved across species. The N195 residue in the *Ps*WbgU (N176; *Cj*Gne), for example, forms two hydrogen bonds via the carbonyl oxygen and the amine group with the NH-group and the oxy bridge of the diphospho moiety of UDP-GlcNAc, respectively. Based on the high conservation and previous functional characterization in other homologs, T122 and N176 of *Cj*Gne are likely to require polar side chains to be either directly or indirectly involved in catalysis and to interact with the substrate bound, respectively.

Along with T122V and N176L, L294 of *Cj*Gne was subjected for mutagenesis studies because there were cases where mutation of the single residue in the same position in its homologs altered substrate specificity. Previous studies demonstrated that Y299C of *Ec*GalE (Thoden, Henderson, Fridovich-keil & Holden, 2002) allowed, but C307Y of *Hs*GalE (Schulz et al., 2004) and S306Y of *Ec*GalE$_{O86:B7}$ (Guo, Li & Wang, 2006) lost the conversion activity of acetylated substrates, respectively. In contrast, mutation of the corresponding residue (S317) from *Pa*WbpP to tyrosine resulted in complete loss of activity, so no further insights regarding side chain size of this residue for substrate specificity was available (Ishiyama et al., 2004). We mutated the residue in the same position, L294, of *Cj*Gne into valine, methionine, and tyrosine to see if various sizes of hydrophobic side chains can affect the activities of any UDP-hexose substrate. We predicted L294M retains bifunctionality of *Cj*Gne, whereas L294V and L294Y preferentially convert acetylated and non-acetylated

sugar substrates, respectively. None of the three mutants of L294 retained the full activity of the wild type, but showed similar levels of reduction between UDP-Glc and UDP-Gal and generally more reduction with UDP-GalNAc. Approximately 50% reduction in L294V activity with non-acetylated substrates can be explained by the smaller hydrophobic side chain of valine increases the active site volume, rendering the binding to non-acetylated and/or acetylated substrates less specific. Also, the hydrophobic interaction between leucine and the methyl group of the *N*-acetylated moiety on C2 on UDP-GlcNAc from the human enzyme seems to be disrupted in *Cj*Gne by the replacement with valine, thereby resulting in a larger reduction in conversion activity of the acetylated substrates. Methionine is in similar size as leucine in L294M mutant and activities for both non-acetylated and acetylated substrates were all at similar levels except for UDP-GlcNAc, but the epimerization is much less efficient than that catalyzed by the wild type, which is probably because of different chemical properties of sulfur replaced from carbon and/or the different position of a methyl group. In L294Y mutant, the bulky side chain of tyrosine seemed to reduce the active site volume, thus leading to the highest loss of epimerization with UDP-GalNAc. Taken together, we demonstrate the variants of the active site residue (L294) can alter substrate specificity, although the previously predicted pattern did not fit here.

Lastly, the inhibition of *Cj*Gne by ebselen provides evidence for this as a small molecule target, but ebselen itself is not a likely route as it has been reported to target myriads of biological pathways (Azad & Tomar, 2014). It will be important to find inhibitors that are specific to *Cj*Gne. The search for such inhibitors is ongoing.

In conclusion, based on a high-resolution *Cj*Gne/NAD^+ crystal structure and biochemical data from mutants, we proposed some critical residues of *Cj*Gne that are catalytically or structurally important. These findings along with its observed susceptibility to a *Hs*GalE inhibitor, suggest a route for antibiotic development.

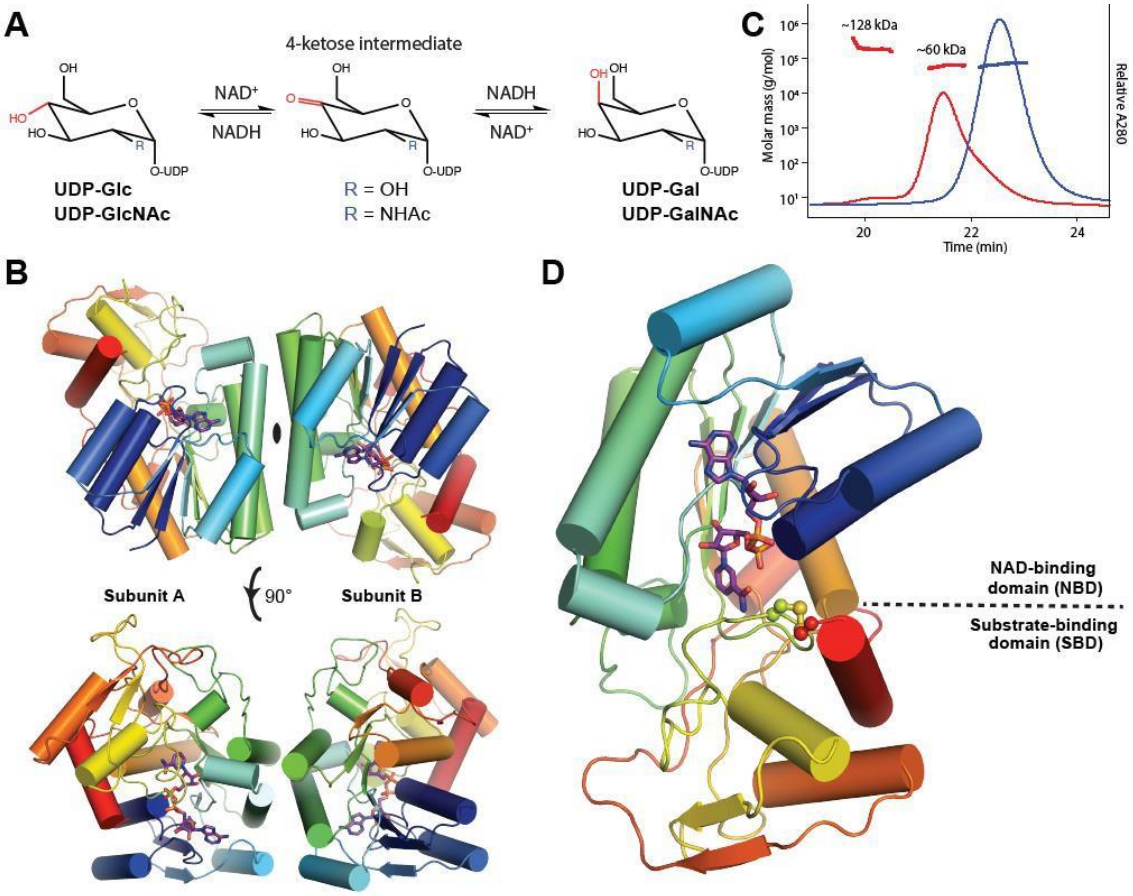

Figure 2.1. The 2.0Å crystal structure of *Cj*Gne in complex with NAD$^+$. A, A catalytic reaction scheme of *Cj*Gne involving four substrates, UDP-Glc, UDP-Gal, UDP-GlcNAc, UDP-GalNAc, and 4-ketose intermediate. B, (Top) An asymmetric unit contains two molecules of *Cj*Gne with a two-fold symmetry. (Bottom) The two molecules are rotated by 90°. In both cases, a monomer in the cartoon representation is colored with rainbow; N-terminus is with red and C-terminus is with blue. One NAD$^+$ molecule in each monomer is shown in stick representation with the purple backbone. C, The SEC-MALL result shows *Cj*Gne forms a dimer (~72 kDa; blue) in solution as compared to the bovine serum albumin (BSA) standard (monomer: ~60 kDa; dimer: ~128 kDa; red). D, A monomer is colored the same as in A. Additionally, the disulfide bond is depicted in yellow spheres. NBD and SBD are also labeled.

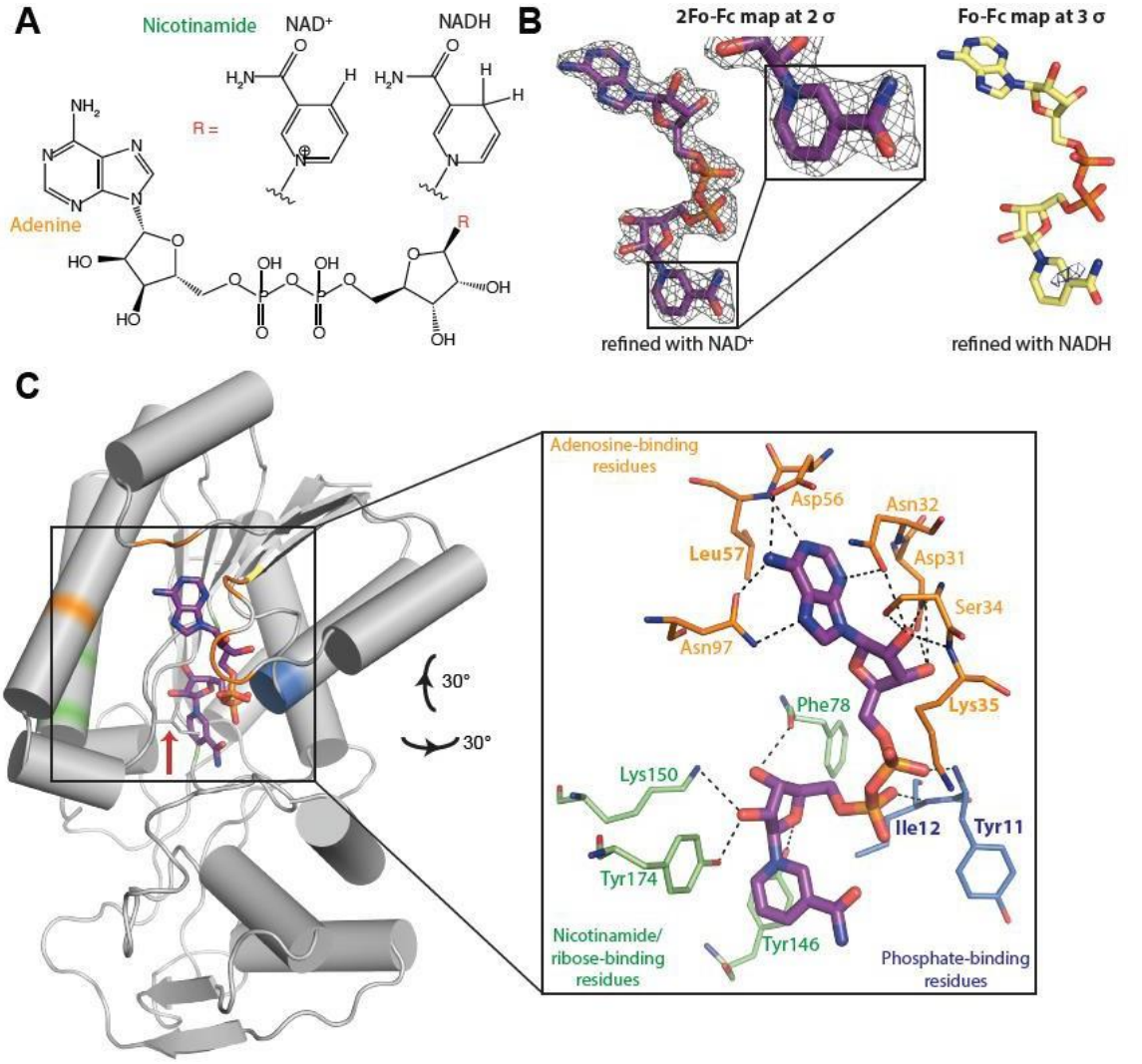

Figure 2.2. NAD$^+$ and NBD. A, Comparison of the chemical structures of NAD$^+$ and NADH. The only difference lies in the nicotinamide ring, depicted as a R group. B, On the left, the $2F_o$-F_c map that is refined with NAD$^+$ is shown in mesh at 2 σ and the magnified view of the nicotinamide ring. On the right, the F_o-F_c map that is refined with NADH is shown in mesh at 3 σ. C, On the left, a monomer of *Cj*Gne in the same orientation as in Fig. 1D is colored gray. NAD$^+$ is shown in stick representation and its backbone is purple. Ile82 is indicated by a red arrow. On the right, this is a magnified view of the NBD. The residues that are in a hydrogen-bonding distance to adenosine, phosphates, and nicotinamide/ribose are shown in stick representation and colored orange, blue, and green, respectively. The residues that interact with NAD$^+$ via backbone are bolded. Hydrogen bonds are shown in black dashed lines.

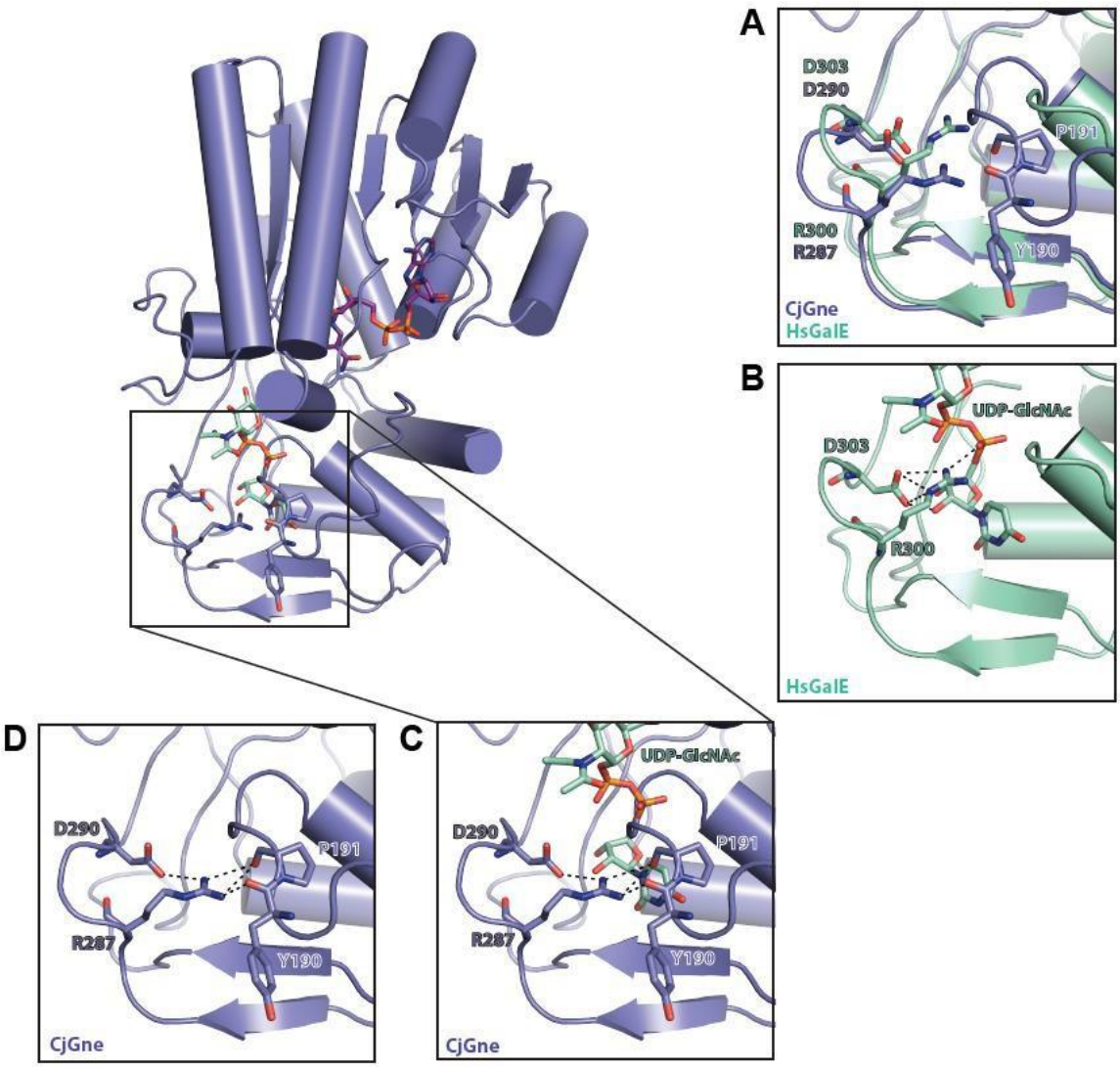

Figure 2.3. Substrate-binding domain (SBD) of *Cj*Gne. A *Cj*Gne monomer is shown in complex with NAD^+ and UDP-GlcNAc from *Hs*GalE (PDB ID: 1HZJ). A, Aligned *Cj*Gne (slate) and the *Hs*GalE SBD (green cyan) in the box region. Y190, P191, R287, and D290 of *Cj*Gne and R300 and D303 of *Hs*GalE are shown in stick representation. B, Same region as A, but only *Hs*GalE is shown along with the polar contacts in black dashed lines. C, Same region as A, but *Cj*Gne and UDP-GlcNAc from *Hs*GalE are shown with polar contacts in black dashed lines. D, Same region as A, but only *Cj*Gne is shown with polar contacts between R287/D290 and Y190/P191.

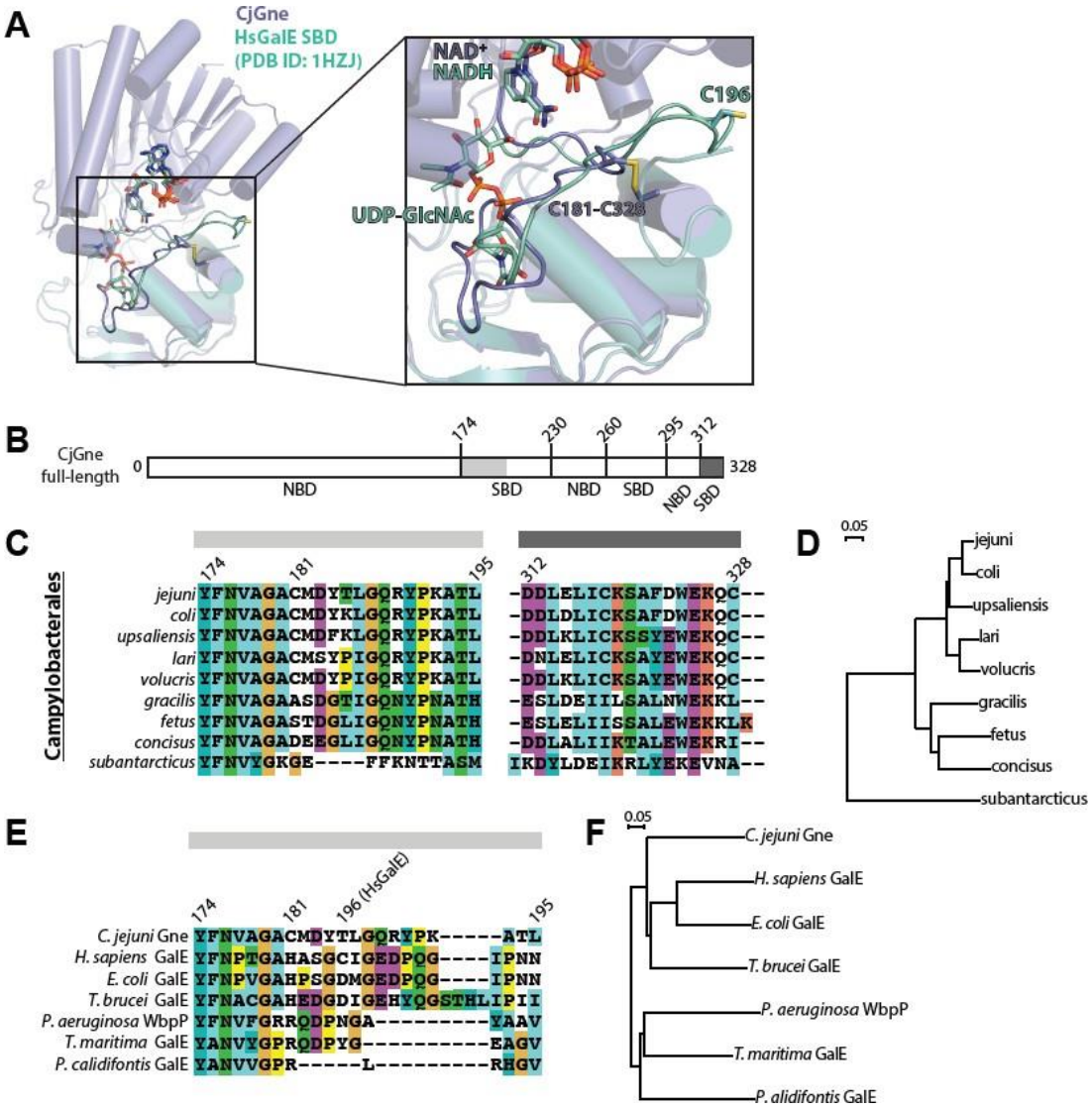

Figure 2.4. Unique features of *Cj*Gne. A, Structural alignment between *Cj*Gne/NAD$^+$ (slate) and *Hs*GalE/NADH/UDP-GlcNAc (PDB ID: 1HZJ) (green cyan). The SBD is magnified in the box. NAD$^+$, NADH, and UDP-GlcNAc are shown in the stick representation. C181 and C328 of *Cj*Gne along with the disulfide bond and C196 of *Hs*GalE are also in the stick representation. B, The full length of *Cj*Gne. C, Multiple sequence alignment of Campylobacterales was performed in ClustalX 2.1. Only two alignment regions (Y174-L195 and D312-C328; *Cj*Gne numbering) that are highlighted in the sequence scheme on the top are shown. The background coloring is as follows: aromatic (cyan), hydrophobic (blue), polar (green), glycines (orange), negative charge (purple), positive charge (red), prolines (yellow), and unconserved (white). D, A phylogenetic tree of the Campylobacterales was drawn in NJplot. E, Multiple sequence alignment of *Cj*Gne and its homologs. The alignment with only the region from Y174 to L195 (*Cj*Gne numbering) is shown here. The sequences that are used here are, from the top, Gne from *Campylobacter jejuni*, GalE from *Homo sapiens*, GalE from *Escherichia coli*, GalE from *Trypanosoma brucei*, WbpP from *Pseudomonas aeruginosa*, GalE from *Thermotoga maritima*, and GalE from *Pyrobaculum calidifontis*. The color scheme is the same as in C. F, A phylogenetic tree of the *Cj*Gne homologs was drawn in NJplot.

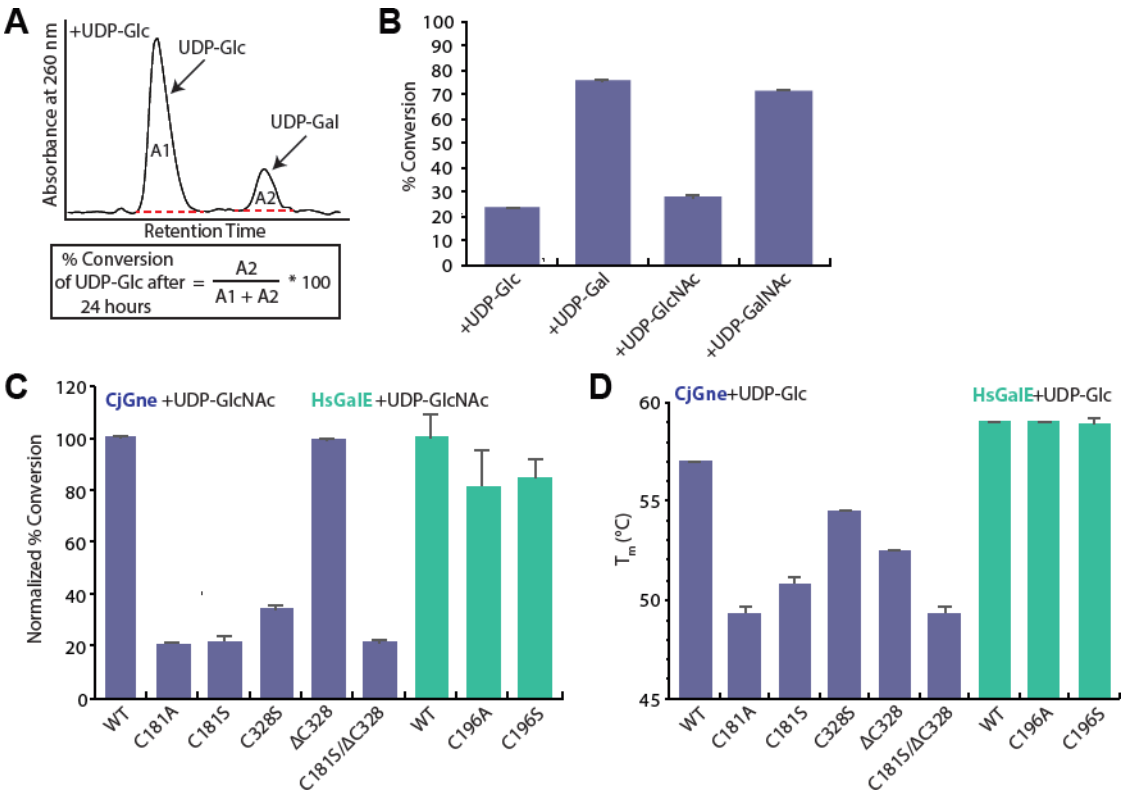

Figure 2.5. Epimerization assays of the wild-type and cysteine mutants of *Cj*Gne and *Hs*GalE. A, An example of the % conversion of UDP-Glc to UDP-Gal by *Cj*Gne measured in capillary electrophoresis. The equation used to calculate % conversion is provided in the box. B, The results of % conversion of each of four substrates by *Cj*Gne. Each reaction was repeated for three times and the error bars indicate the standard deviations. C, Normalized % conversion of UDP-GlcNAc by the wild type and cysteine mutants of *Cj*Gne (slate) and *Hs*GalE (green cyan). Each reaction was repeated for three times and the error bars are based on the calculation provided in Materials and Methods. D, The melting temperature values (T_m) of the wild type and cysteine mutants of *Cj*Gne (slate) and *Hs*GalE (green cyan) in the presence of UDP-Glc. Each measurement was repeated for three times and the error bars indicate the standard deviations.

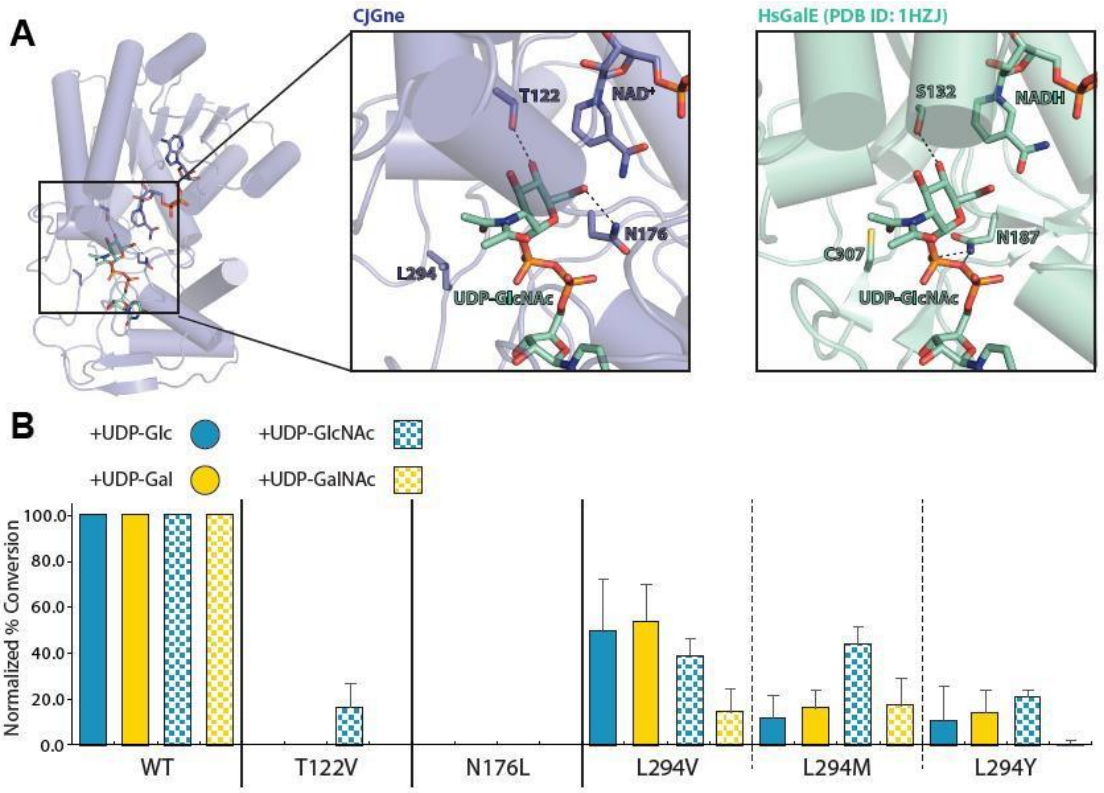

Figure 2.6. Predicted active site of the *Cj*Gne/NAD$^+$ complex. A, On the left, structural alignment between *Cj*Gne/NAD$^+$ and UDP-GlcNAc from the *Hs*GalE/NADH/UDP-GlcNAc (PDB ID: 1HZJ). In the center, the putative active site of *Cj*Gne is magnified with T122, N176, and L294 labeled. On the right, the catalytic site of *Hs*GalE is magnified to the same extent as that with *Cj*Gne with S132, N187, and C307 labeled. Hydrogen bonds are indicated as black dashes. B, Normalized % conversion of UDP-Glc, UDP-Gal, UDP-GlcNAc, and UDP-GalNAc by the wild type and catalytic mutants of *Cj*Gne. Each reaction was repeated for five times except for N176L only for once. The error bars are based on the calculation provided in Materials and Methods.

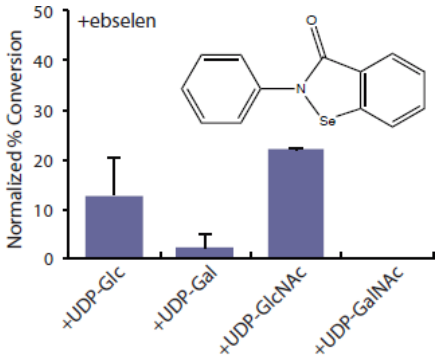

Figure 2.7. Inhibition of *Cj*Gne by ebselen. Chemical structure of ebselen is shown in the top right corner. Each inhibition reaction was repeated for three times except for the one with UDP-GalNAc. The error bars are calculated based on the equation provided in Materials and Methods.

Data collection	
Space group	$P4_1\,2_1\,2$
Cell dimensions	
a/b/c (Å)	87.7/ 87.7/ 261.66
α/β/γ (°)	90/ 90/ 90
Resolution (Å)	65.42–2.04 (2.09–2.04)*
R_{merge}	0.052 (0.677)
I/σ (I)	18.0 (2.4)
Completeness (%)	99.9 (100)
Redundancy	6.6 (6.8)
Refinement	
Resolution (Å)	62.0 – 2.0
No. reflections	65,906
R_{work}/ R_{free}	19.5/ 22.5
No. atoms	5,463
Protein	5,290
Ligand/ion	6
Water	167
Mean B-factor	38.1
RMS deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.1

Table 2.1. Statistics of X-ray data collection and refinement. Values in parentheses are from the highest resolution shell. RMS, root mean square.

Enzymes	UDP-Glc		UDP-Gal		UDP-GlcNAc		UDP-GalNAc	
	raw	normalized	raw	normalized	raw	normalized	raw	normalized
***Cj*Gne (WT)**	**23.2 ± 0.2**	**100.0 ± 1.2**	**75.3 ± 1.0**	**100.0 ± 1.9**	**27.7 ± 0.2**	**100.0 ± 1.1**	**71.4 ± 0.7**	**100.0 ± 1.4**
C181A	1.7 ± 0.5	7.5 ± 2.1	0.6 ± 0.3	0.7 ± 0.4	5.8 ± 0.2	20.8 ± 0.7	1.4 ± 1.2	2.0 ± 1.7
C181S	1.2 ± 0.1	5.3 ± 0.5	1.0 ± 0.4	1.4 ± 0.5	6.1 ± 0.5	22.1 ± 1.9	0.9 ± 0.8	1.3 ± 1.1
C328S	8.8 ± 3.2	37.9 ± 13.8	10.6 ± 3.9	14.0 ± 5.2	9.4 ± 0.4	34.0 ± 1.5	14.4 ± 3.3	20.1 ± 4.6
ΔC328	23.6 ± 0.2	101.9 ± 1.2	76.0 ± 0.3	100.9 ± 1.4	27.5 ± 0.1	99.3 ± 0.9	72.1 ± 0.7	100.9 ± 1.4
C181S/ΔC328	1.1 ± 0.5	4.8 ± 2.3	0.6 ± 0.3	0.8 ± 0.4	5.9 ± 0.2	21.4 ± 0.9	0.7 ± 0.6	0.9 ± 0.8
T122V[*]	0 ± 0	0 ± 0	0 ± 0	0 ± 0	4.5 ± 2.9	16.3 ± 10.3	0 ± 0	0 ± 0
N176L[#]	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
L294V[*]	11.4 ± 5.2	49.4 ± 22.5	40.7 ± 11.9	54.0 ± 15.8	10.8 ± 2.0	39.0 ± 7.2	10.5 ± 6.9	14.7 ± 9.7
L294M[*]	2.7 ± 2.3	11.8 ± 9.7	12.1 ± 5.8	16.1 ± 7.7	12.1 ± 2.1	43.8 ± 7.5	12.5 ± 8.5	17.4 ± 11.9
L294Y[*]	2.4 ± 3.5	10.4 ± 15.3	10.6 ± 7.4	14.1 ± 9.9	5.8 ± 0.7	20.8 ± 2.5	0.4 ± 1.0	0.6 ± 1.4
WT + ebselen	2.9 ± 1.8	12.6 ± 7.8	1.5 ± 2.1	2.0 ± 2.7	6.1 ± 0.1	22.0 ± 0.5	0 ± 0[§]	0 ± 0
***Hs*GalE (WT)**	**25.2 ± 1.0**	**100.0 ± 5.7**	**72.6 ± 0.6**	**100.0 ± 1.2**	**23.3 ± 1.4**	**100.0 ± 8.5**	**56.3 ± 5.9**	**100.0 ± 14.9**
C196A	25.1 ± 1.1	99.8 ± 5.9	62.8 ± 4.0	86.5 ± 5.6	18.9 ± 3.0	81.3 ± 13.9	47.9 ± 3.1	85.1 ± 10.5
C196S	25.7 ± 1.2	102.2 ± 6.4	68.4 ± 1.4	94.2 ± 2.1	19.7 ± 1.1	84.8 ± 6.9	52.0 ± 4.7	92.4 ± 12.8

Table 2.2. Average percent conversion values and their standard deviations of the substrates by *Cj*Gne and *Hs*GalE. The protein samples include the wild type and mutants. Ebselen, an inhibitor, was added only to *Cj*Gne. The normalized average percent conversion values and standard deviations of UDP-hexose substrates are written (Calculations are reported in Materials and Methods). The [#], [§], and [*]marks refer to a single, duplicate, and quintuplicate measurement, respectively. The rest of the values are from triplicates.

Enzymes	Melting Temperature (°C)			
	UDP-Glc	UDP-Gal	UDP-GlcNAc	UDP-GalNAc
*Cj*Gne (WT)	57.0	57.2 ± 0.3	58.5	58.0
C181A	49.3 ± 0.3	49.5	51.3 ± 0.3	50.5
C181S	50.8 ± 0.3	50.7 ± 0.3	53.0 ± 0.5	51.5
C328S	54.5	54.2 ± 0.3	55.5	55.0
ΔC328	52.5	52.0	53.8 ± 0.6	52.5
C181S/ΔC328	49.3 ± 0.3	49.5	50.2 ± 0.3	50.2 ± 0.3
*Hs*GalE (WT)	59.0	59.0*	59.0	59.0
C196A	59.0	59.0	59.0	59.0
C196S	59.0 ± 0.2	59.0	59.0	NA

Table 2.3. Average melting temperatures (°C) and their standard deviations of the wild-type and mutants of *Cj*Gne and *Hs*GalE. A * mark refers to a single measurement. The rest of the values are from at least duplicates. Standard deviations with zero are not written.

MATERIALS AND METHODS

Materials and chemicals

Genomic DNA of *C. jejuni* NCTC11168 (#700819D-5) was purchased from ATCC (Manassas, VA). Two constructs of *Hs*GalE with either N-terminal (pET28a vector) or C-terminal (pET31B vector) hexahistidine (His_6) tags were kindly provided by the Holden Group at the University of Wisconsin, Madison. The former was used for capillary electrophoresis and the latter was used for stability assays, each described below. *E. coli* NiCo21(DE3) competent cells were obtained from New England Biolabs Inc. (Ipswich, MA). UDP-Glucose, UDP-Galactose, UDP-GlcNAc, UDP-GalNAc, NAD^+), dithiothreitol (DTT), tris (2-carboxyethyl) phosphine (TCEP), and bovine serum albumin (BSA) were from Sigma-Aldrich (St. Louis, MO). Ebselen (2-phenyl- 1,2-benzisoselenazol-3(2H)-one) was from Cayman Chemical (Ann Arbor, MI).

Cloning of gne

The plasmid encoding *gne* (*cj1131c*) of *C. jejuni* was constructed by one-step enzymatic DNA assembly through Gibson cloning (Gibson et al., 2009). Briefly, the *gne* gene and the plasmid for insertion were amplified from genomic DNA of *C. jejuni* NCTC11168 and a pET33b-derived vector, respectively, using primers with ca. 40 bp of homology to each other. 5 μL of DNA sample (0.7 μL of the amplified *gne* gene (10 ng/μL) plus 4.3 μL of the amplified vector (10 ng/μL)) was added to 15 μL of a master mix solution including T5 exonuclease, Phusion DNA polymerase, and Taq DNA ligase, then incubated at 50°C for 60 min. To avoid self-colonies from the template-vector, the PCR product amplified from the

vector was treated with *DpnI*. The final construct contained a N-terminal his_6-tag followed by a thrombin cleavage site (N-**M**GGS**HHHHHH**G**LVPRGS**-*gne*-C). All DNA constructs were confirmed by sequencing.

Generation of point mutations

All mutations, including T122V, N176L, L294V, L294M, L294Y, C181A, C181S, C328S, ΔC328, C181S-ΔC328, C196A (*Hs*GalE), and C196S (*Hs*GalE) were prepared in a mixture solution containing Phusion® High-Fidelity PCR Master Mix with HF Buffer from New England Biolabs Inc. (Ipswich, MA), the DNA template of the QK45AA mutant (based on the surface entropy reduction prediction (Cooper et al., 2007)), and primers with or without 5% DMSO (Chester & Marshak, 1993). The recommended PCR protocol for using the Phusion® HF PCR MM from NEB Inc. was used.

Expression and purification of CjGne

Constructs were transformed and expressed in *E. coli* NiCo21(DE3) strain that are originally derived from BL21(DE3). Cells were grown in cultures of 2xYT with kanamycin (35 μg/mL) to an optical density of OD_{600} 0.5–0.7, then induced with 0.4 mM isopropyl β-D-thiogalactoside (IPTG, Anatrace, Maumee, OH) at 37°C for 6 h. Cells were harvested by centrifugation (4000 rpm, 20 min, 4°C), resuspended in lysis buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM DTT), and lysed by three passes through a microfluidizer (18 kpsi), and centrifuged at 12,000 rpm for 20 min at 4°C. Proteins were purified from the supernatant by nickel-affinity chromatography followed by size-exclusion chromatography.

Briefly, the cell lysate was modified to 10 mM imidazole then passed through a 1 mL pre-equilibrated Ni-NTA agarose resin (Qiagen, Germantown, MD), washed with 30x column volumes of 20 mM imidazole-containing buffer, then eluted with 20 mL of 250 mM imidazole-containing buffer. The elution was concentrated using a 10-kD cut-off Amicon® Ultra-4 Centrifugal Filter (Millipore) and further purified by either a Superdex-200 16/60 or Superdex-200 10/300 gel-chromatography column (GE Healthcare, Little Chalfont, UK). The fractionated protein was concentrated again using the Amicon® Centrifugal Filter, flash frozen, and stored at -80°C.

Size-exclusion chromatography and multi-angle light scattering (SEC-MALS)

Purified protein was analyzed by SEC-MALLS (Wyatt Technology, Santa Barbara, CA). Briefly, a Shodex KW-804 column (Showa Denko America, Inc., New York, NY) was equilibrated in running buffer (20 mM HEPES, pH 7.5, 100 mM NaCl, 10% glycerol, 10 mM βME). 100 μg of either BSA as a standard or *Cj*Gne were injected and run at 0.5 mL/min for 30 min. Data analysis was performed using Astro 5.3.4 software. For the sample run, peaks from LS, dRI, and UV (280 nm) detectors were aligned by defining baseline and applying band broadening values from the BSA run.

Crystallization of CjGne and X-ray diffraction

For crystallization, the purified QK45AA optimized variant of *Cj*Gne was pre-incubated with 5 mM UDP-GlcNAc at 4°C overnight. Crystallization screening was performed by sitting-drop vapor-diffusion with commercially available screens (Hampton Research, Qiagen,

Emerald BioSystems) and then incubated at room temperature. Initial conditions were refined by additive screening using the Additive Screen™ (Hampton Research). The final drop consisted of 0.2 μL of mother liquor (1.3 M sodium acetate trihydrate (pH 7.0) with 50 mM sodium malonate) and 0.2 μL of protein (16–20 mg/mL). Crystals grew to full- size after several days. For cryo-protection, crystals were transferred to a drop containing 70% reservoir solution and 30% glycerol for 5 sec then flash frozen in liquid nitrogen. Diffraction data were collected from a single crystal at beamline 12–2 at the Stanford Synchrotron Radiation Lightsource (SSRL).

Structural determination and refinement

Images were collected on a Dectris Pilatus 6M pixel detector. Diffraction data were integrated with XDS (Kabsch, 2010) and scaled with SCALA in CCP4 (Winn et al., 2011). Crystals were in the space group $P4_12_12$ with unit cell dimensions *a=b=87.7 c=261.66* and a complete dataset was collected to 2.0 Å. The asymmetric unit contained two copies of *Cj*Gne (residues 2–328), two NAD^+, two acetate, seven glycerol, and 167 water molecules. Phases were obtained by molecular replacement using the structure of UDP-Glc 4-epimerase from *Bacillus Anthracis* as a search model (PDB entry: 2C20; 42% identity) in Phaser as implemented in Phenix (Adams, Afonine, et al., 2010; McCoy, 2006). Manual model building was performed using Coot (Emsley, Lohkamp, Scott & Cowtan, 2010). *Cj*Gne was refined in Phenix with final R-factor of 19.5% (R_{free} = 22.5%). Statistics for data collection and structure determination are found in **Table 2.1**.

Epimerization assay using capillary electrophoresis

Enzyme reactions were performed in 100 μL of reaction mixture containing 50 mM Tris-HCl (pH 8.0), 1 mM of UDP-sugar, 1 mM of NAD^+, and 50 ng of *Cj*Gne at 37°C. The reaction was stopped after 24 hours by boiling for 5 min and then centrifuged (12,000 rpm, 20 min) to remove protein aggregates. In the case of the inhibition assay with ebselen, enzyme reactions were prepared in 20 μL containing 50 mM Tris-HCl (pH 8.0), a 1 mM of UDP-sugar, 1 mM of NAD^+, 50 ng of *Cj*Gne, and 100 μM ebselen at 37°C for 24 hours. The reactions were then stopped by boiling for 5 min and centrifuged. The samples were monitored by HP 3DCE capillary electrophoresis instrument equipped with a UV-VIS DAD (Agilent Technologies, Santa Clara, CA, USA). The 50 cm-long capillary was packed with fused silica and the running buffer was 20 mM sodium tetraborate decahydrate, pH 9 (J.T. Baker®, Avantor Performance Materials, Center Valley, PA). The capillary was preconditioned for each run by washing with the running buffer for 2 min. Each sample was injected by pressure of 50 mbar for 10 sec and the separation was performed at 30 kV and detected at 260 nm (330 nm as background). The average retention times for UDP-Glc, UDP-Gal, UDP-GlcNAc, and UDP-GalNAc were 8.7 min, 8.9 min, 8.3 min, and 8.5 min, respectively. The peak area was estimated (**Fig. 2.5.A**) and integrated using 3D-CE Chemstation Rev. A.09.03.

A fluorescence-based thermal shift assay

The thermal shift assay was based on Niesen et al. (Niesen, Berglund & Vedadi, 2007). A real-time PCR device (CFX96 from Bio-Rad, Hercules, CA) was used to monitor protein unfolding through fluorescence by SYPRO Orange (Sigma-Aldrich, St. Louis, MO) at 2x

concentration (1:2500 dilution of 5000x stock). Fluorescence was measured using the FRET configuration. This configuration excites and detects in all six channels that the instrument has. This setting is used because no single channel contains that appropriate excitation and emission filter. Protein samples (2 μM) in 20 mM Tris buffer (pH 7.5) containing 100 mM NaCl, 1 mM NAD^+, 1mM DTT, 1mM of a UDP-sugar substrate in a reaction volume of 50 μL were mixed in 96-well PCR plates (Bio-Rad, Hercules, CA). The plates were briefly spun down at 1000 rpm and placed in the device. Melting curves were measured starting with a 15-minute pre-chilling at 15°C then increased in 0.5°C steps to 95°C with 30 second incubation. The fluorescence intensity at the end of each step is plotted as a function of temperature. The resulting sigmoidal curve was best fit to a two-state transition. The inflection point of the melting curve, melting temperature (T_m), was calculated using the internal PCR software.

ACKNOWLEDGEMENTS

We thank Hazel Holden (Wisconsin) for expression plasmids for *Hs*GalE. We are grateful to Nathan Dalleska (Caltech) for assistance with the capillary electrophoresis at the Environmental Analysis Center and Jens Kaiser and Pavle Nikolovski (Caltech) for crystallography help through the Molecular Observatory. Operations at SSRL are supported by the U.S. Department of Energy and the National Institutes of Health (NIH). This work was supported by Pioneer and MraY grant.

C h a p t e r 3

Toward structural and mechanistic understanding of key membrane-bound enzymes, MraY and MurG, in the peptidoglycan biosynthesis pathway

ABSTRACT

Antimicrobial resistance from bacterial infections has become one of the biggest threats to human health. One popular biological pathway to target for development of antibacterial agents has been peptidoglycan (PG) biogenesis. Because the PG layer is uniquely present in bacteria and disruption of its biosynthesis causes cell lysis, designing inhibitor compounds that are specific to an enzyme in the pathway has been a promising route. The focus of this study has been on an integral membrane protein, MraY, and a peripheral protein, MurG, which together synthesize the lipid-linked PG building blocks at the cytoplasmic side of the membrane. The goal here is to obtain three-dimensional molecular pictures of MraY and MurG in the presence of their substrate(s) or inhibitors. More specifically, how the two proteins were purified and used in structural studies using X-ray crystallography will be discussed. The availability of structural details of these enzymes will provide mechanistic insights, as well as contribute to designing and development of selective and potent antimicrobial drugs.

INTRODUCTION

The rapidly growing number of antibiotic-resistant and multidrug-resistant bacterial pathogens pose a great threat to human health (Brown & Wright, 2016; WHO, 2017). A large number of attractive protein targets for antibiotics have been identified from bacterial pathogens. However, their structural and mechanistic details are often missing, slowing design of antibacterial agents that are selective and potent toward each target. Thus, the main goal of this study will be elucidating the catalytic mechanism of selected protein targets by providing its structural details to help the development of novel antibiotics.

Most bacteria possess the peptidoglycan (PG) layer as a major constituent of their cell wall, which protects the cells from the internal turgor pressure and helps them to maintain their cell shape (Lovering, Safadi & Strynadka, 2012). In addition, a functional peptidoglycan layer is required for effective cell division (Vollmer, Blanot & De Pedro, 2008). Gram-positive bacteria have a thick peptidoglycan layer outside their single cell membrane, whereas Gram-negative bacteria have a relatively thinner one in the periplasmic space between the two membranes (Silhavy, Kahne & Walker, 2010). Even the *Mycobacteriaceae*, which has its own unique cell envelope architecture, possess the peptidoglycan layer and its presence is critical in their cell viability (Jankute, Cox, Harrison & Besra, 2015). Because the peptidoglycan layer is unique to bacteria and there is no peptidoglycan in human, many enzymes involved in this pathway have been the site of action of antibacterial agents, including the clinically important β-lactam antibiotics (*e.g.*, penicillin) (Bugg, 1999).

The peptidoglycan biosynthesis pathway (**Fig. 3.1**) starts with UDP-*N*-acetylglucosamine (UDP-GlcNAc) in the cytoplasm, converting into uridine diphosphate-*N*-acetylmuramyl-pentapeptide (UDP-MurNAc-pentapeptide) by a series of enzymes MurA-F (Barreteau et al., 2008). UDP-MurNAc-pentapeptide then is the donor for the transfer of phospho-MurNAc-pentapeptide to a lipid carrier, undecaprenyl phosphate (C_{55}-P), to form undecaprenyl pyrophosphoryl MurNAc-pentapeptide or Lipid I at the cytoplasmic side of the membrane. This reaction is catalyzed by the key integral membrane protein, MraY (phosphor-MurNAc-pentapeptide translocase). Another glycosyltransferase, MurG, attaches a GlcNAc residue to Lipid I, producing Lipid II, which is the building block of the peptidoglycan layer (Bouhss, Trunkfield, Bugg & Mengin-Lecreulx, 2008; van Heijenoort, 2007). Subsequently, Lipid II flips across the membrane to the other side by an integral membrane protein, flippase (MurJ), and undergoes polymerization forming an alternating MurNAc and GlcNAc chain and cross-linking between pentapeptides (typically 3–4 peptide crosslink) to complete the peptidoglycan layer (Matteï, Neves & Dessen, 2010; Sauvage, Kerff, Terrak, Ayala & Charlier, 2008). Over time, a myriad of natural product inhibitors targeting this pathway have been identified and used in clinics, but development of antibacterial resistance toward them is problematic (Silver, 2013).

All of the enzymes in the peptidoglycan biosynthesis pathway have been structurally characterized in one species or another. However, mechanistic understanding of some of them lags behind largely due to the absence of substrate-bound structures. This is especially true for the enzymes in the pathway that have hydrophobic substrates (*e.g.*,

MraY, MurG, MurJ). It has been a challenge to determine their high-resolution structures with the lipid substrate bound. In addition, selective inhibition of specific pathogens requires species-specific structural and functional characterization of these enzymes. For example, in order to treat human tuberculosis, you would want to have an antibiotic that targets enzymes only in *Mycobacterium tuberculosis* rather than those from all bacterial species residing in human body, some of which are beneficial to human health.

Here, I will present the purification of MraY from the thermophiles *Hydrogenivirga sp.* (*Hy*) and *Mycobacterium thermoresistibile* (*Mth*) and the efforts toward determining a structure of the MraY protein bound to a substrate or novel inhibitors. Also, MurG from *Hydrogenivirga sp.* has been successfully purified and a subject of structural characterization using X-ray crystallography technique.

RESULTS

MraY from Hydrogenivirga sp. (HyMraY) was prepared in high purity and yield.

MraY belongs to the polyprenyl-phosphate N-acetyl hexosamine 1-phosphate transferase (PNPT) superfamily that includes other members like WecA. The first structure determined in this superfamily was a crystal structure of an apo form of MraY from the thermophile, *Aquifex aeolicus* (Chung et al., 2013). At the same time, our lab was also actively purifying and crystallizing MraY from *Hydrogenivirga sp.*, selected from an expression test of MraY from many species (data not shown here). The initial purification protocol of the His_6-MraY construct was handed over to me and I modified some of the steps (**Fig. 3.2**). I ended up acquiring protein that is purer and more homogeneous (**Fig. 3.2.C**) than before by introducing 20 mM imidazole wash steps in a nickel-affinity column, concentrating before dialysis of the sample for a cation-exchange column, and extending the first gradient step in the ion exchange column protocol to better separate contaminants. The yield of *Hy*MraY from 24-L culture varies in each batch, but approximately 2 mg is obtained on average. This amount of protein is enough to set up crystallization trays.

Purification of HyMraY without a His-tag.

Literature suggests a His-tag on either N- or C-terminus of a protein can perturb folding, solubility, crystal packing, and native conformation of the protein. However, there are 52 separate pairs of protein structures in the PDB where the protein structure was solved with or without the His-tag and analysis suggests no statistically significant differences (Carson, Johnson, McDonald, Brouillette & DeLucas, 2007). In order to test if the His-

tag affected crystallization, a His_6–3C-GS_5 -*Hy*MraY construct was generated that contained a protease site and a linker between the His-tag and the start of the protein. This was expressed, treated with an HRV 3C protease, and purified (**Fig. 3.3** & **3.4**). After complete digestion with the protease in 2–3 days, the sample was run through either an anion (Uno Q6) (**Fig. 3.3**) or cation (Uno S6) (**Fig. 3.4**) exchange column, each followed by a gel-filtration column. Most of the cleaved protein from both columns were pulled out of the purification columns with other contaminants. The purest protein without a

His-tag was obtained from a small peak that came off from the anion-exchange column at a low concentration of salt. There is still some room to optimize the purification protocol: (a) perform an expression test varying competent cells and expression conditions (*e.g.*, growth temperature, duration, amount of an inducer, amount of L-rhamnose for pLEMO-containing cells), (b) wash the protein-bound Ni-NTA resin with a larger volume of buffer containing low-concentration imidazole, (c) try a gradient of imidazole concentration to wash and elute the protein, (d) use a cobalt-affinity column because its binding is known to be more specific toward a His-tag than Ni-NTA, and (e) after trying all the above, run the sample through both cation and anion-exchange columns again and compare the chromatograms and SDS-PAGE gels.

Co-crystallized of HyMraY with the inhibitor UT-17460.

Our collaborator, Michio Kurosu and his group at the University of Tennessee Health Science Center has focused on development of novel natural product-based inhibitors of MraY. Recently, they developed a novel nucleoside analog, UT-17460, that inhibits the

outgrowth of *Clostridium difficile* spores by targeting MraY with an IC_{50} value of 0.08 ± 4.33 μM (Mitachi et al., 2018) (**Fig. 3.5**). A low IC_{50} value may indicate high binding affinity of the compound to *Hy*MraY, so purified protein was co-crystallized with UT-17460 via a standard vapor-diffusion method as well as lipidic cubic phase (LCP). For the sitting-drop vapor-diffusion method, initial screening of crystallization conditions was performed using commercially available sparse matrices, including MemGold, MemGold2, Morpheus, Index, and Crystal Screens. Previous efforts of crystallizing *Hy*MraY with substrates such as UDP-MurNAc-pentapeptide and/or C_{10}-P, a soluble variant of the native lipid substrate, were not successful. However, in the presence of UT-17460, crystal formation was robust throughout MemGold and Index in the sitting-drop vapor-diffusion method. Some initial crystal hits were mostly rod-shaped in different sizes and most of them were in clusters (**Fig. 3.6**). The crystals were confirmed to have UV fluorescence which suggests that, even though both the protein and the inhibitor will fluoresce, they are not salt and likely contain minimally *Hy*MraY.

To optimize the crystallization hits, several modifications were introduced. In general, initial hits were converted to a grid screen where the pH increases horizontally across a 24-well tray and precipitant concentration increases vertically across the tray. The design goal is to center the original condition in the grid screen and give a wide range of both buffer pH and precipitant concentration. At this point, typical optimization utilized the hanging-drop vapor-diffusion method with bigger drops (2 μL or more) for ease and typically it would lead to increases in the size and changes in the shape of the crystals. Also, with hanging drops crystals can be formed at the drop edge and can be harvested

more easily from precipitate. For example, for the condition that originally contained 0.1 M Tris pH 8.5, 0.1 M potassium chloride, and 39 % PEG 400 (MemGold D7, **Fig. 3.6.B**), a grid screen was prepared with the pH range 7.9–8.9 and PEG 400 concentration 35–43 %. Much bigger rod-shaped crystals were formed throughout the grid screen, but they were formed in clusters, which was not ideal for collecting X-ray data on (**Fig. 3.7.B**). In order to obtain single crystals, one condition was selected from the grid screen (0.1 M Tris pH 8.5, 0.1 M potassium chloride, and 35% PEG 400) and 96 different reagents from Additive screen (Hampton Research) were added. At the same time, the volume-to-volume ratio of the protein to reservoir condition in the drop was varied. The ratio change reduced the clustering of crystals, but some split crystals were still observed (**Fig. 3.7.C**).

In the meantime, over 200 crystals were flash frozen in liquid nitrogen, most of them without any additional cryoprotectants and shipped to the SSRL BL12-2. The crystals diffracted only out to approximately 9 Å in the best cases (**Fig. 3.8**). More efforts are needed to make single crystals.

Expression and purification of MraY from Mycobacterium thermoresistibile (MthMraY). Edwards et al. suggests Mth can be a useful *Mycobacterium* model organism to study *M. tuberculosis* (Mtb) because of following reasons: Mth (a) is thermostable, (b) has generally more soluble orthologs, (c) has a similar genome size as that of Mtb, and (d) can cause granuloma formation in the lung, which is a hallmark of Mtb infection. (Edwards, Liao, Phan, Myler & Grundner, 2012). The sequence alignment between

*Mth*MraY and *Mtb*MraY (also called MurX) resulted in 85% identity and 92% similarity (Altschul, Gish, Miller, Myers & Lipman, 1990). Together, *Mth*MraY is an attractive ortholog to study from the structural perspective. The *Mth*MraY construct initially used for small-scale expression test contained a His-tag, SUMO and a linker (His_6-SUMO-GS_5-*Mth*MraY (51.2 kDa). In Nico21(DE3) pLEMO cells, *Mth*MraY was significantly expressed at two different induction temperature and with three different concentrations of L-rhamnose added (**Fig. 3.9**). The highest levels of expression were when the cells were induced at 22°C in the presence of 0.4 mM L-rhamnose. Using these conditions, the *Mth*MraY expression was repeated in a larger scale, the His-tag was removed by cleavage with the Ulp1 protease, and further purified (**Fig. 3.10**). *Mth*MraY eluted from a nickel-affinity column with 110 mM imidazole, 150 mM imidazole, and 200 mM EDTA. Cleavage with Ulp1 was efficient and purification via a cation-exchange column and gel-filtration column went well. However, the yield obtained at the end was too low to be used for crystallization trials. The obtained protein was used for negative staining and imaged sample using a transmission electron microscope (TEM).

MthMraY aggregated in negative stain.

Purified *Mth*MraY (0.014 mg/ml) was applied on a carbon-coated 400-mesh carbon grid followed by 2% uranyl acetate. Two different areas of the grid were imaged using a Tecnai T12 (120 keVa TEM) at 26,500–67,000x magnification (**Fig. 3.11**). Some aggregation of particles was observed possibly due to the presence of excess decyl maltoside (DM) micelles.

Crystallization of MurG from Hydrogenivirga sp. (Hy) with UDP-GlcNAc.

A construct of MurG from *Hydrogenivirga sp. 128–5-R1–1* (*Hy*Mray) was generated with a His-tag followed by a thrombin cleavage site before the start of the protein. The construct, His_6-thrombin-*Hy*MurG, was expressed in *E. coli* cells and purified in two steps by a cobalt-affinity column and a gel-filtration column from either the cytosolic fraction or that obtained by extraction from the membrane pellet (**Fig. 3.12**). The *Hy*MurG extracted from the *E. coli* membrane pellet resulted in pure and homogeneous protein and confirms the expectation that *Hy*MurG is strongly associated with the membrane. Purified protein was incubated with UDP-GlcNAc, the soluble substrate, and screened for crystallization conditions. Four commercial screens were initially tried and among them, four conditions produced initial crystal hits (**Fig. 3.13**).

Grid screens with varying buffer pH and precipitant concentration were set up for all of the initial four conditions. One condition, MemGold F12, reproduced crystals with improved size and shape in hanging-drop vapor-diffusion method. Furthermore, microseeds were prepared from older crystal trays and used for setting up new trays with fresh protein/substrate solution. Several hundreds of rod-shaped crystals were shipped to the SSRL BL12-2. Four data sets that are worth reporting here have resolution cut-off of 3.28 Å, 2.81 Å, 2.78 Å, and 2.60 Å based on thresholds I followed with overall completeness higher than 99 % (over 70 % in the highest resolution shell), intensity (I/σ) close to or bigger than 1, and correlation coefficient value from random half-datasets ($CC_{1/2}$) over 40 %. The first data set cut off at 3.28 Å was solved using a molecular replacement with a search model of MurG from *Pseudomonas aeruginosa* (PDB ID:

3S2U), followed by three rounds of refinement, resulting in R_{work} = 0.3837 and R_{free} = 0.4494. Although seeing a positive F_o-F_c map in the region UDP-GlcNAc is likely to fit based on the secondary structure alignment with *Pa*MurG bound to UDP-GlcNAc is exciting, there were other problems including high clash score and high number of outliers of the Ramachandran plot. Based on the same criteria mentioned above, the next three data sets were cut off at resolution of 2.81 Å (**Fig. 3.14.AB**), 2.78 Å (**Fig. 3.14.CD**) and 2.60 Å (**Fig. 3.14.EF**). Phases of the 2.60 Å data set were solved by the same search model mentioned above and currently model building and refinement are in progress (**Fig. 3.15**).

DISCUSSION

For peptidoglycan biosynthesis, one of the remaining questions that are challenging to address is how undecaprenyl phosphate (C_{55}-P) is incorporated into the active site of MraY at the cytoplasmic side and how it interacts with UDP-MurNAc-pentapeptide to produce Lipid I. In addition, how Lipid I, that is hydrophobic in nature, is transferred to the MurG active site and transfers a GlcNAc to produce Lipid II. In order to understand the occupancy of the lipid substrates and products in the active sites, structure determination of the enzymes in complex with both of the substrates is required.

Structures of MraY in complex with any of the substrates have not been determined. In collaboration with the Kurosu laboratory, MraY from the thermophile, *Hydrogenivirga sp*., was co-crystallized with a substrate analog (**Fig. 3.16.A**). However, the crystals obtained had low diffraction quality (data not shown here). One of the inhibitor compounds the Kurosu laboratory developed, UT-17460 (**Fig. 3.5**), was also used to co-crystallize with *Hy*MraY. UT-17460 is expected to bind to MraY in a similar way as UDP-MurNAc-pentapeptide does via the uridine. The presence of UT-17460 enhanced crystallization of *Hy*MraY, but the highest resolution obtained the diffraction data was 8 Å. Previously, solubility of UT-17460 in solution was low, but it was improved by preparing it in a salt form by the Kurosu laboratory and renamed as aminouridyl phenoxypiperidinbenzyl butanamide (APPB) (**Fig. 3.16.B**). Co-crystallization with APPB produced similar crystals as before, but the diffraction quality stayed the same.

Unlike MraY, crystal structures of MurG from *E. coli* and *Pseudomonas aeruginosa* in complex with its soluble substrate, UDP-GlcNAc, are available. Lipid I- or Lipid II-bound structures of MurG, however, has not been determined. MurG from *Hyrogenivirga sp*. was expressed and purified in high yield and homogeneity from the *E. coli* membrane pellet. Co-crystallization and high-resolution diffraction data collection of *Hy*MurG with UDP-GlcNAc has been successful. The next goal is co-crystallizing MurG with a soluble analog, C_{10}-P, or MraY with its substrates.

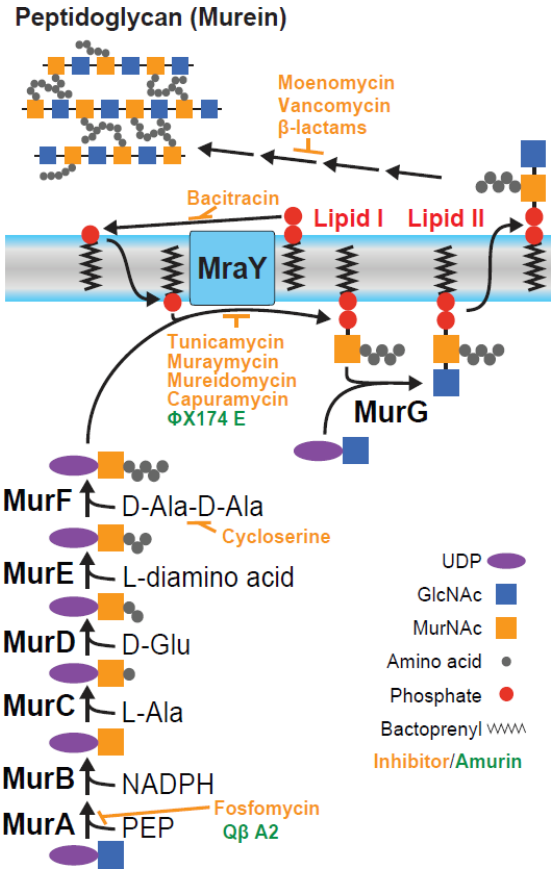

Figure 3.1. Summary of the peptidoglycan biosynthesis pathway. Peptidoglycan biogenesis starts with uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc), which is converted to UDP-*N*-acetylmuramyl pentapeptide (UDP-MurNAc-pentapeptide) by a series of enzymes, MurA-F. MraY catalyzes the transfer of phospho-MurNAc-pentapeptide to undecaprenyl phosphate (C_{55}-P), forming Lipid I. MurG adds a GlcNAc to produce Lipid II, which is then flipped to the periplasmic side and undergoes polymerization and cross-linking to synthesize the peptidoglycan layer. Inhibitors and phage proteins that target specific enzymes in the pathway are shown in orange and green, respectively. PEP, phosphoenolpyruvate; NADPH, nicotinamide adenine dinucleotide phosphate. **Created by Prof. Bil Clemons.**

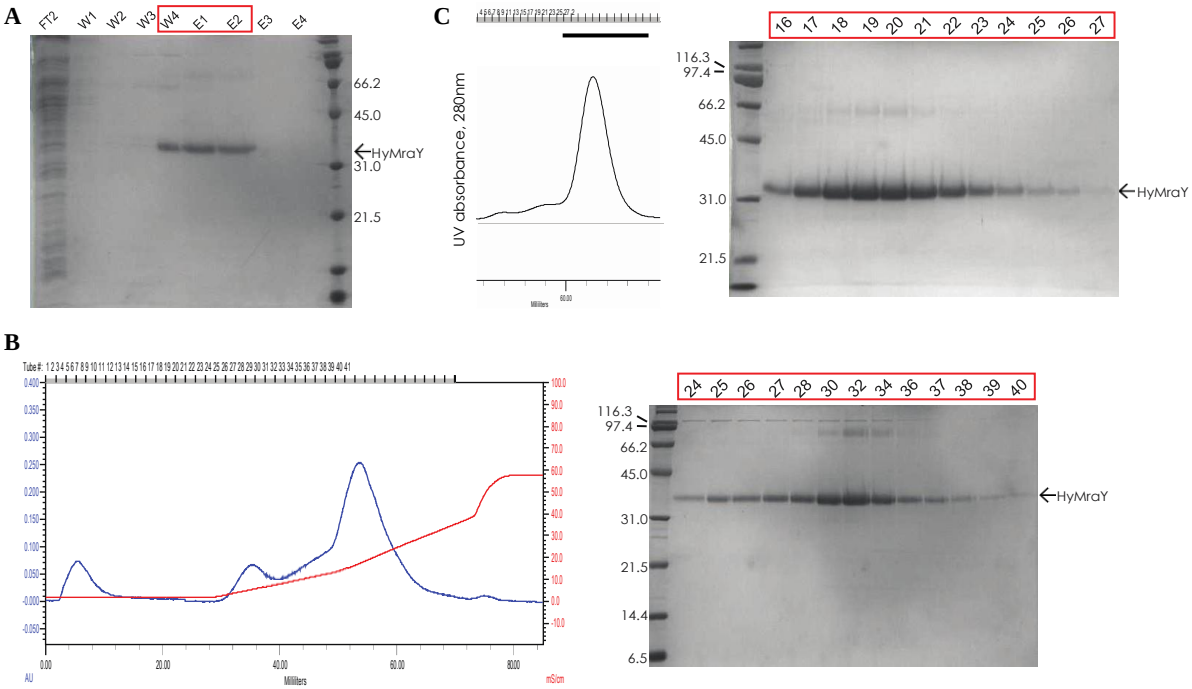

Figure 3.2. Optimized purification of *Hy*MraY. A, an SDS-PAGE gel from a nickel-affinity column run with *Hy*MraY. *Hy*MraY came off in W4, E1, and E2 fractions. FT2, second flow-through; W1–2, wash with 20 mM imidazole and 10 mL each; W3–4, wash with 20 mM EDTA and 10 mL each; E1–4, elute with 20 mM EDTA and 5 mL each; the last lane is Broad Range protein marker in kDa from Bio-Rad; B, purification of *Hy*MraY using a cation-exchange column. Left, a cation-exchange (Uno S6) chromatogram shows *Hy*MraY (peak at ~55 mL) is separated from other contaminants based on charge difference. Right, fractions 24–40 were run on an SDS-PAGE gel and *Hy*MraY is almost pure; C, purification of *Hy*MraY using a gel-filtration column. Left, a gel-filtration (Superdex 200 16/600) chromatogram shows a symmetrical peak of *Hy*MraY. Right, fractions 16–27 were run on an SDS-PAGE gel and *Hy*MraY is pure.

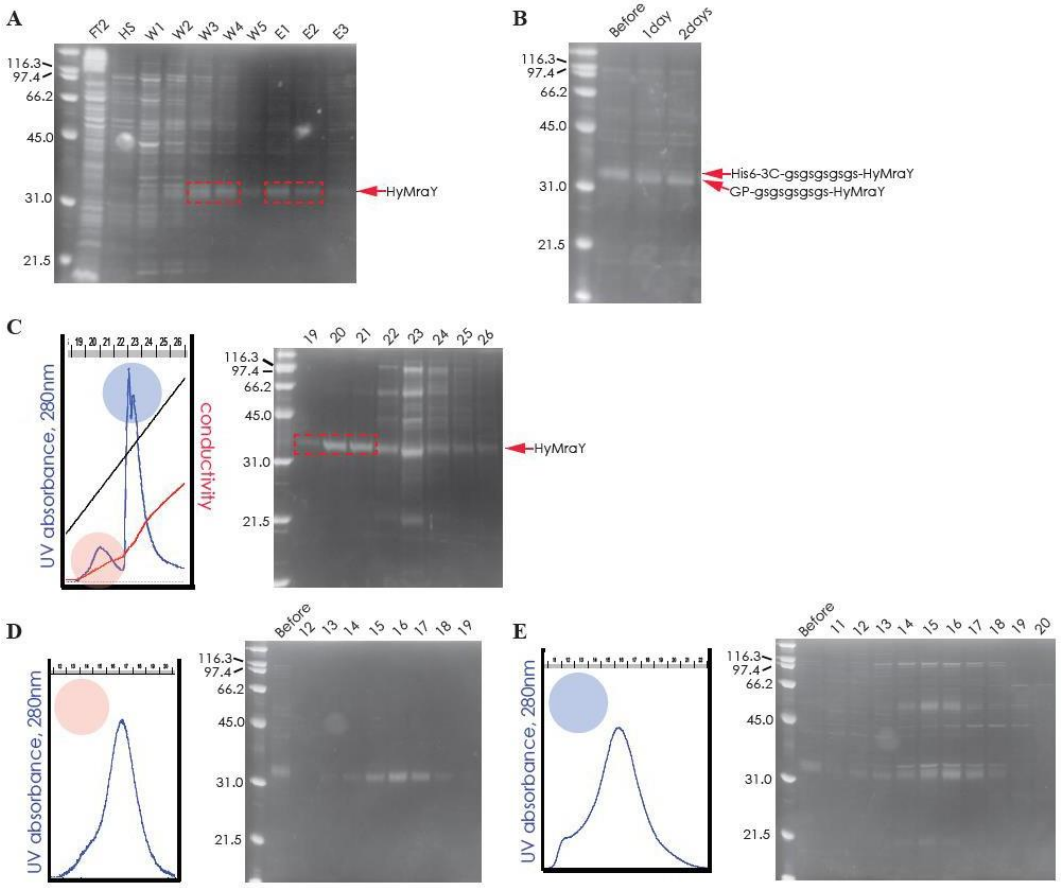

Figure 3.3. Cleavage of the His-tag from *Hy*MraY and purification via an anion-exchange column. A, an SDS-PAGE gel from a nickel-affinity column run with *Hy*MraY. *Hy*MraY came off in W3-W4 and E1-E2 fractions. FT2, second flow-through; HS, high-salt wash with 1 M NaCl and 5 mM imidazole and 50 mL; W1–5, wash with 20 mM EDTA and 10 mL each; E1–3, elute with 200 mM EDTA and 5 mL each; the first lane is Broad Range protein marker in kDa from Bio-Rad; B, a SDS-PAGE gel after 3C proteolysis. Before, before adding HRV 3C protease to the protein sample; 1 day, cleavage for 1 day; 2 days, cleavage for 2 days. C, purification of *Hy*MraY without the His-tag using an anion- exchange column. Left, an anion-exchange (Uno Q6) chromatogram shows *Hy*MraY came off pure in the first smaller peak. Right, fractions 19–26 were run on an SDS-PAGE gel and *Hy*MraY is pure in fractions 19–21; D, purification of *Hy*MraY without the His-tag using a gel-filtration column. Left, a gel-filtration (Superdex 200 10/300) chromatogram shows an almost symmetrical peak of the protein. The red circle indicates the fractions from the first peak in C. Right, fractions 12–19 were run on an SDS-PAGE gel and *Hy*MraY without the His-tag is pure. Before, the sample before adding HRV 3C protease was run to confirm cleavage again. E, Left, a gel-filtration (Superdex 200 10/300) chromatogram shows a peak with a left shoulder. Right, fractions 11–20 were run on an SDS-PAGE gel and *Hy*MraY without the His-tag is pure. The blue circle indicates the fractions from the second peak in C.

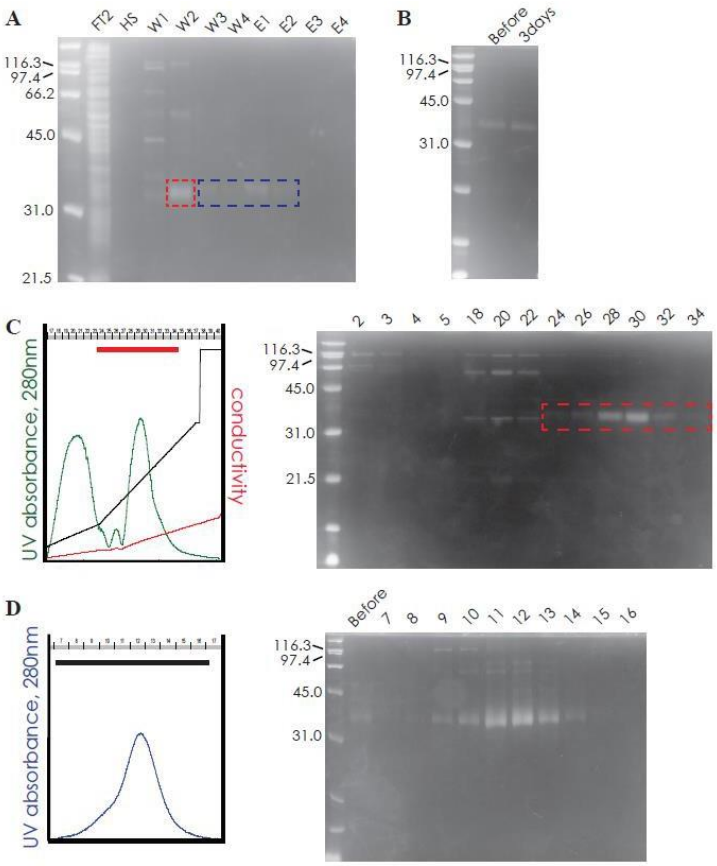

Figure 3.4. Cleavage of the His-tag from *Hy*MraY and purification via a cation-exchange column. A, an SDS-PAGE gel from a nickel-affinity column run with *Hy*MraY. *Hy*MraY came off mostly in W2. FT2, second flow-through; HS, high-salt wash with 1 M NaCl and 5 mM imidazole and 50 mL; W1–4, wash with 20 mM EDTA and 10 mL each; E1–4, elute with 200 mM EDTA and 5 mL each; the first lane is Broad Range protein marker in kDa from Bio-Rad; B, a SDS-PAGE gel after 3C proteolysis. Before, before adding HRV 3C protease to the protein sample; 3 days, cleavage for 3 days. C, purification of *Hy*MraY without the His-tag using a cation-exchange column. A cation-exchange (Uno S6) chromatogram shows *Hy*MraY came off fairly pure in fractions 24–34. D, *Hy*MraY without the His-tag using a gel- filtration column. Left, a gel-filtration (Superdex 200 10/300) chromatogram shows an almost symmetrical peak.
Before, the sample before adding HRV 3C protease was run to confirm cleavage again. Fraction 7–16 was run on an SDS-PAGE gel.

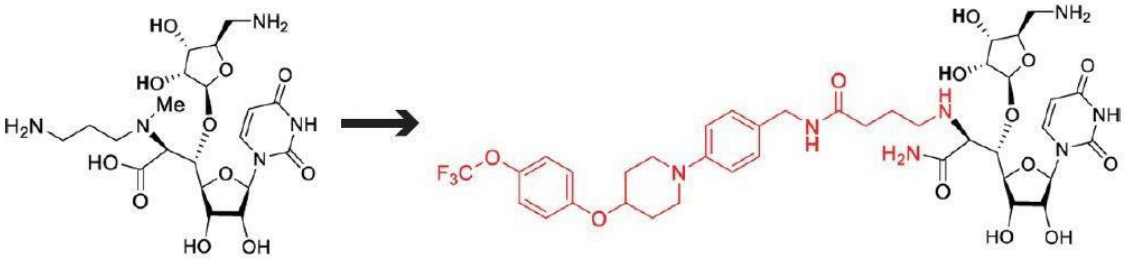

Figure 3.5. The development of UT-17460. The existing antibiotic, FR-900493 from *Bacillus cereus* (left), was chemically modified and became UT-17460 (right) that has improved inhibitory activity against *Hy*MraY (IC_{50} = 0.08 ± 4.33 μM). This figure is adapted from the published work in collaboration with the Kurosu lab (Mitachi et al., 2018).

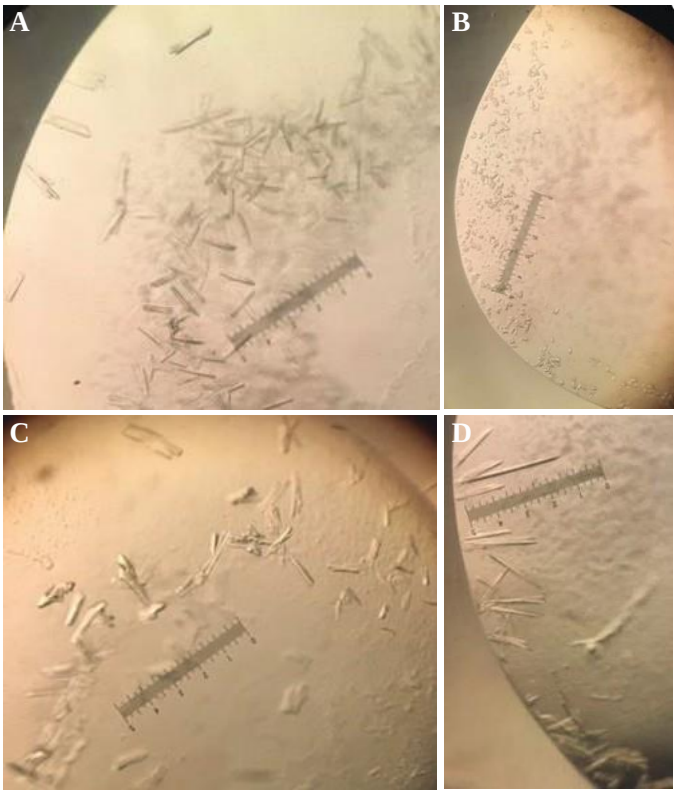

Figure 3.6. Initial crystal hits of *Hy*MraY with UT-17460 from commercial screens. A, 0.2 M sodium acetate trihydrate, 0.2 M potassium chloride, 0.1 M HEPES pH 7.0, 22% PEG 3000; B, 0.1 M Tris pH 8.5, 0.1 M potassium chloride, 39% PEG 400; C, 0.2 M sodium chloride, 0.1 M Bis-Tris pH 5.5, 25% PEG 3350; D, 0.05 M ammonium sulfate, 0.05 M Bis-tris pH 6.5, 30% Pentaerythritol ethoxylate (15/4 EO/OH). A&B from MemGold (Molecular Dimensions). C&D from Index (Hampton Research).Each tick of the ruler in the images is 4 μm long.

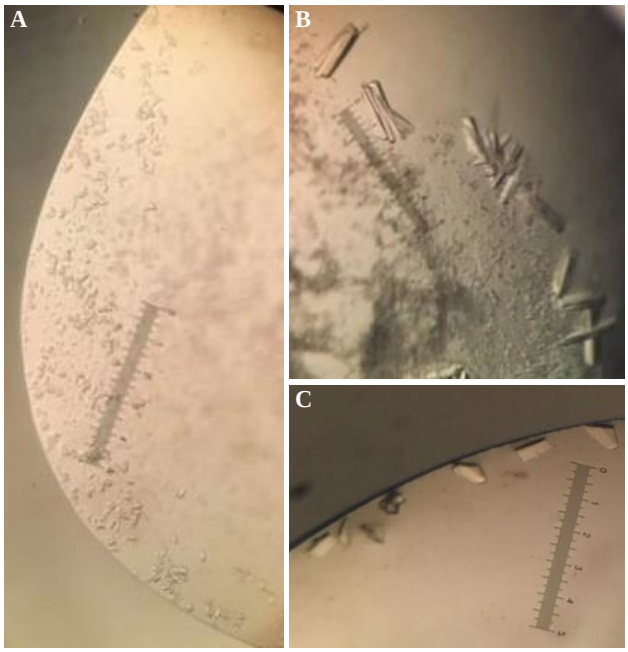

Figure 3.7. Optimization of a crystal condition of *Hy*MraY with UT-17460. A, an initial crystal hit from MemGold in 0.1 M Tris pH 8.5, 0.1 M potassium chloride, and 39% PEG 400 via the sitting-drop vapor-diffusion method with a 0.4 μL drops. B, pH was optimized to 8.9 in the hanging-drop vapor-diffusion method with 2 μL drops. C, the PEG 400 concentration was optimized to 35% in the hanging-drop vapor-diffusion method with 2 μL of the sample and 1 μL of the well solution in each drop. An additive, 0.03 M glycyl-glycyl-glycine was also added to the well and the drop.

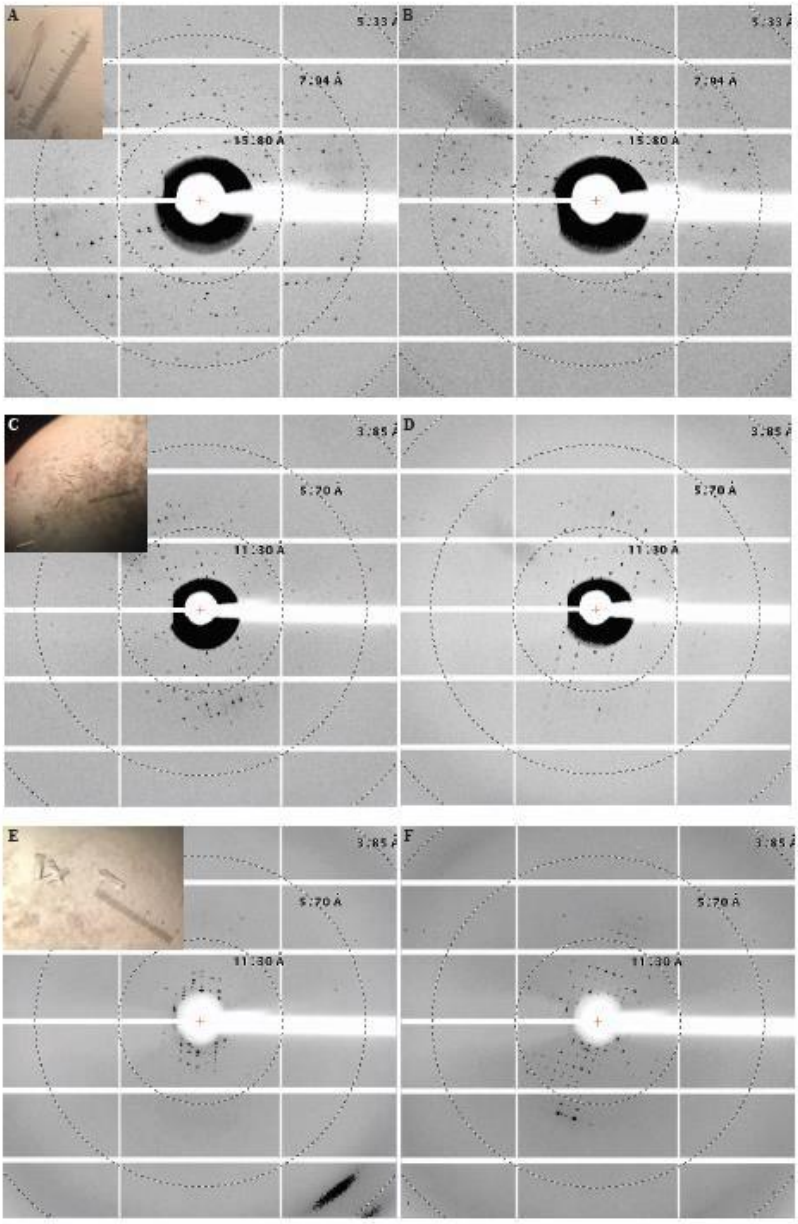

Figure 3.8. X-ray diffraction screening of *Hy*MraY co-crystallized with UT-17460 from the grid screen at the SSRL BL12-2. A&B, The final concentration of the protein and UT-17460 in the drop was 5 mg/ml and 500 µM, respectively. The crystallization condition was 0.1 M Tris pH 8.7, 0.1 M potassium chloride, and 43% PEG 400. Note, X-ray screening on Nov. 8, 2018; the cassette number was 333; the port number was E5. C&D, The final concentration of the protein and UT-17460 in the drop was 2.6 mg/ml and 500 µM, respectively. The crystallization condition was 0.1 M Tris pH 8.3, 0.1 M potassium chloride, and 43% PEG 400. Note, X-ray screening on Nov. 19, 2018; the cassette number was 210; the port number was E6. E&F, The final concentration of the protein and UT-17460 in the drop was 2.6 mg/ml and 500 µM, respectively. The crystallization condition was 0.1 M Tris pH 8.3, 0.1 M potassium chloride, and 39% PEG 400. Note, X-ray screening on Nov. 19, 2018; the cassette number was 210; the port number was D6. The crystals harvested are shown in A, C, and E at the top left corner. A and B/ C and D/ E and F are pairs of diffraction images that were acquired from exposing the crystals 90° apart. The resolution rings were drawn in using Adxv.

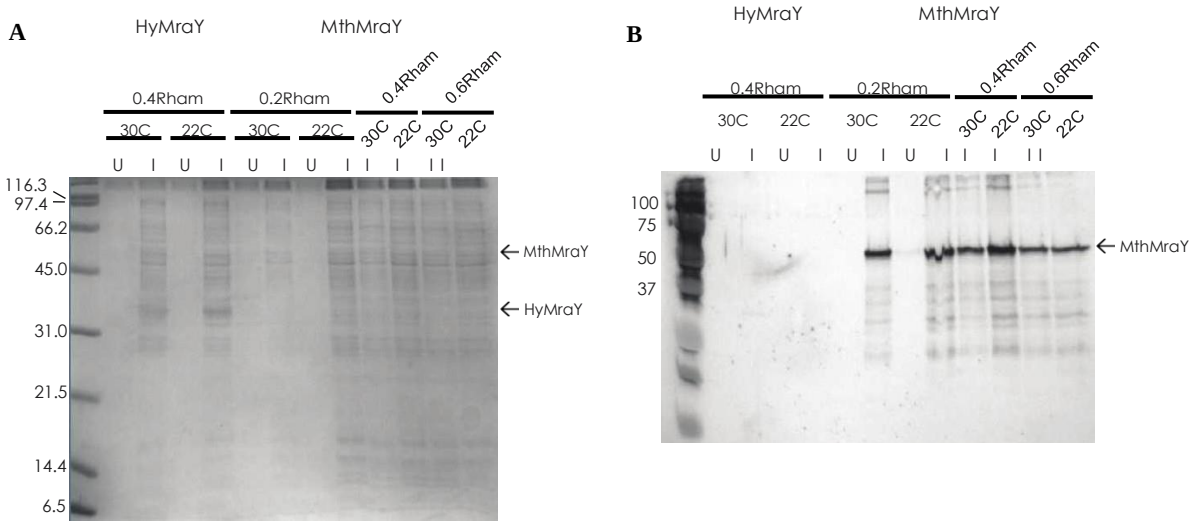

Figure 3.9. Expression test of *Mth*MraY. A, the His_6-SUMO-GS_5-*Mth*MraY construct was expressed in Nico21(DE3) pLEMO cells in the presence of 0.2, 0.4, or 0.6 mM L-rhamnose and they were induced at either 30 °C or 22 °C. The His_6-*Hy*MraY was expressed as a control. U, uninduced; I, induced. A, an SDS-PAGE gel shows both *Hy*MraY and *Mth*MraY were expressed in all conditions. B, a Western blot against α-His_5 antibody shows *Mth*MraY is clearly expressed in all of the induced fractions. The reason MraY does not appear in the Western blot is due to inaccessibility of the His-tag in this construct.

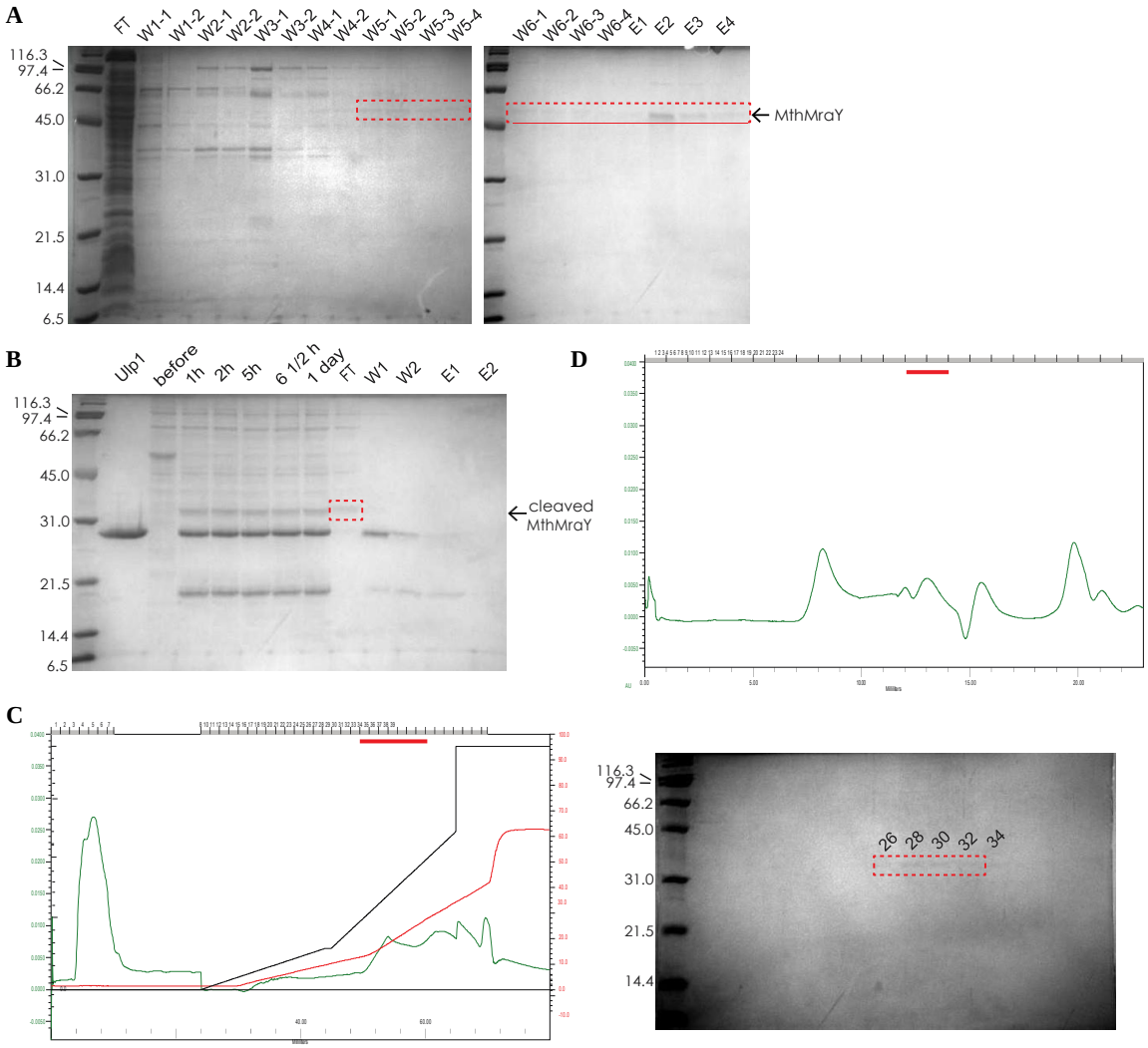

Figure 3.10. Cleavage of the His-tag and purification of *Mth*MraY. A, an SDS-PAGE gel from a nickel-affinity column run with *Mth*MraY. *Mth*MraY came off in W4–6 and E1–4 fractions. FT, flow-through; W1, wash with 20 mM imidazole and 10 mL each; W2, wash with 30 mM imidazole and 10 mL each; W3, wash with 50 mM imidazole and 10 mL each; W4, wash with 80 mM imidazole and 10 mL each; W5, wash with 110 mM imidazole and 10 mL each; W6, wash with 150 mM imidazole and 10 mL each; E1–4, elute with 200 mM EDTA and 5 mL each; the first lane is Broad Range protein marker in kDa from Bio-Rad; B, cleavage of the His-tag off *Mth*MraY using Ulp1 and a reverse nickel-affinity column. C, Left, a cation-exchange (Uno S6) chromatogram shows multiple peaks. Right, fractions 26–34 run on an SDS-PAGE gel confirm *Mth*MraY was present; D, purification of *Mth*MraY using a gel-filtration (Superdex 200 10/300) chromatogram shows multiple peaks. An SDS-PAGE gel was run and faint bands at the right size of *Mth*MraY appeared in fraction 13–14 (not shown here).

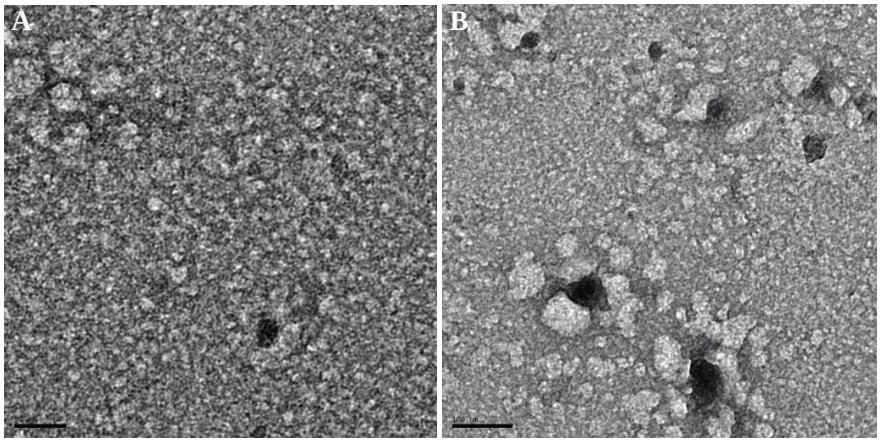

Figure 3.11. Micrographs of negative-stained *Mth*MraY on a carbon-coated copper grid. *Mth*MraY purified in DM on a carbon-coated 400-mesh copper grid is stained with 2% uranyl acetate. A, the scale bar is 50 nm long; B, the scale bar is 100 nm long. A Tecnai T12 was used to image the grid.

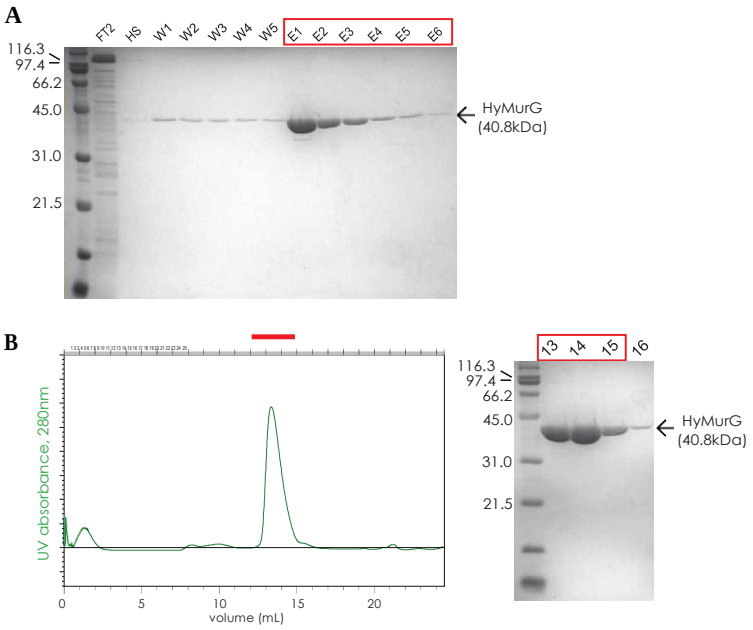

Figure 3.12. Purification of MurG from *Hydrogenivirga sp.* A, An SDS-PAGE gel from a cobalt-affinity column. Most *Hy*MurG came off in E1–6 fractions. FT2, second flow-through; HS, high-salt wash, wash with 1 M NaCl and 50 mL; W1–5, wash with 30 mM imidazole and 10 mL each; E1–6, elute with 200 mM imidazole and 5 mL each; the first lane is Broad Range protein marker in kDa from Bio-Rad; B, purification of *Hy*MurG using a gel-filtration column. Left, a gel-filtration (Superdex 200 10/300) chromatogram shows a symmetrical peak of *Hy*MurG. Right, fractions 13–16 were run on an SDS-PAGE gel and *Hy*MurG was pure.

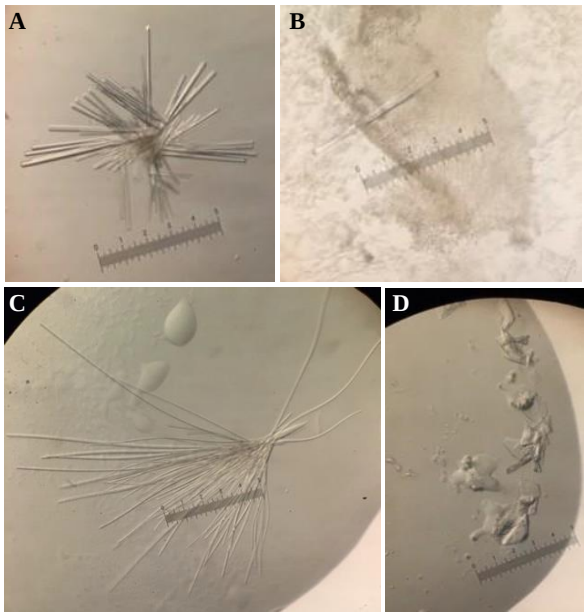

Figure 3.13. Initial crystal hits of *Hy*MurG co-crystallized with UDP-GlcNAc. A, 0.2 M sodium malonate, 20 % PEG 3350 pH 7.4 (Index H3); B, 1.2 M sodium citrate tribasic dehydrate, 0.01 M Tris pH 8.0 (MemGold A2); C, 0.1 M ADA pH 7.0, 31 % PEG 600 (MemGold2 H3); D, 0.07 M sodium chloride, 0.05 M sodium citrate pH 4.5, 22 % PEG 400 (MemGold F12).

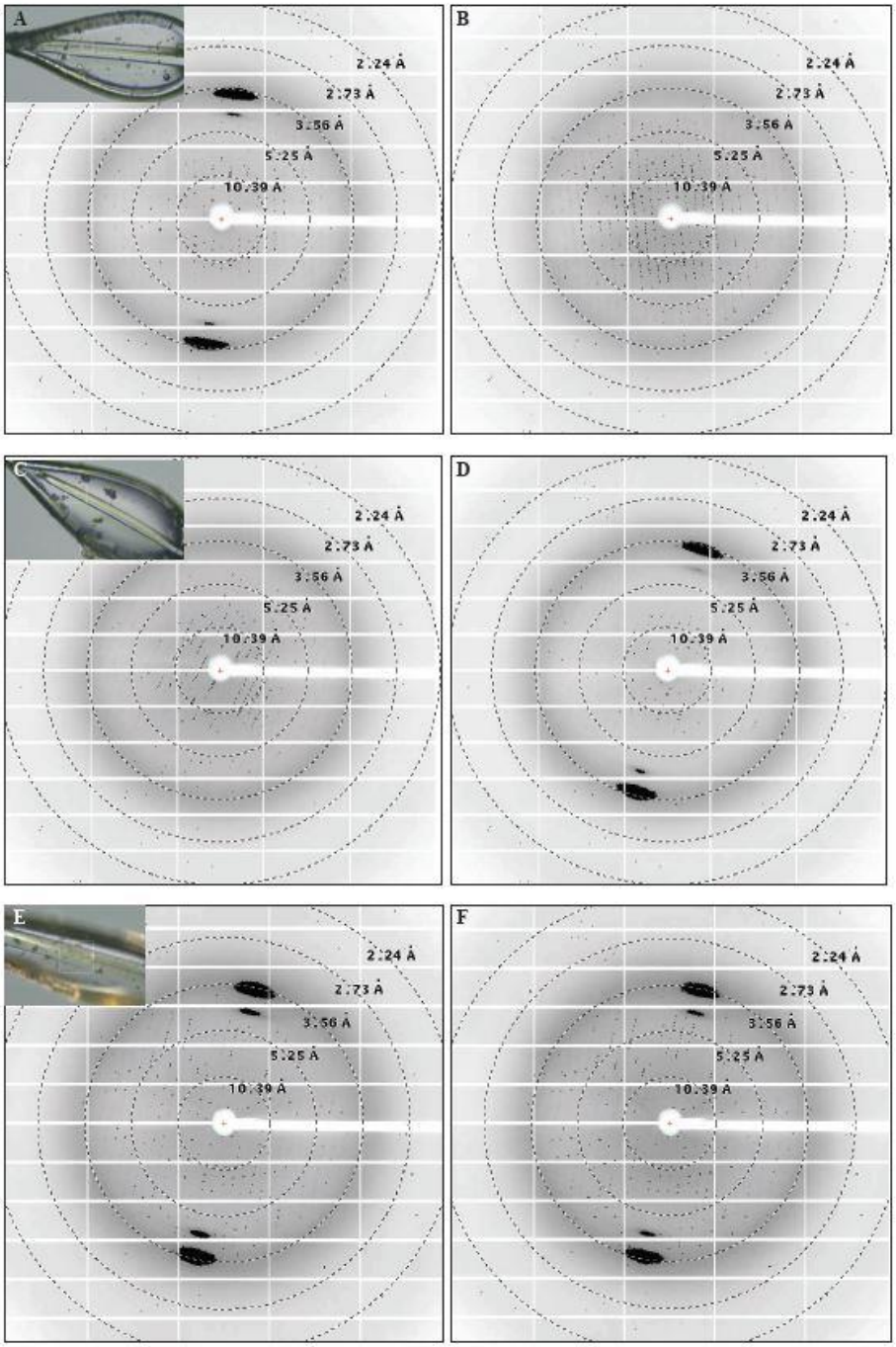

Figure 3.14. 2.81 Å, 2.78 Å, and 2.60 Å X-ray diffraction images of *Hy*MurG co-crystallized with UDP-GlcNAc at the SSRL BL12-2. The final concentration of the protein and UDP-GlcNAc in the drop was 5.5 mg/ml and 5.0 mM, respectively. The crystallization condition for A-D was 0.05 M sodium citrate pH 5.3, 0.07 M sodium chloride, and 22 % PEG 400. The crystallization condition for E and F was 0.05 M sodium citrate pH 5.5, 0.07 M sodium chloride, and 21 % PEG 400. The crystals that were harvested without additional cryoprotectant and exposed to X-ray are shown in the top left corner of A, C, and E. A and B/ C and D/ E and F are pairs of diffraction images that were acquired from exposing the crystal 90° apart. The data set collected from A and B diffracted to 2.81 Å, that from C and D to 2.78 Å, and that from E and F to 2.60 Å. The resolution rings were drawn in using Adxv. Note. X-ray screening for A-D on March 26, 2019; the cassette number was 379; the port number was A2 and A3. X-ray screening for E-F on May 3, 2019; the cassette number was 203; the port number was B7.

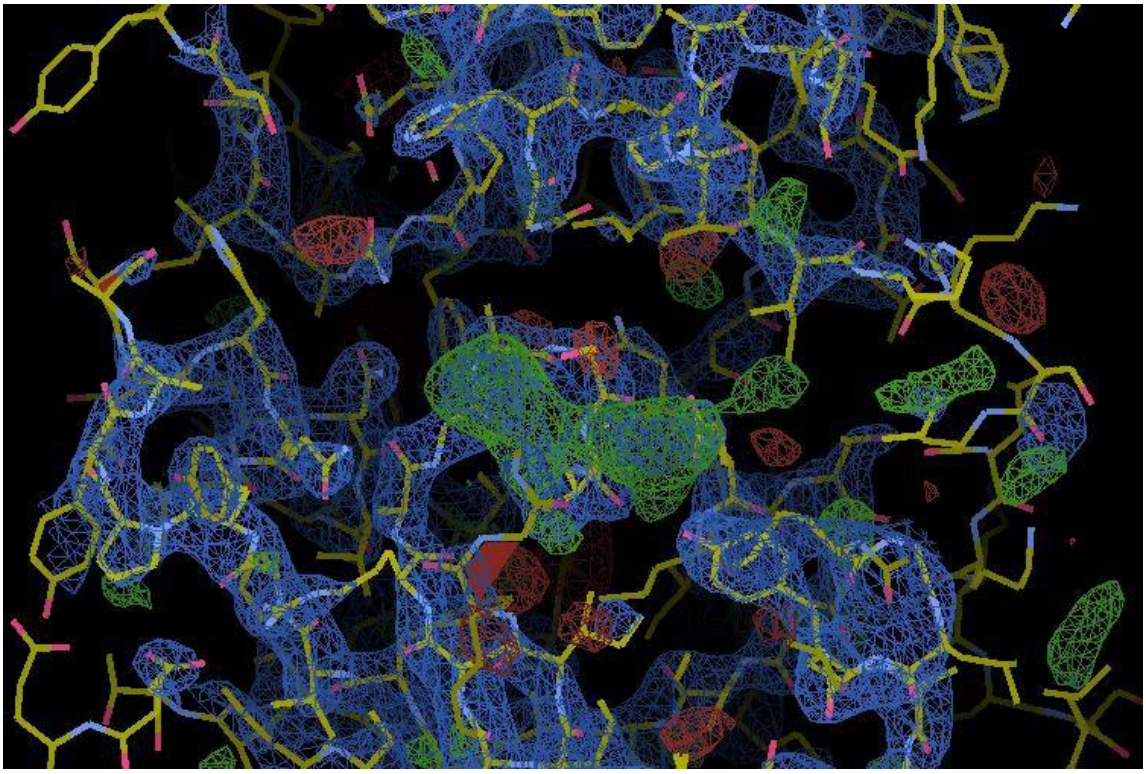

Figure 3.15. Partially refined 2.60 Å X-ray diffraction data of *Hy*MurG reveals some electron density for UDP-GlcNAc. The search model used to solve phases of the diffraction data was MurG from *Pseudomonas aeruginosa* bound to UDP-GlcNAc (PDB ID: 3S2U). Refined *Hy*MurG structure reveals a positive F_o-F_c map (green) in the center of the image, which is likely to be where UDP-GlcNAc binds.

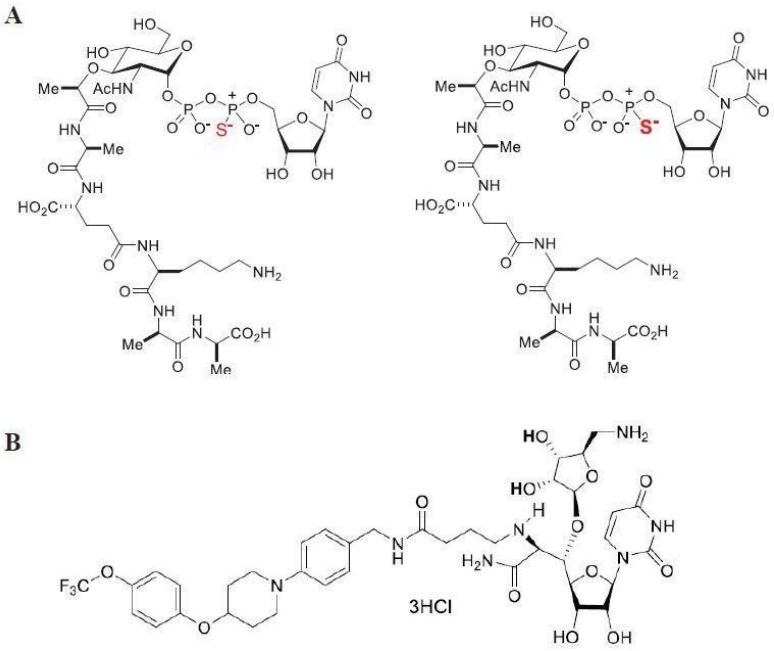

Figure 3.16. Chemical structures of the S-analog of UDP-MurNAc-pentapeptide in a diastereomer mixture and APPB-HCl salt. A, The Kurosu laboratory synthesized the S-analog and purified using a HPLC HYPERSIL GOLD™ column with the solvent ratio of MeCN : 0.05 M NH_4HCO_3 (aq) = 5 : 95 and 2.0 mL/min flow rate. The product was detected at 254 nm. B, UT-17460 was renamed with aminouridyl phenoxypiperidinbenzyl butanamide (APPB) and solubility was improved.

Data collection	
Synchrotron	SSRL
Beamline	BL12-2
Space group	$P6_1\ 2\ 2$
Cell dimensions	
a/b/c (Å)	197.65/197.65/52.74
α/β/γ (°)	90/90/120
Wavelength (Å)	1.0000
Resolution (Å)	39.27 – 2.60
R_{merge} (%)	11.6 (420.1)
R_{meas} (%)	11.7 (425.4)
R_{pim} (%)	1.9 (66.3)
$CC_{1/2}$	100 (65.4)
$\langle I/\sigma(I) \rangle$	23.9 (1.7)
Completeness (%)	100 (100)
No. of observations	753478 (92092)
No. of unique	19219 (2283)
Multiplicity	39.2 (40.3)
Average mosaicity	0.10

Table 3.1. Statistics of X-ray data collection of *Hy*MurG co-crystallized with UDP-GlcNAc at 2.60 Å. Values in parenthesis are from the highest resolution shell.

MATERIALS AND METHODS

Expression and purification of HyMraY

The pET22b-His$_6$-*Hy*MraY plasmid was transformed into Nico21(DE3) pLEMO cells the night before expression. In the morning of the expression day, all the colonies from a plate were scraped off and added to a small culture flask with 200 mL of LB + 35 µg/ml carbenicillin for a few hours. Then, 1 L 2x yeast tryptone (YT) (16 g/L tryptone, 10 g/L yeast extract, and 5 g/L NaCl) media and 35 µg/ml carbenicillin was inoculated with approximately 10 mL of the starter culture in the presence of 0.4 mM L- rhamnose. The cells grew at 37°C and when optical density at 600 nm (OD_{600}) reaches 0.4, the temperature was decreased to 30°C. When OD_{600} becomes 0.6–0.7, the cells were induced with 0.4 mM isopropyl-β-1-thiogalactopyranoside (IPTG) and grown for four more hours. The culture was harvested using a JLA-8.1 rotor at 4,000 rpm for 15 min.

Cell pellets were resuspended and homogenized into 20 mM Tris pH 7.5, 300 mM NaCl, 10% glycerol, 5 mM βME, and protease inhibitors (*e.g.,* PMSF, benzamidine). The cell resuspension underwent four passes in a microfluidizer. The cell lysate was spun in a JLA-16.250 rotor at 12,000 rpm for 30 min to pellet unbroken cells and cell debris. Subsequently, the supernatant was spun in an ultracentrifuge Ti-45 rotor at 45,000 rpm for 30 min and membrane pellet was scraped off from the bottom of the tubes. The membrane pellet was stored at -80°C or went forward and was resuspended in 20 mM HEPES pH 7.5, 300 mM NaCl, 10 % glycerol, 5 mM βME, and protease inhibitors along with 1 % DM and 5 mM imidazole. Extraction was achieved for two hours, rocking at

4°C. The sample was spun in an ultracentrifuge Ti-50.2 rotor at 45,000 rpm for 30 min, from which the supernatant was saved and incubated with Ni-NTA resin (1 mL resin was used for the sample from 6 L culture.) overnight, rocking at 4°C.

The sample bound to the Ni-NTA resin was first washed with 20 CV of 20 mM HEPES pH 7.5, 300 mM NaCl, 10 % glycerol, 5 mM βME, 0.15 % DM, and 20 mM imidazole, which was followed by another 20 CV wash with the same buffer, except for 20 mM EDTA instead of 20 mM imidazole. Then, proteins were eluted in 5 mL fractions of 20 CV with 20 mM HEPES pH 7.5, 100 mM NaCl, 10 % glycerol, 5 mM βME, 0.15 % DM, and 200 mM EDTA. All of the fractions from the nickel-affinity column were run in an SDS-PAGE gel. The fractions with proteins were concentrated using an Amicon 50 kDa cut-off concentrator (Millipore) until the volume reached about 5 mL. Then, the sample was added to a 10 kDa cut-off SnakeSkin^{TM} dialysis tubing (Thermo Scientific) and dialyzed in 1 L of 20 mM HEPES pH 7.5, 10 mM NaCl, 10 % glycerol, 5 mM βME, and 0.15 % DM overnight at 4°C.

On the next day, the dialyzed sample was filtered and injected into a cation-exchange column (Uno S6 from Bio-Rad) connected to a FPLC (BioLogic DuoFlow v5.3 from Bio-Rad). Proteins were released from the column during a gradient from 100 % Buffer A (same as the dialysis buffer) to 65 % Buffer B (same as Buffer A except for 1 M NaCl). Fractions in the peak area were run on an SDS-PAGE gel. The fractions with pure protein was concentrated using a concentrator to be injected onto a pre-equilibrated gel-filtration

column (Superdex 200 16/600 (GE Healthcare)) in a buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 10 % glycerol, 0.15 % DM, 5 mM βME, and 10 mM $MgCl_2$ at 0.7 mL/min. The protein eluted at approximately 65 mL and fractions were analyzed by SDS-PAGE and Coomassie staining. The final step was pooled fractions that were concentrated and either used immediately for structural and biochemical studies or flash frozen in liquid N_2 and stored at -80°C.

Purification of HyMraY without a His-tag

The construct His-3C-*Hy*MraY (pET22b-His_6-3C (Leu-Glu-Val-Leu-Phe-Gln ↓ Gly-Pro)-GS_5-*Hy*MraY (42.6 kDa)) was expressed as described above for *Hy*MraY. Ni-NTA resin (1 mL resin for 6 L culture) was incubated with the extracted sample overnight at 4°C, rocking. The bound sample was washed with 50 CV of 20 mM HEPES pH 7.5, 1 M NaCl, 10 % glycerol, 0.15 % DM, 5 mM βME, 5 mM imidazole, and protease inhibitors, followed by 50 CV of the same buffer, except for 300 mM NaCl, and 20 mM EDTA. Then, the proteins were eluted in three fractions of 5 mL each with the buffer containing 150 mM NaCl and 200 mM EDTA. W3–4 and E1–2 were collected and dialyzed into 50 mM Tris pH 7.5, 150 mM NaCl, 10 % glycerol, 5 mM βME, and 0.15 % DM at 4°C for one day. HRV 3C protease (2 units/L; ThermoFisher Scientific) was added to the dialyzed sample and rocked at 4°C for about two days with the ratio of the protease (60 µL) : protein (9 mg) = 1: 150. A reverse nickel-affinity column was run after incubating the digested sample with Ni-NTA resin at 4°C, overnight. The flow-through was taken from the reverse nickel-affinity column and dialyzed to run on an anion-exchange column

(Uno Q6) with Buffer A (20 mM HEPES pH 7.5, 10 % glycerol, 5 mM βME, and 0.15 % DM) and Buffer B, containing1 M NaCl. Lastly, selected fractions from the anion-exchange column were run on a gel-filtration column (Superdex 200 10/300) with the buffer containing 150 mM NaCl and 10 mM $MgCl_2$.

For subsequent expression, the same plasmid was expressed again and its His-tag was cleaved using the ratio of HRV 3C protease (28 μL) : protein (2.8 mg) = 100 :1. Then, the sample was purified in the same process as above, except for using a cation (Uno S6) instead of anion-exchange column. All the buffer compositions stayed the same.

Co-crystallization of HyMraY with UT-17460 using vapor-diffusion method and lipidic cubic phase (LCP)

For initial crystallization screening, 1 mM UT-17460 was added to 5 mg/ml purified *Hy*MraY. Aggregates were removed from the mixture on Amicon 0.22 μm filter by spinning at 8,000 rpm for 5–15 min at 4°C. For the standard sitting-drop vapor-diffusion method, the sample (0.2 μL) was dispensed into individual wells on Swissci 96-well MRC plates (Molecular Dimensions), followed by well solution (0.2 μL), which was drawn from the 50 μL reservoir, using Mosquito (TTP Labtech). Some commercial sparse matrices used were MemGold (Molecular Dimensions), MemGold2 (Molecular Dimensions), Morpheus (Molecular Dimensions), Index (Hampton Research), and Crystal Screens (Hampton Research).

Crystallization conditions that gave initial crystal hits were usually categorized as a buffer with specific pH, salt, and a precipitant reagent. A grid screen was designed by varying pH and precipitant concentration around the condition in which crystals were observed, as salt concentration was kept constant. For grid screens, hanging drops were used and the volume increased in the drop by adding 1 µL of protein sample that contains UT-17460 and 1 µL of well solution. As optimization continues, a finer range of pH and precipitant concentration were used, the volume ratio of protein sample to condition that go into drop varied, and the whole drop size was varied. Also, a 96-well Additive screen (Hampton Research) was used to see if any of the additives improve the size and change the shape of crystals.

The initial crystallization screening for the LCP method started with mixing 8 mg/ml purified *Hy*MraY in addition of 1 mM UT-17460 and monoolein (sigma) to 1:2, 2:3, 4:5 volume-to-volume ratio using a syringe lipid mixer. The protein/lipid mixture was dispensed on a glass plate using a Gryphon (Art Robbins Instruments). The volume ratios of the sample to crystallization condition used were 0.05 µL to 0.8 µL, 0.05 µL to 1 µL, 0.1 µL to 1 µL, 0.15 µL to 1 µL, and 0.2 µL to 1 µL. Two commercial sparse matrices, MemMeso (Molecular Dimensions) and MemGoldMeso (Molecular Dimensions), were used. The plates were stored at room temperature.

Expression and purification of MthMraY

The pET22b-His_6-SUMO-GS_5-*Mth*MraY plasmid was transformed into the Nico21(DE3) pLEMO competent cells. The colonies were grown at 37°C until OD_{600} reaches 0.4 and

the flasks were transferred to 22°C. When OD_{600} reaches 0.6–0.8, the cells were induced with 0.4 mM IPTG and grown overnight. The cells were harvested using a JLA-8.1 rotor at 4,000 rpm for 15 min and resuspended in lysis buffer (20 mM Tris pH 7.5, 300 mM NaCl, 10 % glycerol, 5 mM βME, and protease inhibitors). The cells were lysed by flowing through a microfluidizer for four times. The lysate was spun down in a JLA-16.250 rotor at 12,000 rpm for 30 min. Pellets were discarded and the supernatant was spun down in a Ti-45 rotor at 45,000 rpm for 30 min.

The membrane pellets were either stored at -80°C or resuspended in an extraction buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 10 % glycerol, 5 mM βME, 10 mM imidazole, 1 % DM, and protease inhibitors). Solubilized membrane fraction was incubated for 2 hours at 4°C and spun down in a Ti-50.2 rotor at 45,000 rpm for 30 min. Pellets were discarded and the supernatant was incubated with 2.5 mL Ni-NTA resin (for 6L culture) overnight at 4°C. The sample was applied to a gravity column, washed with an imidazole gradient covering 20 mM, 30 mM, 50 mM, 80 mM, 110 mM, and 150 mM imidazole-containing buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 10 % glycerol, 5 mM βME, 0.15 % DM) and eluted with the similar buffer that contained 100 mM NaCl and 200 mM EDTA. The fractions were run on a SDS-PAGE gel, collected and dialyzed into 1 L of 20 mM Tris pH 7.5, 100 mM NaCl, 10 % glycerol, 1 mM DTT, and 0.15 % DM overnight at 4°C. Ulp1 (expressed and purified in the lab) was added to the dialyzed sample and incubated for one day at 4°C. Then, the sample was incubated with the 2.5 mL Ni-NTA resin overnight at 4°C. A reverse nickel-affinity column was performed by collecting the flow-through that contained the protein without a tag and washing the sample with 20

mM HEPES pH 7.5, 300 mM NaCl, 10 % glycerol, 0.15 % DM, 5 mM βME, and 30 mM imidazole. The rest of the sample was eluted with the sample buffer containing 200 mM EDTA. An SDS-PAGE gel was run to make sure the sample with no tag came off in the flow-through.

The tag-free protein was dialyzed into Buffer A (20 mM HEPES pH 7.5, 10 mM NaCl, 10 % glycerol, 5 mM βME, and 0.15 % DM) overnight at 4°C and ran on a cation-exchange column (Uno S6) with a gradient of 100 % Buffer A to 100 % Buffer B, containing 1 M NaCl. After an SDS-PAGE gel was run, the sample was run on a gel-filtration column (Superdex 200 10/300) with a running buffer, 20 mM HEPES pH 7.5, 150 mM NaCl, 10 % glycerol, 0.15 % DM, 5 mM βME, and 10 mM $MgCl_2$. Fractions were analyzed by SDS-PAGE and Coomassie staining and the fractions with pure protein were collected and concentrated.

Preparation of negative-stained grids of MthMraY

2 µL of purified *Mth*MraY (0.014 mg/ml) was added to a 400-mesh copper grid that was carbon coated using a Cressington 208carbon and glow discharged using an Emitech K100X (15 mA, 1 min). After 50 seconds, the excess protein sample on the grid was removed using Whatman No. 1 filter paper. Then, 2 µL of 2 % uranyl acetate was added on top. After 50 seconds, the excess was again blotted away at the edge of the filter paper.

Imaging a negative-stained grid of MthMraY using a 120 keV TEM

A Tecnai T12 equipped with an electron gun, LaB_6, and Gatan Ultrascan 2k x 2k CCD was used to image the negative-stained grid of *Mth*MraY. The 26,500- 52,000 magnification was used to visualize particles.

Expression and purification of HyMurG

The pET33b-His_6-thrombin (Leu-Val-Pro-Arg ↓ Gly-Ser)-*Hy*MurG plasmid was transformed into the Nico21(DE3) competent cells. The cells were grown at 37°C, shaking and induced with 0.4 mM IPTG when OD_{600} reached 0.6–0.8. After growing for four more hrs, cells were harvested using a JLA-8.1 rotor at 4,000 rpm for 20 min and resuspended in lysis buffer (20 mM Tris pH 8.0, 150 mM NaCl, 10 % glycerol, 5 mM βME, and protease inhibitors). The cells were lysed by flowing through a microfluidizer for three times. The lysate was spun down in a JLA-16.250 rotor at 12,000 rpm for 30 min. Pellets were discarded and the supernatant was spun down in a Ti-45 rotor at 45,000 rpm for 30 min. The membrane pellets were either stored at -80°C or resuspended in an extraction buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 10 % glycerol, 5 mM βME, 10 mM imidazole, 1 % DM, and protease inhibitors). It was incubated for 2 hours at 4°C and spun down in a Ti-50.2 rotor at 45,000 rpm for 30 min. Pellets were discarded and the supernatant was incubated with 1 mL cobalt resin (for 6L culture) for 2 hrs at 4°C, rocking. The sample was flown through a gravity column and washed with 50 CV of 20 mM HEPES pH 7.5, 1 M NaCl, 10 % glycerol, 10 mM imidazole, 0.15 % DM, 5 mM βME, and protease inhibitors, 50 CV of the same buffer with 300 mM NaCl and 30 mM imidazole and eluted with 150 mM NaCl and 200 mM imidazole. After the sample was checked for purity by SDS-PAGE and Coomassie staining, the sample was injected onto a gel-filtration column (either

Superdex 200 10/300 or 16/600) with 20 mM Tris pH 7.5, 150 mM NaCl, 10 % glycerol, 0.15 % DM, and 5 mM βME. Fractions from the peak were analyzed by SDS-PAGE, and some of the fractions were selected based on the purity for concentration using an Amicon 50 kDa cut-off concentrator and used for structural studies right away or stored at -80°C.

Co-crystallization of HyMurG with UDP-GlcNAc by sitting-drop vapor-diffusion method

6.74 mg/ml *Hy*MurG and 10 mM UDP-GlcNAc were incubated on ice for 30 min, filtered, and set up with crystallization conditions in a ratio of one-to-one (0.2 µl + 0.2 µl) with a 50-µL reservoir on a 96-well MRC plate using Mosquito. Thus, the final concentration of *Hy*MurG and UDP-GlcNAc in the drop was 3.37 mg/ml and 5 mM, respectively. Three commercial screens used were Index (Hampton Research), MemGold (Molecular Dimensions), and MemGold2 (Molecular Dimensions). The plates were stored at room temperature.

Optimizing co-crystallization of HyMurG with UDP-GlcNAc using hanging-drop vapor-diffusion method

Initial crystals hits were observed in MemGold F12, MemGold A2, MemGold2 H3, and Index H3. Among them, MemGold F12, MemGold A2, and Index H3 conditions were used to design 24-well grid screens by varying buffer pH and precipitant concentration. 5.5–11mg/ml *Hy*MurG and 10 mM UDP-GlcNAc were incubated on ice for 30 min, filtered, and set up with crystallization conditions in a ratio of one-to-one (1 µl + 1 µl) with a 300-µL reservoir on a 24-well VDX Plate with sealant (Hampton Research). For several trays, 0.5 µl microseeds were introduced into each drop. Thus, the final

concentration of *Hy*MurG and UDP-GlcNAc in the drop was 2.7–5.5 mg/ml and 3.6-5.0 mM, respectively.

ACKNOWLEDGEMENTS

We are grateful to Jens Kaiser (Caltech) for crystallography help through the Molecular Observatory. Operations at SSRL are supported by the U.S. Department of Energy (DOE) and the National Institutes of Health (NIH). This work was supported by MraY grant.

C h a p t e r 4

Structural elucidation of MraY in complex with phage ΦX174 protein E, a novel inhibitor to combat antibiotic-resistant *Escherichia coli* infection

ABSTRACT

Antibiotic- and multidrug-resistant pathogens pose a great threat to public health worldwide. Even though *Escherichia coli* is one of the most studied microorganisms, resistance mechanisms of its pathogenic strains are not fully understood. As one way of tackling antibiotic-resistant pathogenic *E. coli* strains, a key integral membrane protein called MraY, which is involved in peptidoglycan (PG) biogenesis, has been an attractive target for developing antibacterial drugs. Unfortunately, none of the inhibitor compounds targeting MraY is in clinical use due to their low cell permeability and cellular concentration. MraY is also the target in *E. coli* for the lysis protein E from a small single-stranded bacteriophage ΦX174. However, their molecular interactions have not been structurally elucidated. Here, I describe how *E. coli* MraY in complex with protein E was purified, reconstituted in non-detergent systems, and was subject of structural studies using X-ray crystallography and electron cryomicroscopy (cryo-EM) single particle analysis. The goal of this study is to determine a high-resolution structure of the *Ec*MraY-protein E complex and this will pave a way to developing a novel type of anti-bacterial drugs.

INTRODUCTION

Escherichia coli is a commensal organism in the human gut and uropathogenic *E. coli* (UPEC) is the most common cause of urinary tract infection with emerging antibiotic resistance (Flores-Mireles, Walker, Caparon & Hultgren, 2015). A common strategy for developing antibacterial drugs is to target the peptidoglycan (PG) that forms a major component of the bacterial cell wall. However, UPEC has already shown resistance to some β-lactam antibiotics (Blango & Mulvey, 2010) that inhibit penicillin-binding proteins (PBPs) preventing cross-linking of the GlcNAc-MurNAc polymer in the periplasm (*e.g.*, penicillin, nafcillin, cefadroxil). Blocking enzymes involved in earlier steps of the PG synthesis is one way to design new antibacterial agents. A promising candidate is MraY, a key integral membrane protein that is essential for cell viability and catalyzes the transfer of phospho-MurNAc-pentapeptide from a nucleotide-activated form, UDP-MurNAc-pentapeptide, to the lipid carrier, C_{55}-P.

Bernhardt et al. provided genetic evidence showing that *E. coli* MraY is the cellular target of protein E, encoded from a single lysis gene in the small single-stranded bacteriophage ΦX174 that lyses *E. coli* cells. (T. G. Bernhardt, Roof & Young, 2000). E-mediated lysis requires the host *slyD* gene, which encodes an FK506 binding protein (FKBP)-type peptidyl-prolyl cis-trans isomerase (PPIase). SlyD is thought to stabilize protein E, which allows it to be accumulated in the membrane and helps it to lyse *E. coli* cells (Thomas G. Bernhardt, Roof & Young, 2002). Rodolis et al. constructed a helical wheel model and proposed the possible interaction site between the transmembrane domain of protein E and transmembrane helix 9 of *E. coli* MraY (Rodolis et al., 2014). In

order to confirm the site of action of protein E on *Ec*MraY, determining a high-resolution structure of the *Ec*MraY in complex with protein E and *Ec*SlyD is necessary. Details of their molecular interactions in the structure will shed light on developing protein E as a therapeutic against *E. coli*.

This study discusses our success in expressing and purifying the protein complex that contains three protein components (*Ec*MraY, protein E, *Ec*SlyD), followed by how the purified protein complex was used for structural characterization using X-ray crystallography and cryo-EM single particle analysis. Two constructs that were used are a bicistronic vector containing protein E from an isoform of ΦX174, ID21, that has a His-tag at the C-terminus and *E. coli* MraY and the other vector that contains the first 154 residues of *E. coli* SlyD. Protein E from ID21 was chosen as it is shorter (76-residues) compared to the 91-residue protein E from ΦX174, while the transmembrane domain is conserved. Also, ID21 protein E in the complex contains the mutation, L19F, which was reported to bypass the SlyD requirement for E-mediated lysis (T G Bernhardt, Roof & Young, 2000) and also recover the cell-lysis activity that was lost in a truncated construct of ΦX174 protein E, likely by promoting protein-protein interactions (Tanaka & Clemons, 2012). Together, this protein complex will be called EYS21-L19F (protein E, *Ec*MraY, *Ec*SlyD; protein E is from ID21 with L19F mutation).

RESULTS

Expression and purification of the EYS21-L19F complex.

Two plasmids that contain three components of the EYS21-L19F complex (*Ec*MraY, protein E from ID21, and *Ec*SlyD1–154) were co-transformed and expressed in *E. coli* cells. Following cell lysis, membrane pellet was obtained and the protein was solubilized with 1% DM (weight-to-volume) containing buffer. The protein complex was purified via a nickel-affinity, anion-exchange, and gel-filtration columns (**Fig. 4.1**). A shoulder on the right side of the peak in the gel-filtration chromatogram was observed for some cases (**Fig. 4.1.C**) while the peak was perfectly symmetrical in other cases (**Fig. 4.1.D**). Note that some contaminants that were larger than *Ec*MraY were carried over from the nickel-affinity column to gel-filtration column (**Fig. 4.1**).

Co-crystallization of EYS21-L19F with the S-analog of UDP-MurNAc-pentapeptide.

We were provided with an S-analog of UDP-MurNAc-pentapeptide, one of the oxygens in the α-phosphate group is replaced by a sulfur (**Fig. 4.2**), by the Kurosu laboratory at the University of Tennessee Health Science Center. EYS21-L19F was co-crystallized with this analog with the rationale that it would behave as a substrate without being catalyzed. Based on the preliminary testing, when more than 100 μM of the analog was added to the reaction of MraY from *Hydrogenivirga sp.*, approximately 10 % inhibition was observed (Kurosu laboratory).

While the substrate analog was not tested directly on *E. coli* MraY, we predicted UDP-MurNAc-pentapeptide would bind to *Ec*MraY similar to *Hy*MraY. In the presence of 100

μM of the analog, crystal formation was readily obtained int two commercial screens, MemGold2 and Morpheus. There were three different crystal morphologies during the initial screening with the sparse matrices: cubes, pyramids, and three-dimensional trapezoids (**Fig. 4.3.A-C**). The conditions that produced crystals were varied. During diffraction screening, two diffraction images were collected 90° apart on several of the crystals with a cube-looking crystal diffracting to 7.4 Å (**Fig. 4.4.A**) and trapezoid-looking crystals to 6.3 Å (**Fig. 4.4.B**) and 6.6 Å. X-ray diffraction data were remotely collected at the SSRL BL12-2 and data processing was performed using XDS (Kabsch, 2010). With all three data sets, XDS suggested the space group number 155 (H32) with a three-fold symmetry operator.

Attempts to solve the phases using molecular replacement were not successful using Phenix (Adams, Pavel, et al., 2010).

Initial crystallization conditions were further optimized by designing and setting up grid screens varying buffer pH and precipitant concentration. Switching from sitting-drop to hanging-drop vapor-diffusion method with larger drop size also increased the crystal size (**Fig. 4.3.D-G**). Some of the crystals from hanging-drop trays were screened at the APS 23-ID-B (**Fig. 4.5**). Unfortunately, most of the crystals appeared to have been kept in the trays too long and were likely dehydrated, causing streaky spots on the diffraction images or no diffraction at all. After this trip, the EYS21- L19F crystals were not able to be reproduced in the presence of the substrate analog for some time.

EYS21-L19F in amphipols imaged in negative stain.

Amphipathic polymers (amphipols or Apols) were developed (Tribet, Audebert & Popot, 1996) as solubilizing agents to replace detergents and stabilize membrane proteins with the most popular being Amphipol A8-35. Amphipols wrap around a membrane protein stabilizing them in buffer without requiring any detergent. I previously observed that the EYS21-L19F sample purified in dodecyl maltoside (DDM) aggregated when it was negative-stained and imaged using a TEM, Tecnai T12 operating at 120 keV. I speculated that aggregation was possibly coming from excess detergent micelles, I mixed the detergent solubilized protein with amphipols and were removed the detergent by using beads. The amphipol stabilized protein was injected onto a gel-filtration column (Superdex 200 10/300) was run with buffer without any detergent resulting in a nearly symmetrical peak (**Fig. 4.6.A**). The fractions in the peak area were concentrated, incubated with the substrate analog, and added to a copper grid, followed by the addition of 2 % uranyl acetate. The grid was imaged using a Tecnai T12 and the particles looked homogeneous and no other contaminants were detected (**Fig. 4.6.BC**). The same protein sample with the substrate analog was used to prepare a cryo-EM grid and a data set was collected from a Talos Arctica operating at 200 keV. Data processing on the collected micrographs was initiated with Relion-2 (Kimanius, Forsberg, Scheres & Lindahl, 2016; Scheres, 2012), but it was difficult to pick out individual particles and 2D classes were not obtained (data not shown here).

DDM was selected from a detergent screening for extracting EYS21-L19F.

In order to select a detergent that can extract the EYS21-L19F complex as native form as possible, detergent screening was performed during the extraction step of purification. 1 % weight-to-volume amount of each of six detergents, DDM, DM, LDAO, β-OG, Fos-choline-12, and Cymal-5, was added to the buffer that was used to resuspend the membrane pellet and rocked for two hours at 4 °C. The samples were spun in an ultracentrifuge rotor and the supernatant from each detergent-extracted sample was saved to incubate with 250 µL of Ni-NTA resin. Six nickel-affinity columns were run in parallel and the purity of the fractions were checked by SDS-PAGE (**Fig. 4.7.A**). The amount of *Ec*MraY extracted was similar among DDM-, DM-, and Cymal-5-extracted samples, whereas the other three detergent-extracted samples showed thin double bands at the size, in which *Ec*MraY is supposed to appear as a smeared band. Elution fractions were collected from each nickel column, concentrated, and run on a gel-filtration column (**Fig. 4.7.BD**). The obvious differences were the height of the main peaks, indicating the amount of the protein complex extracted, and the size of the shoulders. DDM was the best among them because the main peak is the highest, almost symmetrical, and had a shoulder only on one side of the peak. Purity of all of the detergent samples from the columns was checked by SDS-PAGE with the DDM fractions shown here (**Fig. 4.7.C**). After this screening, DDM instead of DM was used to extract the EYS21-L19F complex in the subsequent purification preparations.

Optimization of purification of the EYS21-L19F complex.

Previous purification protocol of the EYS21-L19F provided high yield of almost pure protein sample, but contaminants that are bigger than *Ec*MraY always appeared on a

SDS-PAGE gel after every step of purification (**Fig. 4.1**). This problem was resolved by using a cobalt-affinity instead of nickel-affinity column at the expense of some yield (**Fig. 4.8.A**). The protein complex looks pure on an SDS-PAGE gel after a gel-filtration column, but there were two small shoulders on the left side of the peak, indicating purity can be improved further (**Fig. 4.8.C**). Running an anion-exchange column (**Fig. 4.8.B**) in between a metal-affinity and gel-filtration chromatography helped reduce the size of the shoulders.

Purification of three membrane scaffold protein (MSP) variants.

It has been shown that interactions between membrane proteins and its surrounding lipids are important to maintain protein function (Phillips, Ursell, Wiggins & Sens, 2009; Saliba, Vonkova & Gavin, 2015; Zhou & Cross, 2013). The nanodisc technology was developed to provide a lipid bilayer environment to membrane proteins and has been widely used for structural and functional studies of proteins (Denisov & Sligar, 2016, 2017). Nanodiscs refer to the disc-shape formed when the amphipathic helices of a MSP (Bayburt, Grinkova & Sligar, 2002) wrap around phospholipids like a belt. Depending on the number of transmembrane helices in a protein of interest and the phospholipids used, MSP variants of different lengths are selected and that are compatible with the expected diameter of nanodiscs that will solubilize the protein in lipids. Here, three MSP containing plasmids, pMSP1D1, pMSP1E3D1, and pMSP2N2, were overexpressed in *E. coli* cells and purified using a nickel-affinity column (**Fig. 4.9.ACE**) and gel-filtration column (data not shown). In the purification process, a His-tag was cleaved off by digesting the samples with a TEV

protease and then removed via a reverse nickel-affinity column (**Fig. 4.9.BDF**). Purified MSPs were stored at -80°C until they were used for nanodisc assembly. Phospholipids selected were DMPC, DMPG, POPC, and POPG because their phase transition temperatures are close to room temperature or 4°C, which can be controlled easily during assembly. Either single or a pair of phospholipids were used for assembly.

Reconstitution and imaging of EYS21-L19F into nanodisc composed of MSP1E3D1 and DMPC.

Reconstitution of EYS21-L19F into nanodiscs was performed by incubating purified protein, purified MSP1E3D1, and DMPC in the presence of detergent, sodium cholate, at room temperature for an hour. Detergent was removed using polystyrene beads. The sample containing proteins reconstituted in nanodiscs were incubated with Ni-NTA resin overnight at 4°C. On the next day, a nickel-affinity column was run and what newly appeared on an SDS-PAGE gel in addition to the EYS21-L19F complex was MSP1E3D1 (**Fig. 4.10.A**). The main purpose of running the protein-nanodisc assembly via a nickel-affinity column was to remove all the empty nanodiscs as the His-tag on the EYS21-L19F complex would only bind to the resin. Subsequently, the eluates were run through a gel-filtration (Superdex 200 10/300) column and the fractions with all components were collected for concentration (**Fig. 4.10.BC**). The peak from the gel-filtration chromatogram had two large shoulders that came off after the main peak indicating the sample was not completely homogeneous. The purified EYS21-L19F assembled in nanodiscs with MSP1E3D1 and DMPC was added to a glow-discharged copper grid and negative-stained with 2 % uranyl acetate. The images collected from a Tecnai T12

operating at 120 keV showed homogenous particles (**Fig. 4.11**). The same sample was used to prepare cryo-EM grids.

Cryo-EM sample preparation and data collection of EYS21-L19F-nanodisc assembly. The EYS21-L19F-nanodisc sample was added to a glow-discharged Quantifoil R2/2 grid using a Vitrobot by varying blot time (2, 4, 6, 8 sec) and keeping 100 % humidity and -5 blot force with two different concentrations of the protein sample (4 mg/ml or 0.4 mg/ml). Grids were subsequently plunge frozen into liquid ethane that was cooled by liquid nitrogen. All of the grids were screened in a Talos Arctica operating at 200 keV and in only one grid (0.4 mg/ml, 100% humidity, -5 blot force, 2 sec blot time) were particles present in the holes of the grid. The selected cryo-EM grid was used to collect two-day movies at a 165,000x magnification on a Titan Krios operating at 300 keV.

The micrographs collected in this data set was motion corrected using MotionCor2 (S. Q. Zheng et al., 2017) implemented in Relion-3 and CTF estimated. Then, 1,089 particles were manually picked from 89 selected micrographs (**Fig. 4.12.A**). Micrographs with no particles were manually omitted from the whole data set and 2D classification was performed with the manually picked particles (**Fig. 4.12.B**). Six 2D classes were selected and averaged to be used as a template to auto-pick the rest of the particles (**Fig. 4.12.C**). Initial 3D initial models were derived from the picked particles (**Fig. 4.12.D**). The 3D class averages from one or more classes were visualized using Chimera (Pettersen et al., 2004) (**Fig. 4.13**). The resolution of the densities was too low to fit known structural components of the protein complex. The expectation is that some of the soluble protein domains would

be sticking out of the nanodiscs; yet this density was inconsistent throughout the initial models, suggesting this part may be flexible. For future data collection, more particles for each orientation would be a goal.

Other nanodisc assemblies with different combinations of MSP and phospholipid were imaged in negative stain.

After the cryo-EM data analysis from the first EYS21-L19F-nanodisc assembly was assessed, nanodiscs with different combinations of MSPs and phospholipids were prepared. EYS21-L19F reconstituted in nanodiscs with MSP1D1 and DMPC was added to a 200-mesh copper grid along with 2 % uranyl acetate. The images collected in a Tecnai T12 showed heterogeneous populations of particles, especially particles that look smaller than the expected size of 10–15 nm (data not shown).

Next, nanodisc prepared and purified with EYS21-L19F consisted of MSP2N2 and DMPC/DMPG. Imaging on a Tecnai T12 after negative staining showed that particles tend to clump together with smaller particles still present (data not shown).

The next nanodisc prepared and purified with EYS21-L19F consisted of MSP1E3D1 and POPC/POPG (**Fig. 4.14**). Imaging on a Tecnai T12 after negative staining showed again that particles tend to clump together and small circular particles were still present (**Fig. 4.15**).

In an attempt to improve heterogeneity, the nanodisc (POPC/POPG, MSP1E3D1) assembly of EYS21-L19F was run on an anion-exchange column (Uno Q6). Three

different fractions from the column were individually added to 400-mesh copper grids with a carbon layer, stained with 2 % uranyl acetate and they were imaged on a Tecnai T12. However, heterogeneity that is seen in **Fig. 4.15** did not improve (data not shown here). This indicates that the sample was likely heterogeneous in the assembly step and could not be purified further by size or charge in nanodisc. Heterogeneity could be due to partially pure MSPs and/or mixing two different phospholipids for nanodisc assembly. For the former case, I previously combined MSP2N2 purified from soluble and insoluble fractions and used it for nanodisc assembly above, while MSP1E3D1 used above was purified only from cell debris pellet. However, I obtained mixed results in terms of heterogeneity from the nanodiscs with MSP2N2 and MSP1E3D1. For the latter case, I started with DMPC to prepare nanodisc assembly, but later I used DMPC/DMPG and POPC/POPG. Using two different phospholipids could have caused heterogeneity in the samples.

Further purification of MSP2N2 for the nanodisc assembly.

The next step taken to achieve homogeneity of nanodiscs was taking individual 1-mL fractions of MSP2N2 from a gel-filtration column and using each fraction and either POPC or POPG alone to assemble with EYS21-L19F (**Fig. 4.16**). The compositions of nanodiscs tried here were (a) MSP2N2 gel-filtration fraction 11 with POPC, (b) MSP2N2 gel-filtration fraction 11 with POPG, (c) MSP2N2 gel-filtration fraction 12 with POPC, (d) MSP2N2 gel-filtration fraction 12 with POPG, (e) MSP2N2 gel-filtration fraction 13 with POPC, and (f) MSP2N2 gel-filtration fraction 13 with POPG. For this preparation, the

nanodisc assemblies were not run through another gel-filtration column to remove empty nanodiscs. From negative-stain imaging of these assemblies on a Tecnai T12 (**Fig. 4.17**) there was more clumping of particles when MSP2N2 gel-filtration fraction 11 was used, whereas smaller particles were clearly observed when fraction 13 was used. The difference in the samples prepared from either POPC or POPG was not very obvious and inconclusive. This could be due to inconsistent coverage of carbon coating that was manually introduced to the grids. Nevertheless, the best grid among them was from the nanodisc assembly with MSP2N2 gel-filtration fraction 12 and POPC (**Fig. 4.17.C**).

Further optimization of crystals of EYS21-L19F with the S-analog.

In efforts to reproduce crystals previously acquired (**Fig. 4.3**), grid screens with a broad range of pH and precipitant concentrations were attempted using the prior conditions that gave crystals for the EYS21-L19F complex incubated with the analog. However, crystals with sharp edges were no longer observed. In the meantime, the purification protocol of the EYS21-L19F complex was optimized as described above. The final buffer salt concentration was also considered by lowering NaCl concentration (25, 50, 75 mM) for the gel-filtration column. In comparison to 100 mM NaCl-containing buffer that had initially been used, as NaCl concentration decreases, the sample comes off the column slightly faster (**Fig. 4.18**). The main protein peak is present across the various NaCl concentration, whereas the peak is missing when no NaCl is used (**Fig. 4.18**).

Subsequently, more EYS21-L19F was purified in the final buffer with 25 mM NaCl, incubated with the analog for 30 min on ice, and trays were set up using several

commercially available sparse matrix screens by the sitting-drop vapor-diffusion method. Among them, crystals were finally reproduced in either pyramid- or cube-shape, although they appeared much smaller than before. The conditions that crystallized the protein complex were MemGold2 D5, F3, F7, G1, H2, and Index H2, H3, H10. For some of these, grid screens with a range of buffer pH and precipitant concentration were designed and set up using the hanging-drop vapor-diffusion method. Among them, we were able to consistently reproduce crystals in the grid screens of MemGold F7 (0.1 M MES pH 6.5, 0.37 M potassium nitrate, 22 % PEG 400) and G1 (0.05 M HEPES pH 6.5, 0.5 M potassium chloride, 20 % PEG 400). Several hundreds of crystals were frozen and shipped to the SSRL BL12-2 and some X-ray diffraction data sets were collected. The best data set diffracted to 4.20 Å (**Fig. 4.19** and **Table 4.1**).

DISCUSSION

Determining a high-resolution structure of the EYS21-L19F complex had been a challenge, and this could be due to many reasons. This study discussed some modifications introduced to the purification process and new approaches to overcome various issues. First, some high molecular weight impurities that were present in the protein sample after purification could have disrupted crystal packing. By switching from a nickel-affinity to cobalt-affinity column, binding specificity toward the His-tag was increased and those contaminants were removed. Second, the addition of the S-analog to the protein sample greatly enhanced crystallization. This indicates that the S-analog may be binding to the active site in a similar way as the native substrate, UDP-MurNAc-pentapeptide, which may lead to stabilizing certain conformations of *E. coli* MraY. Third, replacing all the detergent molecules with amphipathic polymers to stabilize the integral membrane protein MraY prevented aggregation of particles when imaged after negative staining. We collected and processed movies from a Talos Arctica operating at 200 keV, but the cryo-EM map we obtained was at a low resolution and we were unable to improve the resolution (data not shown here).

Amphipols surrounding the protein complex do not add much size or have unique structural features and the overall complex size is challenging for the software to align and sort out particles in different orientations. This led to a switch to the nanodisc reconstitution system, which has potentially many benefits over the use of detergent and amphipols. One advantage of using nanodiscs over amphipols for the cryo-EM single particle analysis was its disc-shaped feature helped to pick out particles and easily

recognize their different orientations on the micrographs. Nevertheless, the cryo-EM data set collected from the fist nanodisc assembly with EYS21-L19F did not significantly improve the resolution. A possible explanation could be the flexible cytoplasmic side of the complex, where the soluble domain of protein E and SlyD interact. Determining a structure only of this region would help us to understand how protein E and SlyD interact and we could then, for example, introduce cross-links to stabilize this region. Increasing the number of particles and homogeneity could also improve the low-resolution EM maps.

More efforts toward improving heterogeneity of the EYS21-L19F-nanodisc assembly sample and varying conditions for vitrification of cryo-EM grids are needed. The EYS21-L19F complex is estimated to be slightly larger than 100 kDa, from which long has been impossible to solve high-resolution cryo-EM structures. Recently, technical advancements with cryo-EM enabled obtaining cryo-EM maps for alcohol dehydrogenase (82 kDa) at 2.9 Å and methemoglobin (64 kDa) at 2.8 Å (Herzik, Wu & Lander, 2019). Along with traditional X-ray crystallography, cryo-EM is a powerful technique to solve structures of protein complexes and, for us, a high-resolution structure of *Ec*MraY-protein E complex will contribute to development of novel antibacterial drugs in the time of antibiotic resistance crisis.

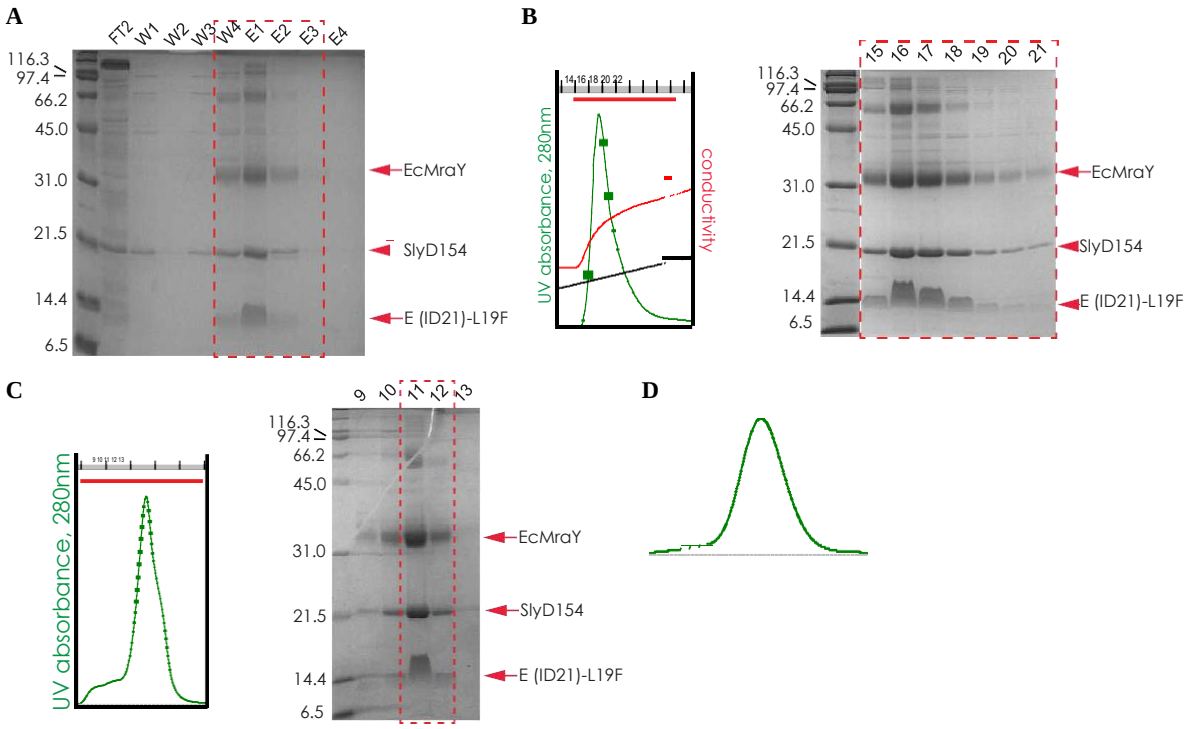

Figure 4.1. Purification of EYS21-L19F. A, an SDS-PAGE gel from a nickel-affinity column run with EYS21-L19F. The protein complex came off in the W4-E3 fractions. FT2, second flow-through; W1–2, wash with 20 mM imidazole and 10 mL each; W 3–4, wash with 20 mM EDTA and 10 mL each; E1–4, elute with 200 mM EDTA and 5 mL each; the first lane is Broad Range protein marker in kDa from Bio-Rad; B, purification of EYS21-L19F using an anion-exchange column. Left, an anion-exchange (Uno Q6) chromatogram shows a single peak. Right, fractions 15–21 were run on an SDS-PAGE gel and there are high MW contaminants above the *Ec*MraY bands; C, purification of EYS21-L19F using a gel-filtration column. Left, a gel-filtration (Superdex 200 10/300) chromatogram shows a peak with a right shoulder.
Right, fractions 9–13 were run on an SDS-PAGE gel and high MW contaminants are still there; D, in another purification batch, a gel-filtration column also gave a symmetrical peak without any shoulder.

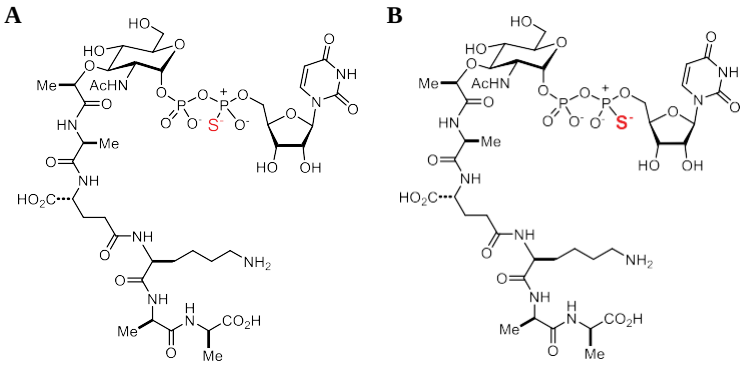

Figure 4.2. Chemical structure of the S-analog of UDP-MurNAc-pentapeptide in a diastereomer mixture. The Kurosu laboratory synthesized this analog and purified using a HPLC HYPERSIL GOLD™ column with the solvent ratio of MeCN : 0.05 M NH_4HCO_3 (aq) = 5 : 95 and 2.0 mL/min flow rate. The product was detected at 254 nm.

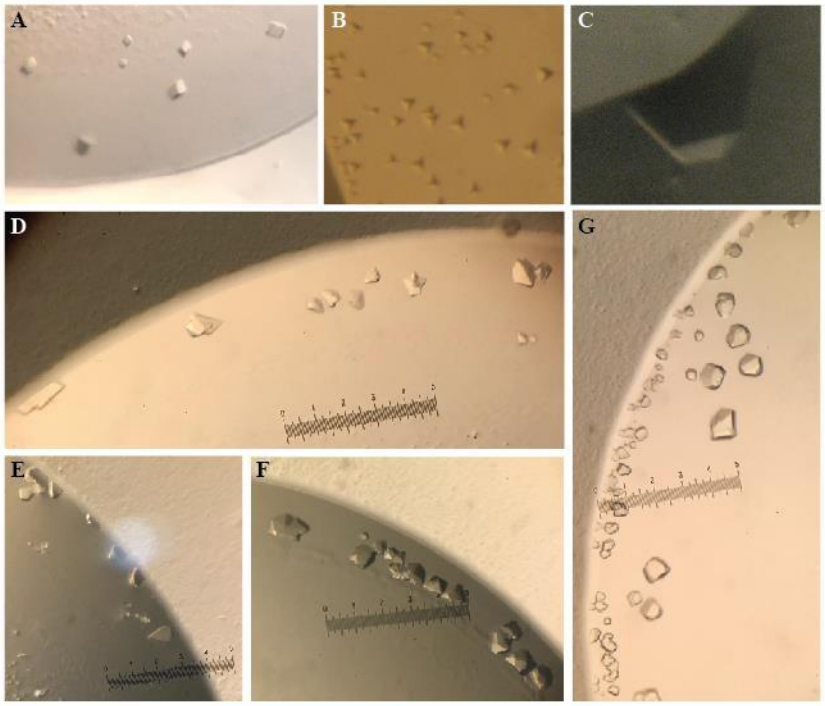

Figure 4.3. Initial and optimized crystal hits of EYS21-L19F co-crystallized with the S-analog. 6 mg/ ml of EYS21-L19F and 100 µM analog formed crystals via the sitting-drop (A-C) or hanging-drop (D-G) vapor-diffusion method. A-C, three forms of crystals observed in the commercial screens (MemGold2 and Morpheus). D&E, 0.1 M MOPS pH 7.5, 0.2 M NaCl, and 38 % PEG 400 (MemGold2 E7). F&G, 0.05 M Tris pH 9.0, 0.3 M ammonium formate, and 33 % PEG 500 MME (MemGold2 F2).

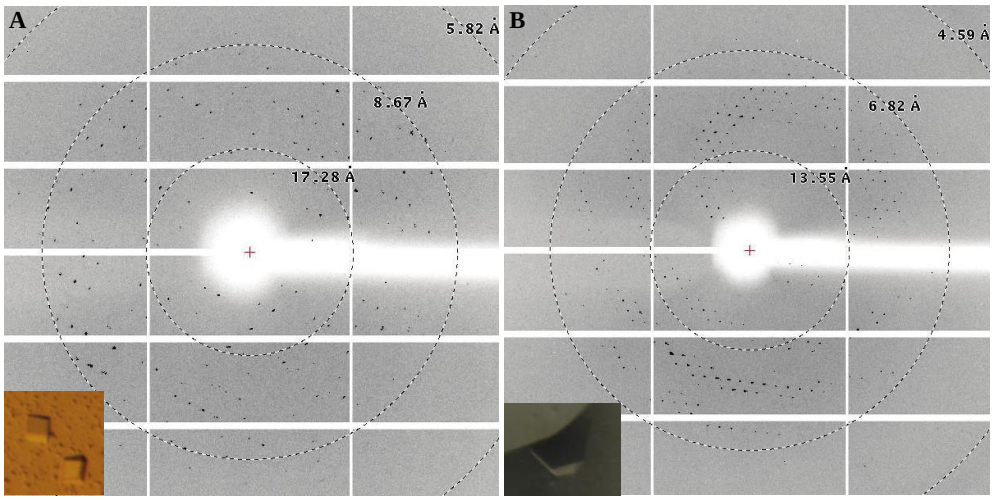

Figure 4.4. 7.4 Å and 6.3 Å X-ray diffraction data collected on EYS21-L19F co-crystallized with the S-analog at the SSRL BL12-2. 6 mg/ml of EYS21-L19F and 100 µM analog formed crystals via the sitting-drop vapor-diffusion method. A, 0.01 M MES pH 6.5, 0.1 M sodium chloride, 0.15 M ammonium sulfate, and 19 % PEG 1000 (MemGold2 B1). B, 0.1 M Carboxylic Acids, 0.1 M Buffer System 2, pH 7.5, and 50 % Precipitant Mix 3 (Morpheus G7).

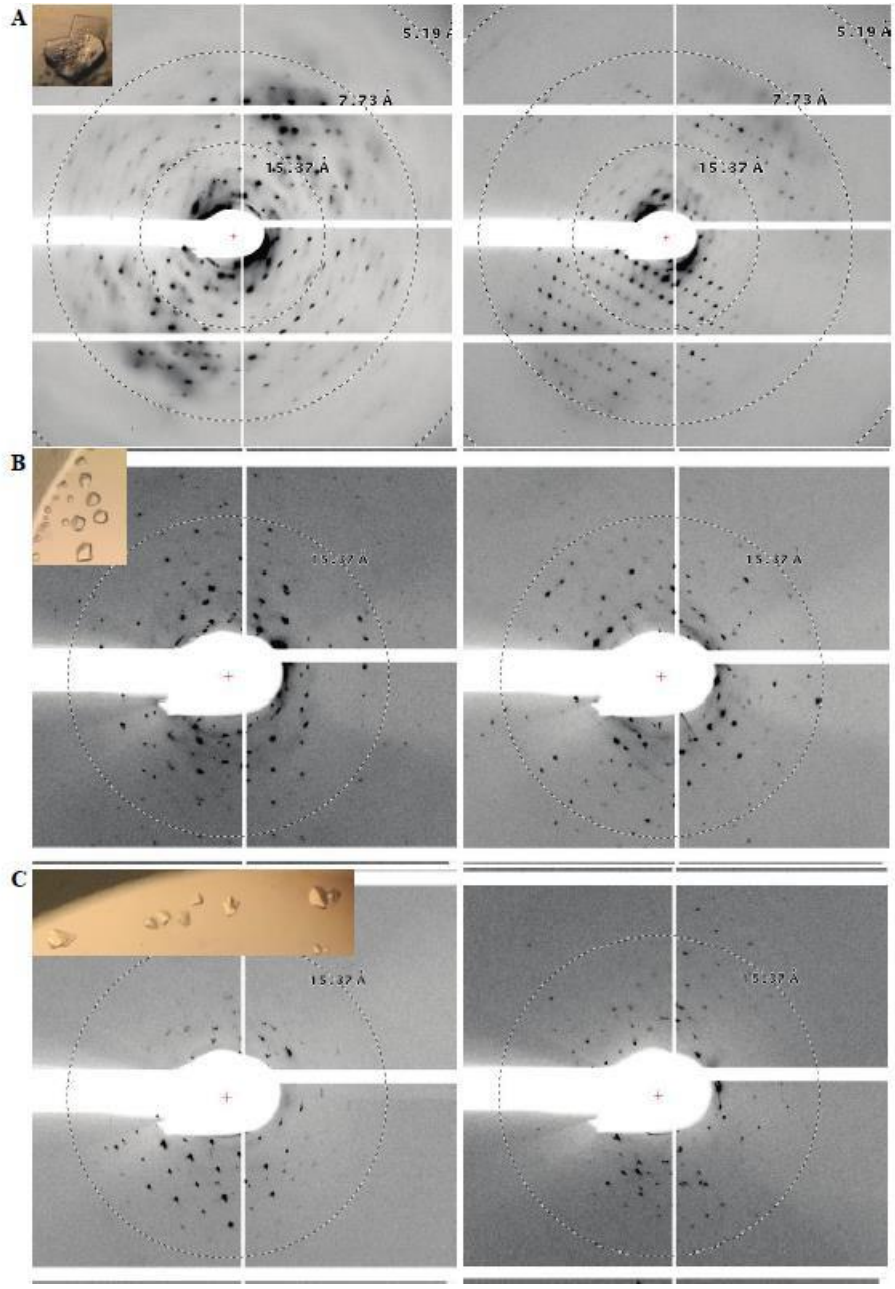

Figure 4.5. Crystals of EYS21-L19F with the S-analog screened at the APS 23-ID-B. Two X-ray diffraction images were collected 90° apart from each crystal. The drops where individual crystals were harvested are shown in the top left corner of the first diffraction images. A, 0.1 M HEPES pH 7.5, 0.04 M $MgCl_2$, and 32 % PEG 400 (MemGold2 F9). B, same drop from **Fig. 4.3.G**. C, same drop from **Fig. 4.3.D**.

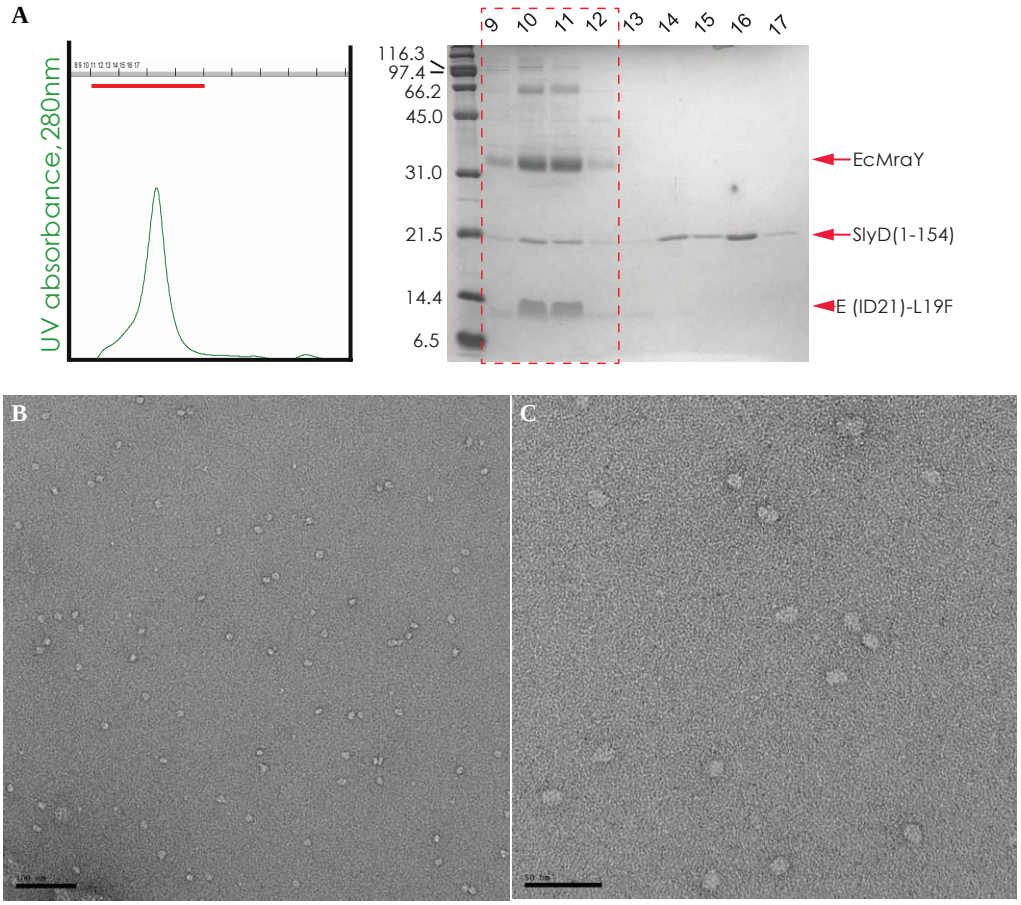

Figure 4.6 Reconstitution of EYS21-L19F in amphipols and negative-stained images of the complex with the S-analog. A, a gel-filtration (Superdex 200 10/300) chromatogram and an SDS-PAGE gel with the gel-filtration fractions show the EYS21-L19F complex is stable in Apols. B&C, a negative-stained grid with 1.14 mg/ml protein and 210 µM analog was imaged using a Tecnai T12. The scale bar is 100 nm in B and 50 nm in C.

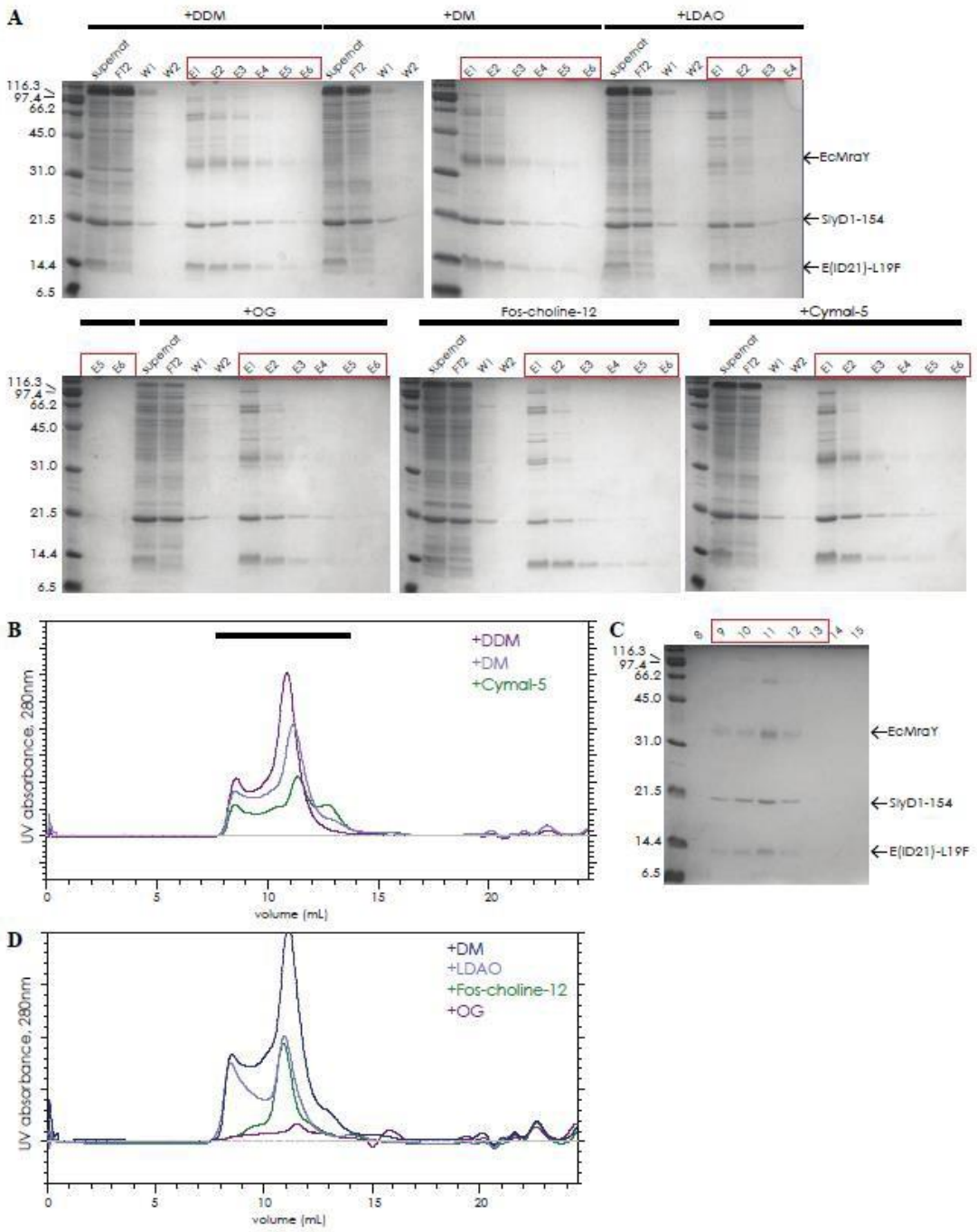

Figure 4.7. Detergent screening for extraction of EYS21-L19F. A, an SDS-PAGE gel of individual nickel-affinity columns runs with the EYS21-L19F complex extracted by six different detergents (DDM, DM, LDAO, β-OG, Fos-choline-12, and Cymal-5). Each detergent added is written on the top of the gels. Supernat, supernatant of the last ultracentrifuge run that contains solubilized protein; FT2, second flow-through; W1–2, 5 mL each; E1–E5, 1 mL each; E6, 5 mL. B, a gel-filtration (Superdex 200 10/300) chromatogram with overlaid peaks from DDM-, DM-, and Cymal-5- containing running buffers. The black bar on the top indicates which area of the peak or fractions from the DDM run were used on a SDS-PAGE gel in C. D, a gel-filtration (Superdex 200 10/300) chromatogram with overlaid peaks from DM-, LDAO-, Fos-choline-12-, and β-OG-containing running buffers.

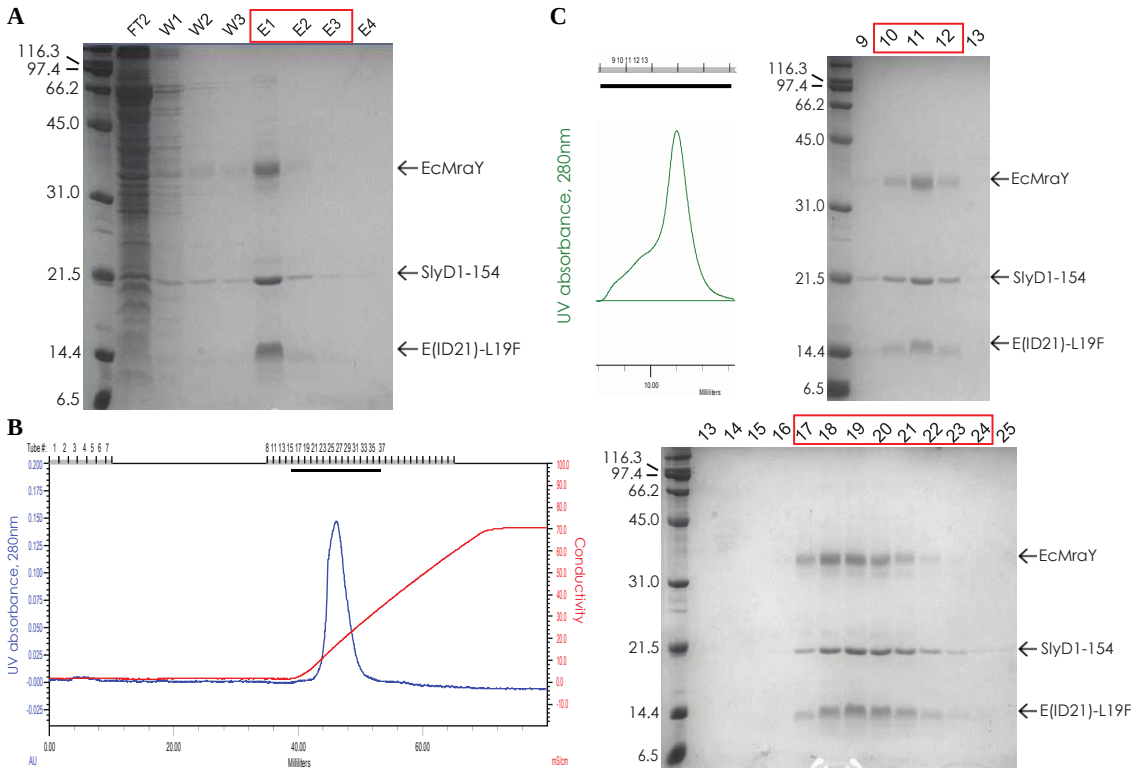

Figure 4.8. Optimized purification of EYS21-L19F. A, an SDS-PAGE gel from a cobalt-affinity column run with EYS21- L19F. The protein complex came off in the E1–3 fractions. FT2, second flow-through; W1–3, wash with 20 mM imidazole and 10 mL each; E1–4, elute with 200 mM imidazole and 5 mL each; the first lane is Broad Range protein marker in kDa from Bio-Rad; B, purification of EYS21-L19F using an anion-exchange column. Left, an anion-exchange (Uno Q6) chromatogram shows a single peak. Right, fractions 13–25 were run on an SDS-PAGE gel and the protein complex looks pure; C, purification of EYS21-L19F using a gel-filtration column. Left, a gel-filtration (Superdex 200 10/300) chromatogram shows a peak preceding with two little shoulders. Right, fractions 9–13 were run on an SDS-PAGE gel and the protein complex is pure.

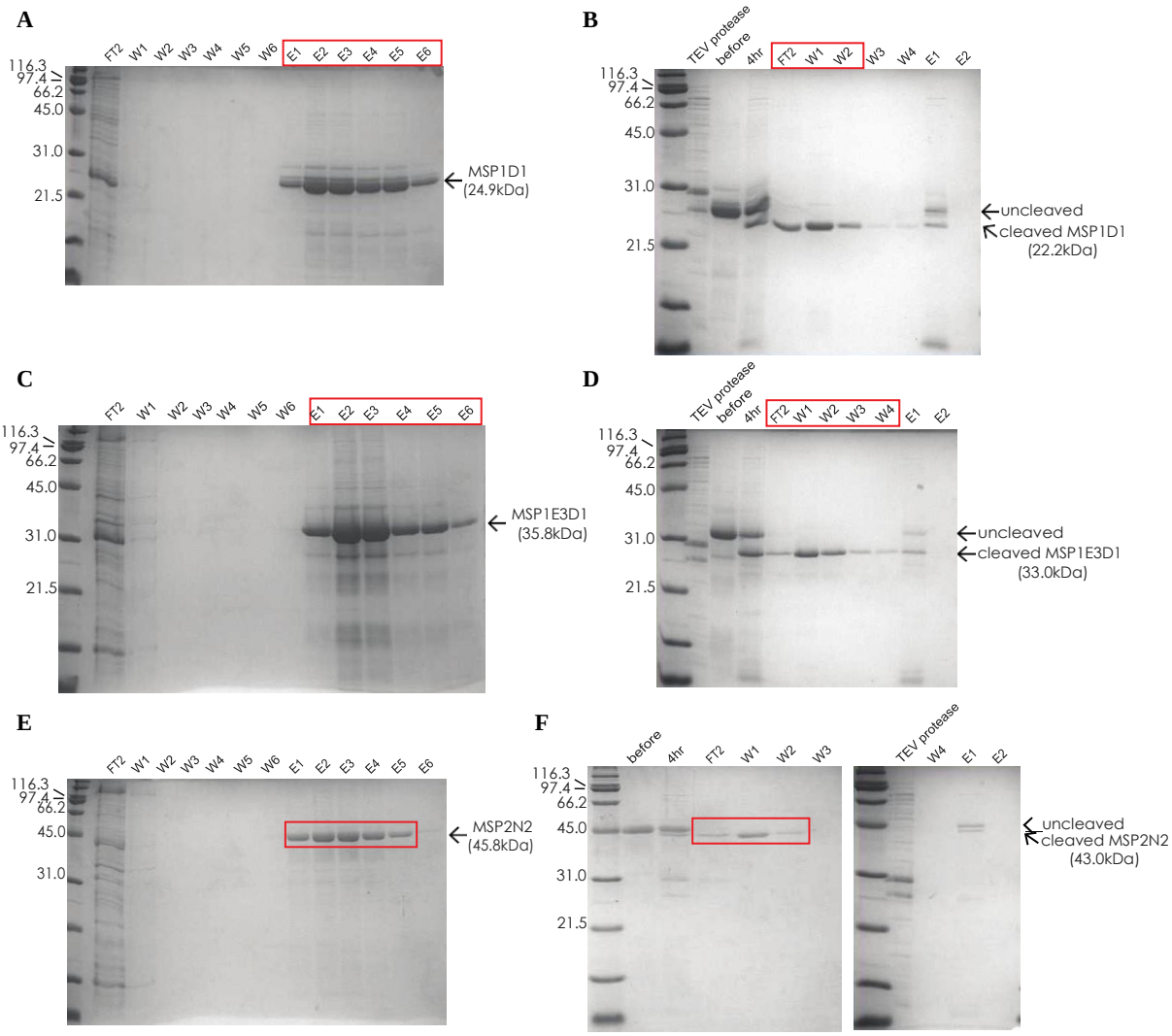

Figure 4.9. Purification of membrane scaffold proteins (MSPs) from *E. coli* cell debris. A, an SDS-PAGE gel of a nickel-affinity column from a 3L culture of MSP1D1 cell debris pellet. The fractions collected are enclosed in a red box. FT2, second flow-through; W1–6, please see the Materials and Methods section for the buffer composition, 10 mL each; E1–6, elute with 200 mM imidazole and 2 mL each; the first lane is Broad Range protein marker in kDa from Bio-Rad; B, a SDS-PAGE gel after TEV cleavage and a reverse nickel-affinity column. The fractions collected are enclosed in a red box. TEV protease, alone; before, the sample before TEV protease was added; 4 hr, 4 hours after the addition of TEV protease; FT2, second flow-through containing proteins without a His-tag; W1–4, wash with 10 mM imidazole, 5 mL each; E1–2, elute with 200 mM imidazole, 10 mL each; C&E, MSP1E3D1, and MSP2N2, respectively, in the same nickel-affinity purification conditions as in A; D&F, MSP1E3D1, and MSP2N2, respectively, in the same cleavage and reverse nickel-affinity purification as in B.

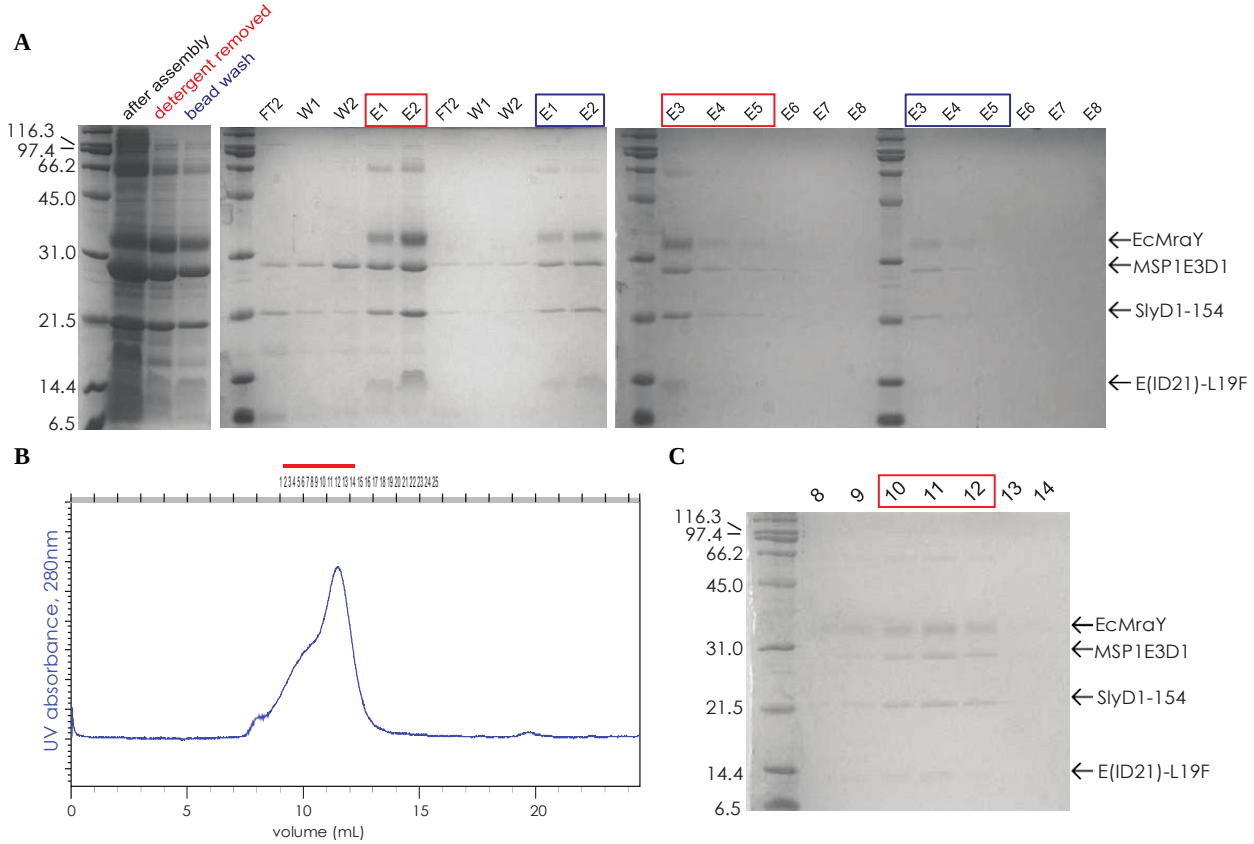

Figure 4.10. Reconstitution of EYS21-L19F into nanodiscs composed of MSP1E3D1 and DMPC. A, an SDS-PAGE gel of a nickel-affinity column after the nanodisc assembly of EYS21-L19F in MSP1E3D1 and DMPC (DMPC : MSP1E3D1 : EYS21-L19F = 40 : 2 : 1). After the assembly, Bio-Beads (Bio-Rad) were added to remove any detergent. Beads were filtered out from the sample and it was incubated with Ni-NTA resin for a nickel-affinity column (fractions collected enclosed in red boxes). Then, beads were washed with MSP buffer and the remaining sample was incubated with Ni-NTA resin separately for another nickel-affinity column (fractions collected enclosed in blue boxes). FT2, second flow-through; W1–2, wash with 20 mM imidazole, E1–8, elute with 200 mM imidazole; B, a gel-filtration (Superdex 200 10/300) chromatogram showing a peak with two left shoulders. C, fractions 8–14 from the gel-filtration column were run on an SDS-PAGE gel. All three components of the protein complex as well as MSP1E3D1 appeared. Fractions in a red box were collected and concentrated.

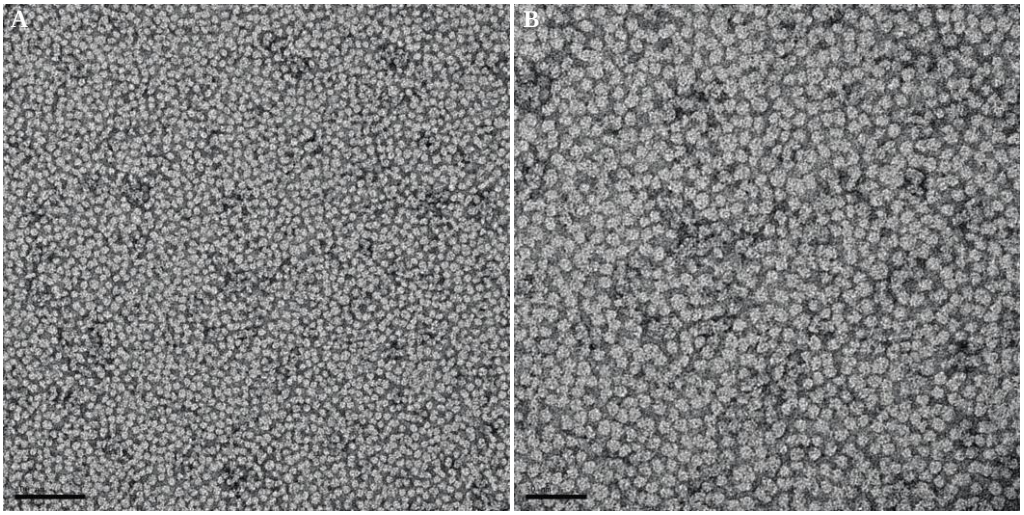

Figure 4.11. Negative-stained images of the EYS21-L19F-nanodisc (DMPC, MSP1E3D1) assembly. The EYS21-L19F-nanodisc assembly was added to a 200-mesh copper grid covered with formvar/carbon film (#01801, Ted Pella, Inc.) followed by 2 % uranyl acetate. The scale bar is 100 nm in A and 50 nm in B.

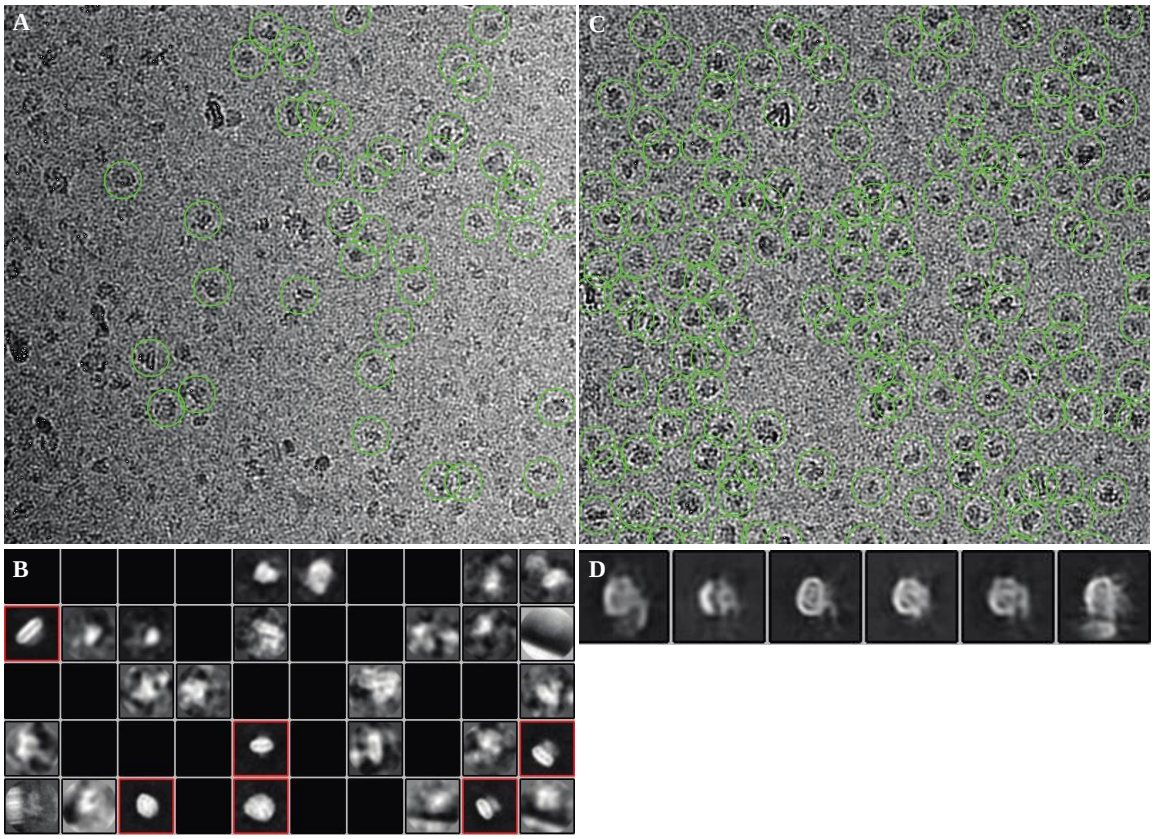

Figure 4.12. Data analysis of cryo-EM images of the EYS21-L19F-nanodisc assembly using Relion-3. A, an example of a cryo-EM micrograph at a 19,965 Å defocus with particles manually picked (green circles) to make a template for 2D classification. B, 2D classification based on the template created from 1,089 particles in 89 selected micrograms. Six classes were selected (red boxes) for the next steps of data processing. C, based on the 2D average, particles from all of the micrograms were auto-picked (green circles). D, 3D classification from the auto-picked particles.

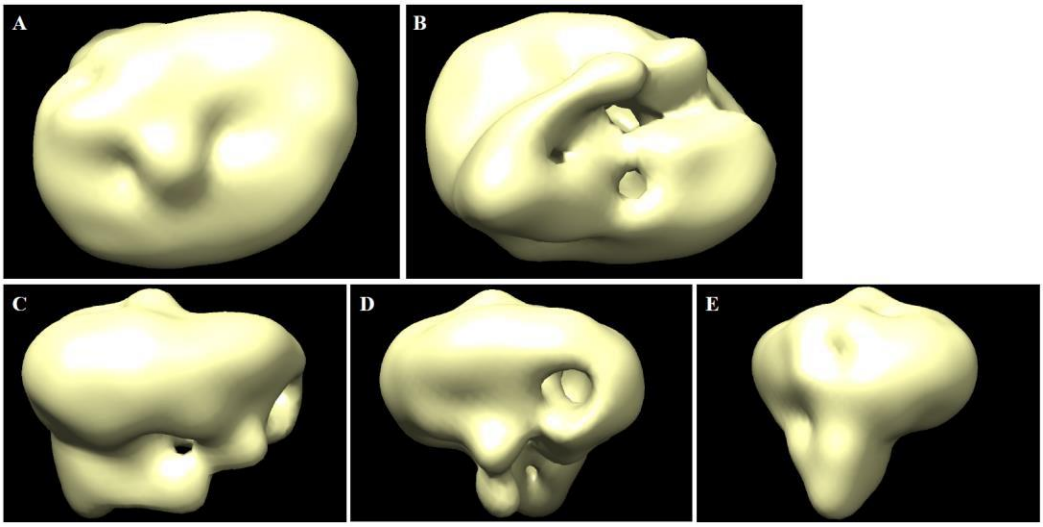

Figure 4.13. A 3D initial model of the EYS21-L19F-nanodisc assembly using Relion-3. This model was generated from the fourth class (24,323 particles) in **Fig. 4.12.D** and visualized with volume level 0.0741 in Chimera. A, a view from the periplasmic side. B, a view from the cytoplasmic side. C-E, side views.

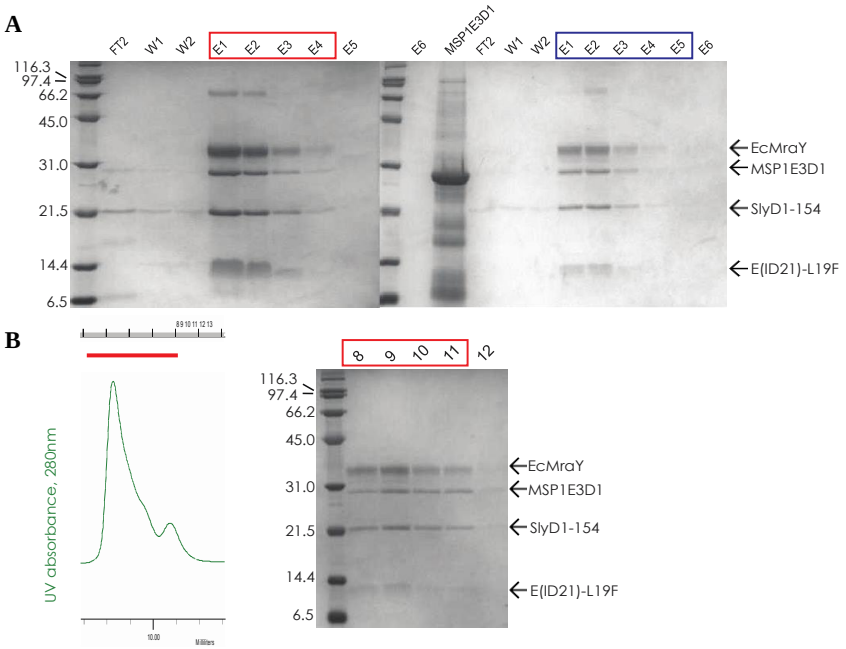

Figure 4.14. Reconstitution of EYS21-L19F into nanodiscs composed of MSP1E3D1 and POPC/POPG. A, an SDS- PAGE gel of a nickel-affinity column after the nanodisc assembly of EYS21-L19F along with MSP1E3D1 and POPC/POPG (POPC/POPG : MSP1E3D1 : EYS21-L19F = 80 : 2 : 1). After the assembly, Bio-Beads (Bio-Rad) were added to remove any detergent. Beads were filtered out from the sample and it was incubated with Ni-NTA resin for a nickel-affinity column (fractions collected enclosed in red boxes). Then, beads were washed with MSP buffer and the remaining sample was incubated with Ni-NTA resin separately for another nickel-affinity column (fractions collected enclosed in blue boxes). FT2, second flow-through; W1–2, wash with 20 mM imidazole, E1–6, elute with 200 mM imidazole; B, a gel-filtration (Superdex 200 10/300) chromatogram showing a peak with a shoulder on the right side, followed by another small peak. C, fractions 8–12 from the gel-filtration column was run on an SDS-PAGE gel. All three components of the protein complex as well as MSP1E3D1 appeared. Fractions in a red box were collected and concentrated.

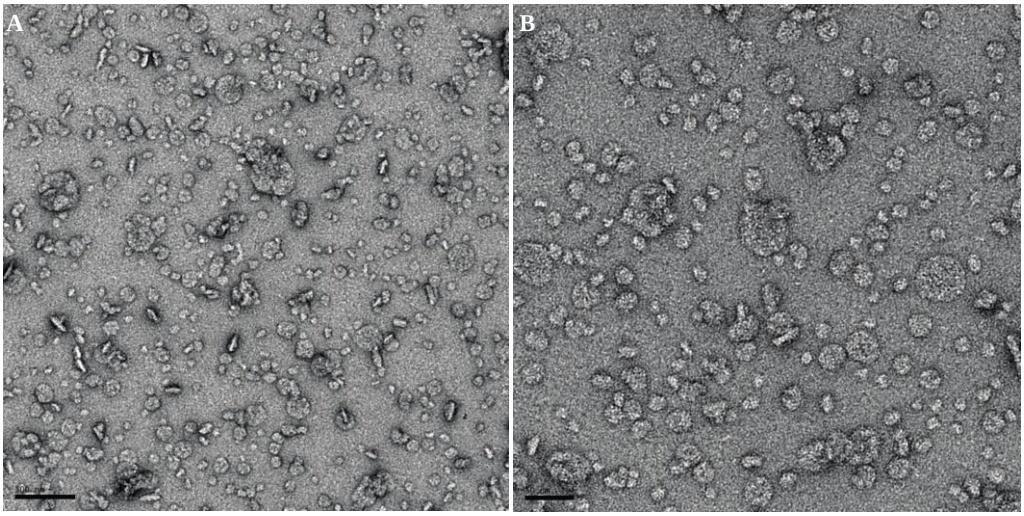

Figure 4.15. Negative-stained images of the EYS21-L19F-nanodisc (POPC/POPG, MSP1E3D1) assembly. The EYS21-L19F-nanodisc assembly was added to a 300-mesh copper grid (#2130C-XA, SPI Supplies) with a carbon layer, followed by 2 % uranyl acetate. The scale bar is 100 nm in A and 50 nm in B.

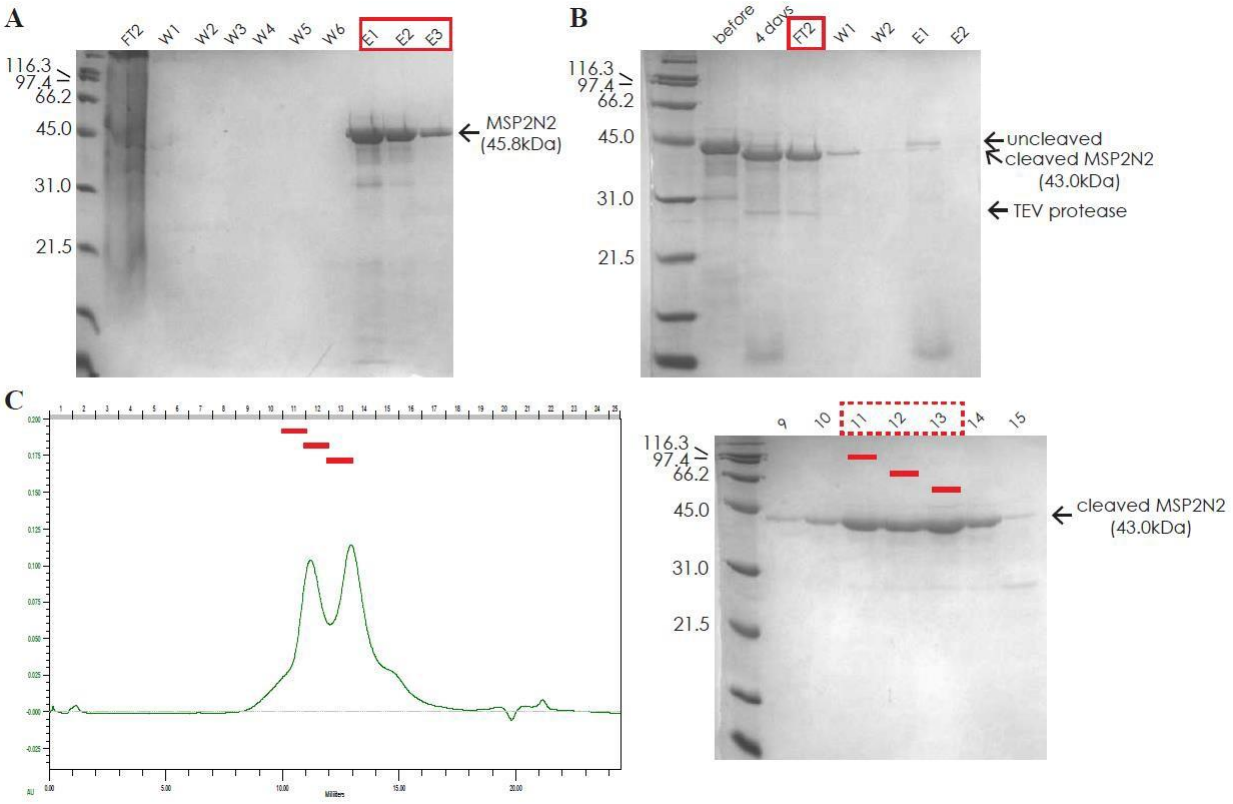

Figure 4.16. Another purification batch of MSP2N2 from *E. coli* cell debris. A, an SDS-PAGE gel from a nickel-affinity column for a 6L culture of MSP2N2 from cell debris pellet. The fractions collected are enclosed in a red box. FT2, second flow-through; W1–6, please see the Materials and Methods section for the buffer composition, 50 mL each; E1–6, elute with 200 mM imidazole and 10 mL each; the first lane is Broad Range protein marker in kDa from Bio-Rad;
B, an SDS-PAGE gel after TEV cleavage and a reverse nickel-affinity column. The fraction collected is enclosed in a red box. before, the sample before TEV protease was added; 4 days, 4 days after the addition of TEV protease; FT2, second flow-through containing cleaved proteins; W1–2, wash with 20 mM imidazole, 10 mL each; E1–2, elute with 200 mM imidazole, 10 mL each. C, left, a gel-filtration (Superdex 200 10/300) chromatogram with red bars indicating which area of the peaks corresponds to fractions 11, 12, and 13. The delay volume was 0.345 mL. Right, an SDS-PAGE gel with fractions 9–15 from the gel-filtration chromatogram. Individual fractions of 11, 12, and 13 were proceeded to the nanodisc assembly step.

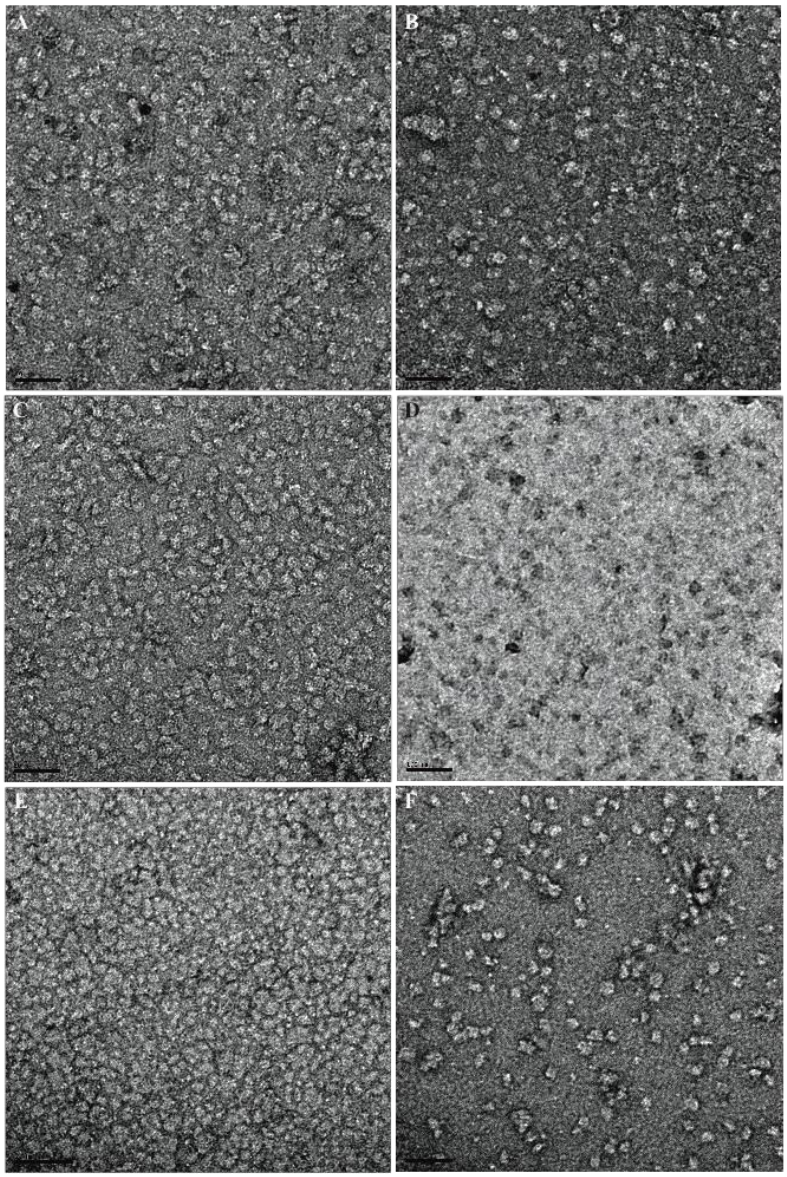

Figure 4.17. Negative-stained images of EYS21-L19F reconstituted in a variety of nanodiscs. The EYS21-L19F-nanodisc assembly was added to a 300-mesh copper grid (#2130C-XA, SPI Supplies) with a carbon layer, followed by 2 % uranyl acetate. Nanodiscs were composed of: (A) MSP2N2 gel-filtration fraction 11 and POPC, (B) MSP2N2 gel-filtration fraction 11 and POPG, (C) MSP2N2 gel-filtration fraction 12 and POPC, (D) MSP2N2 gel-filtration fraction 12 and POPG (Note that this image is hard to interpret. The grid used with this sample did not have enough area covered with the carbon layer.), (E) MSP2N2 gel-filtration fraction 13 and POPC, (F) MSP2N2 gel-filtration fraction 13 and POPG. Scale bars are 50 nm long for all panels, except for D with 100 nm.

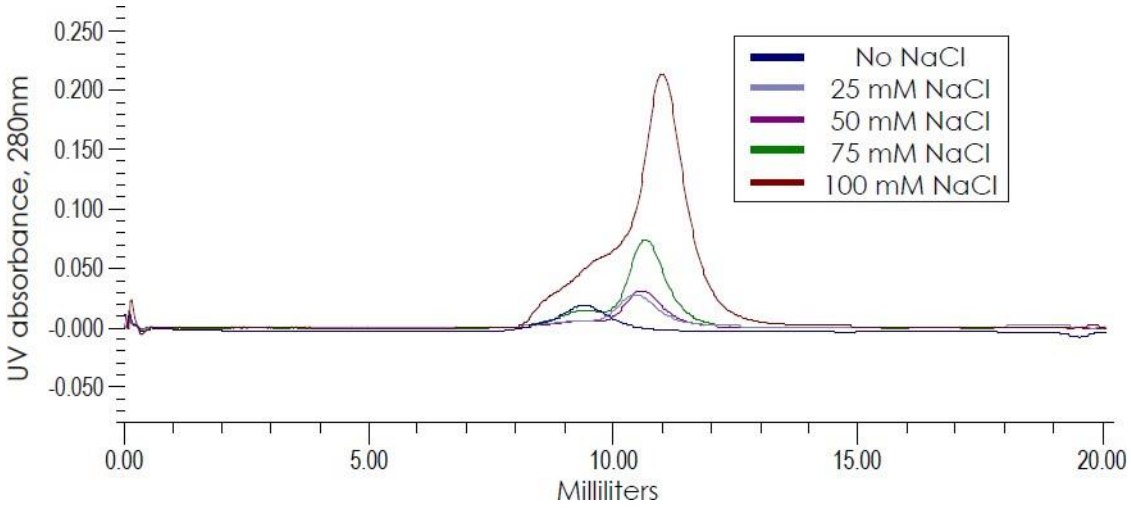

Figure 4.18. Overlay of gel-filtration chromatograms of EYS21-L19F purified at different NaCl concentrations. Protein sample ran with 0, 25, 50, 75 mM NaCl-containing buffer was from 3 L cell culture as opposed to the one with 100 mM NaCl from 12 L cell culture. Superdex 200 10/300 was used.

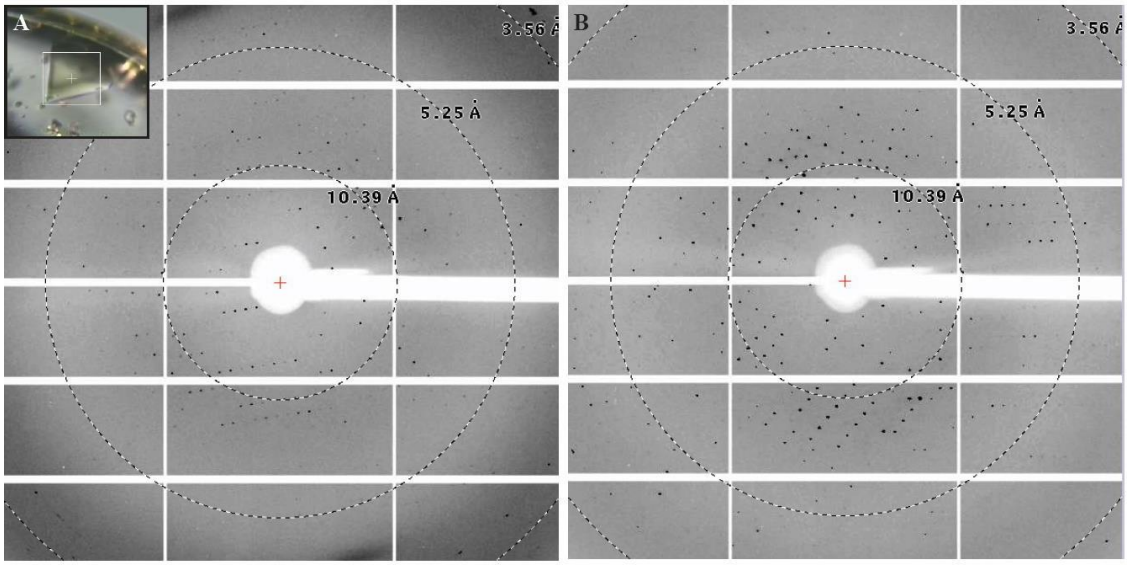

Figure 4.19. The 4.20 Å X-ray diffraction data collected on EYS21-L19F co-crystallized with the S-analog at the SSRL BL12-2. The final concentration of the protein and the analog in the drop was 4 mg/ml and 500 µM, respectively. The crystallization condition was 0.1 M MES pH 6.6, 0.37 M potassium nitrate, and 21.5% PEG 400. A crystal was harvested from the drop that is shown in the top left corner of A. A and B are diffraction images that were acquired from exposing the crystal 90° apart. The resolution rings were drawn in using Adxv. Note. X-ray screening on May 3rd, 2019; the cassette number was 203; the port number was G6.

Data collection	
Synchrotron	SSRL
Beamline	BL12-2
Space group	C1 2 1
Cell dimensions	
a/b/c (Å)	247.79/143.54/188.83
α/β/γ (°)	90/116.07/90
Wavelength (Å)	1.0000
Resolution (Å)	39.45 – 4.20
R_{merge} (%)	10.4 (73.3)
R_{meas} (%)	12.4 (86.3)
R_{pim} (%)	6.6 (45.0)
$CC_{1/2}$	95 (78.8)
$\langle I/\sigma(I) \rangle$	6.9 (1.9)
Completeness (%)	98.8 (99.5)
No. of observations	151168 (16202)
No. of unique	42915 (4487)
Multiplicity	3.5 (3.6)
Average mosaicity	0.09

Table 4.1. Statistics of X-ray data collection of EYS21-L19F co-crystallized with the S-analog at 4.20 Å. Values in parenthesis are from the highest resolution shell.

MATERIALS AND METHODS

Expression and purification of the EYS21-L19F complex

Two plasmids that were used are pRSFDuet-E(ID21)(L19F)-His_6-*Ec*MraY ($kanamycin^R$) and pET22b-SlyD1-154 ($carbenicillin^R$). They were co-transformed into the BL21(DE3)ΔSlyD competent cells. The colonies were scraped off into ~150 mL lysogeny broth media with kanamycin (35 µg/ml) and carbenicillin (35 µg/ml) and shaken for a few hours at 37°C. Each 1 L 2x YT media (16 g/L tryptone, 10 g/L yeast extract, and 5 g/L NaCl) was inoculated with ~10 mL of the starter culture and the cells grew at 37°C, shaking until optical density at 600 nm (OD_{600}) reaches 0.6-0.7. The culture flasks were transferred to 22°C and induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) when OD_{600} reaches 0.9–1.0 (to prevent lysis effect of E expression). The cells were grown overnight (16–20 hrs) and harvested using a JLA-8.1 rotor at 4,000 rpm for 10 min. Cell pellet was resuspended in lysis buffer (20 mM Tris pH 7.5, 300 mM NaCl, 10 % glycerol, 5 mM βME, and protease inhibitors (*e.g.*, PMSF, benzamidine)) and homogenized using a douncer. The cells were lysed by flowing through a microfluidizer four times. The lysate was spun down in a JLA-16.250 rotor at 12,000 rpm for 30 min. Pellets were discarded and the supernatant was spun down in a Ti-45 rotor at 45,000 rpm for 30 min or 30,000 rpm for 1 hr. The membrane pellet was either stored at -80°C or resuspended in extraction buffer (10 mM HEPES pH 7.5, 300 mM NaCl, 5 % glycerol, 5 mM βME, 10 mM imidazole, 1 % DM, and protease inhibitors). It was incubated for 2 hours at 4°C and spun down in a Ti-50.2 rotor at 45,000 rpm for 30 min. The pellet was discarded and the supernatant was incubated with 2.5 mL Ni-NTA resin (for 6L culture) overnight at 4°C. The sample was flown through a

gravity column twice and washed with 10 CV of a buffer (10 mM HEPES pH 7.5, 150 mM NaCl, 5 % glycerol, 0.03 % DDM, 5 mM βME), containing 20 mM imidazole, 10 CV of the same buffer with 20 mM EDTA, and eluted with the buffer containing 200 mM EDTA. SDS-PAGE gel was used to validate the eluted sample. The eluate was dialyzed into Buffer A (10 mM HEPES pH 7.5, 100 mM NaCl, 5 % glycerol, 5 mM βME, 0.03 % DDM) overnight at 4°C and ran on an anion-exchange column (Uno Q6) in a gradient from 100 % Buffer A to 100 % Buffer B containing 500 mM NaCl. Purity of the sample was checked by SDS-PAGE and Coomassie staining and the fractions were run on a gel-filtration column (Superdex 200 16/600) with 10 mM HEPES pH 7.5, 100 mM NaCl, 5 % glycerol, 5 mM βME, and 0.03% DDM. The purity and homogeneity were checked again by SDS-PAGE and the corresponding fractions were collected and used for structural studies right away or frozen at -80°C.

Co-crystallization of EYS21-L19F with the S-analog.

Initial crystallization screening started with adding 200 μM substrate analog to 12 mg/ml purified EYS21-L19F. The mixture was then filtered using an Amicon 0.22 μm filter by spinning at 8,000 rpm for 5 min at 4°C. For a standard sitting-drop vapor-diffusion method, the sample (0.2 μL) was dispensed into individual wells on Swissci 96-well MRC plates (Molecular Dimensions), followed by well solution (0.2 μL), which was drawn from a 50 μL reservoir, using a Mosquito (TTP Labtech). Some of the commercial sparse matrices used were MemGold (Molecular Dimensions), MemGold2 (Molecular Dimensions), and Morpheus (Molecular Dimensions).

The crystallization conditions that gave initial crystal hits were categorized by the buffer, pH, salt, and a precipitant reagent. A grid screen was designed by varying pH and precipitant concentration around the condition in which crystals were observed, as salt concentration was kept constant. For setting up grid screens, I switched from the sitting-drop to the hanging-drop method and increased the drop size by adding 1 µL of protein sample that contains the substrate analog and 1 µL of the well solution. As optimization continued, the range of pH and precipitant concentration became finer and the volume ratio of protein sample to condition that go into drop varied along with the whole drop size.

Initial crystallization screening for the LCP method started with mixing 12 mg/ml or 24 mg/ml purified EYS21-L19F in addition to 200 µM substrate analog and monoolein (Sigma) at a 2:3 volume-to-volume ratio using a syringe lipid mixer. The protein/lipid mixture was dispensed on a glass plate using a Gryphon robot (Art Robbins Instruments). The volume ratios of the sample to crystallization condition used were 0.045 µL to 0.5 µL, 0.045 µL to 1 µL, 0.2 µL to 0.8 µL, and 0.2 µL to 1 µL. One commercial sparse matrix, MemMeso (Molecular Dimensions), was used. The plates were stored at room temperature.

Detergent screening for extraction of EYS21-L19F.

The membrane pellet was prepared as above and resuspended into 10 mM HEPES pH 7.5, 300 mM NaCl, 5 % glycerol, 5 mM βME, protease inhibitors, 10 mM imidazole, and

1 % of each of six detergents (DDM, DM, Cymal-5, β-OG, LDAO, and Fos-choline-12). After 2 hrs of rocking at 4°C, each of the samples were spun in a bench-top ultracentrifuge MLA-80 rotor at 30,000 rpm for 1 hr. Each supernatant was incubated with 250 µL of Ni-NTA resin rocking overnight at 4°C then passed through a gravity column. The samples were washed with 40 CV of 10 mM HEPES pH 7.5, 300 mM NaCl, 5 % glycerol, 5 mM βME, 20 mM imidazole, and 2.5x CMC of a detergent of choice and eluted with 40 CV of the buffer with all the same components, except for 200 mM imidazole. Purity of the fractions were checked by SDS-PAGE and proceeded with a gel-filtration column (Superdex 200 10/300) with 10 mM HEPES pH 7.5, 150 mM NaCl, 5 % glycerol, 5 mM βME, and 2.5x CMC of a detergent of choice. The fractions from peaks were chosen based on their purity by SDS-PAGE and concentrated using an Amicon 100 kDa MWCO concentrator. Varying amounts of white precipitates were observed for the DDM- and β-OG-solubilized sample during concentration.

Reconstitution of the protein complex into a non-detergent polymer, amphipols

Purified EYS21-L19F was incubated with amphipathic polymers (amphipols or Apols) in the ratio of 1: 2.9 (3.2 mg of protein and 9.6 mg of Apols) for 4 hours at 4°C. 69 mg of Bio-Beads (About 500 times excess amount of detergent DDM in the buffer) washed with 1 mL methanol twice and 1 mL milliQ water five times were added to the sample and incubated overnight at 4°C. An extra 100 µL of the buffer without DDM was added to 400 µL of the sample with Bio-Beads. Beads were removed by filter and the flow-through was injected into a gel-filtration column (Superdex 200 10/300) and run with the buffer without DDM (10 mM HEPES pH 7.5, 100 mM NaCl, 5% glycerol, 5 mM βME in

the presence or absence of 10 mM $MgCl_2$). In the presence of $MgCl_2$ in the running buffer, a shoulder appeared on the left side of the peak in the chromatogram (data not shown here). The fractions without $MgCl_2$ (1.14 mg/ml) were incubated with 210 µM S-analog and used to prepare a negative-stained grid.

Optimized of purification of the EYS21-L19F complex

The overall purification process was similar to the one above with some modifications. Extraction was performed with 1 % DDM instead of DM after the detergent screening. Cobalt resin was used instead of Ni-NTA and the sample was washed with 30 CV of 10 mM HEPES pH 7.5, 300 mM NaCl, 5 % glycerol, 5 mM βME, 0.03 % DDM, and 20 mM imidazole, and eluted with the buffer containing 100 mM NaCl and 200 mM imidazole.

Expression and purification of membrane scaffold proteins (MSPs)

pMSP1D1 (Addgene plasmid # 20061), pMSP1E3D1 (Addgene plasmid # 20066), and pMSP2N2 (Addgene plasmid # 29520) were gifts from Stephen Sligar. All three plasmids had a His_7 tag followed by a TEV cleavage site at the N-terminus of MSP in pET28a and they were transformed into Nico21(DE3) cells. The colonies were scraped off, added to 150 mL of LB + kanamycin (35 µg/ml), and shaken at 37°C for a few hours. Culture flasks with 1 L 2xYT + kanamycin (35 µg/ml) were inoculated with about 10 mL of the starter culture. The cells were grown in a 225 rpm shaker at 37°C. When OD_{600} reached 0.6–0.8, the cells were induced with 1 mM IPTG and grown for four more hours.

Cells were resuspended and homogenized in 50 mM Tris pH 8.0, and 500 mM NaCl, 1 % Triton X-100, 1 mM EDTA, and 5 mM $MgCl_2$ using a douncer. Two pills of protease inhibitor cocktail (cOmplete™, Roche) were added and stirred for 30 min at 4°C. The cells were lysed as they ran through the microfluidizer for three passes at 15–18 kpsi. The cell lysate was spun in an ultracentrifuge Ti-45 rotor at 20,000 rpm for 45 min. The cell debris pellet was scraped off and set aside for separate purification later. The supernatant was incubated with Ni-NTA resin (2–2.5 mL resin was used for proteins from 6 L culture.) for 2 hours, stirring at 4°C. The sample with Ni-NTA resin was run through a gravity column. The bound proteins were first washed with 10 CV of 1 % Triton X- 100-containing Buffer A (50 mM Tris pH 8.0 and 500 mM NaCl), followed by 50 mM cholate-containing Buffer A, 20 mM imidazole-containing Buffer A, and 50 mM imidazole-containing Buffer A. The proteins were eluted with 10 mL of 200 mM imidazole-containing Buffer A and collected in 2mL fractions. Purity of all of the fractions from the nickel-affinity column was checked by SDS-PAGE and Coomassie stain.

The cell debris pellet was resuspended and homogenized in 6 M Guanidine HCl, 50 mM Tris pH 8.0, and 500 mM NaCl using a douncer. This resuspension was spun in an ultracentrifuge Ti-45 rotor at 20,000 rpm for 45 min. The supernatant was incubated with Ni-NTA resin (2–2.5 mL resin was used for cell debris from 6L culture.) for 2 hours, rocking at 4°C. The sample was added to a gravity column. The bound denatured proteins were first washed with 10 CV of 6 M GuHCl-containing Buffer A (W1), followed by 10 CV of Buffer A (W2), 1 % Triton X-100-containing Buffer A (W3), 50 mM cholate-

containing Buffer A (W4), Buffer A itself (W5), and 20 mM imidazole-containing Buffer A. The refolded proteins in the column was eluted with 10 mL of 200 mM imidazole-containing Buffer A and collected in 2 mL fractions. Purity of all of the fractions from the nickel-affinity column were checked by SDS-PAGE and Coomassie staining.

The samples from soluble and insoluble portions of lysate were kept separated throughout the purification process. The fractions with proteins were dialyzed in 50 mM Tris pH 8.0, 20 mM NaCl, and 0.5 mM EDTA overnight at 4°C to remove imidazole for subsequent TEV cleavage. Some of the dialyzed sample was saved for SDS-PAGE. 0.5–1 mg of TEV protease (expressed and purified in the lab) along with 1 mM DTT were added to each of the dialyzed samples. Some of the sample after 4 hours was saved and TEV cleavage continued overnight while rocking at 4°C. All DTT was dialyzed out overnight at 4°C in 20 mM Tris pH 8.0 and 500 mM NaCl. Then, the sample was passed over nickel-affinity column to remove uncleaved sample and then with four 5-mL washes with 20 mM Tris pH 8.0, 100 mM NaCl, and 10 mM imidazole were collected followed by ten 2-mL elutions with 200 mM imidazole collected as fractions. After SDS-PAGE to ensure cleavage, cleaved MSPs were separated from TEV protease and other contaminants. Final dialysis was done with 20 mM Tris pH 7.5, 100 mM NaCl, and 0.5 mM EDTA.

Reconstitution of the protein complex into nanodiscs.

MSP buffer (20 mM Tris pH 7.5, 100 mM NaCl, 0.5 mM EDTA) and MSP buffer with 100 mM cholate were first prepared. 35.1 mg of DMPC, 18.1 mg/17.2 mg of

DMPC/DMPG, 21.5 mg/19.3 mg of POPC/POPG were each solubilized in 1 mL of the cholate-containing buffer, making 50 mM lipid stocks. Then, all the components were mixed, so that the ratio of the protein complex : MSP : lipid becomes 1 : 2 : 80 (40 also tried) and the final concentration of cholate is 20 mM. Incubation of this mixture for an hour at room temperature for DMPC and DMPC/DMPG-containing assembly, whereas, with POPC/POPG, incubation was at 4°C. About 250 mg of Bio-Beads (Bio-Rad) were added into each nanodisc assembly sample and was incubated at the same temperature at least for an hour and up to four hours. Bio-Beads were filtered out and the sample was saved. The Bio-Beads were washed with extra MSP buffer and the washed sample was also saved. After incubation with 250 µL Ni-NTA resin overnight at 4°C, all of the samples were run on nickel-affinity columns, washed with 4 mL 20 mM imidazole-containing MSP buffer and eluted with 6 mL 200 mM imidazole-containing MSP buffer. After checking their purity with SDS-PAGE, certain fractions were selected for a gel-filtration column (Superdex 200 10/300), with 20 mM Tris pH 7.5, 100 mM NaCl, and 0.5 mM EDTA.

Preparation of negative-stained grids of the EYS21-L19F complex

2 µL of purified EYS21-L19F (0.4 mg/ml) was pipetted onto either 300-mesh or 400-mesh copper grids that were carbon coated using a Cressington 208carbon and glow discharged using an Emitech K100X (15 mA, 1 min). After 50 sec, the excess protein sample on the grid was wicked away using Whatman No. 1 filter paper. Then, 2 µL of 2 % uranyl acetate was pipetted on top of the grid and again, after 50 sec, the excess blotted away using the edge of the filter paper.

Imaging a negative-stained grid of the EYS21-L19F complex using a 120 keV transmission electron microscope (TEM)

A Tecnai T12 equipped with an electron gun, LaB_6, and Gatan Ultrascan 2k x 2k CCD was used to image the negative-stained grid of the protein complex. The SA range magnification, 26,500- 52,000x, was used to visualize particles.

Cryo-EM grid preparation, data acquisition, and image processing

A copper Quantifoil R2/2 grid was glow discharged at 15 mA for 1 min using Pelco easiGlow. Using a FEI Vitrobot (Mark v4), 3 µL of the protein sample (0.4 mg/ml) reconstituted in nanodisc (DMPC/MSP1E3D1) was applied to the grid and blotted with Whatman No. 1 filter paper with the blot force -5 for 2 sec at 4°C and with 100 % humidity. Immediately after blotting, the grid was plunge frozen in liquid ethane cooled by liquid nitrogen. The grid was first screened using Talos Arctica equipped with a FEG operating at 200 keV and the 4k x 4k FEI Falcon III direct electron detector.

Movies were collected using a Titan Krios that operates at 300 keV, equipped with the Gatan K3 4k x 4k direct electron detector in counting mode (0.834 Å/pixel) at a magnification of 165,000x using a defocus range of -0.8 to -2.7 µm. Movies were collected over an 8 second exposure (40 frames) with an exposure rate of 5.9 e^-/pixel/sec, resulting in total exposure of 1.7 $e^-/Å^2$/frame. All image processing was done using Relion-3 (Scheres, 2012; Zivanov et al., 2018). Motion correction was performed using the MotionCor2 frame alignment program (Chen et al., 2013). CTF determination was

performed using CTFFIND4 (Cardone, Heymann & Steven, 2013). A total of 1,089 particles were manually picked for 2D classification. Six 2D classes were selected for a template for automatic particle picking. Then, all the images were manually inspected to remove selection on things that were not particles. Using a 2D class average, 3D classification was performed, and 3D initial models were obtained from 23,121 ~ 45,285 particles/class.

ACKNOWLEDGEMENTS

We thank Andrey Malyutin, Songye Chen, and Alasdair McDowall for all the help with the use of electron microscopes in the Beckman Institute Resource Center for Transmission Electron Microscopy at Caltech. We are grateful to Jens Kaiser (Caltech) for crystallography help through the Molecular Observatory. Operations at SSRL are supported by the U.S. Department of Energy (DOE) and the National Institutes of Health (NIH). This research also used resources of the Advanced Photon Source, a U.S. DOE Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. This work was supported by MraY grant.

Chapter 5

Insight into the catalytic mechanism of the bacterial phosphotransferase MraY

ABSTRACT

The peptidoglycan is a protective layer conserved in nearly all bacteria, but it is not present in eukaryotic cells. Enzymes involved in the biosynthesis pathway of peptidoglycan have been attractive targets for designing antibacterial agents for decades. One of the most promising targets (Silver, 2013) is the first membrane-bound enzyme in the pathway, called MraY. Despite structural and biochemical information available for MraY, more investigation to elucidate the catalytic mechanism and binding modes of native substrates awaits. In this study, a dUMP exchange reaction supports the model that MraY from *Hydrogenivirga sp*. undergoes a one-step mechanism, in which both of the substrates, UDP-MurNAc-pentapeptide and C_{55}-P, are required to form a non-covalent ternary complex to produce Lipid I and UMP. Based on the result of fluorescence-based continuous activity assay of MraY, we suggest MraY undergoes, more specifically, an ordered Bi–Bi mechanism. Furthermore, we performed extensive mutagenesis on invariant aspartate, lysine, and histidine residues located in the putative active site of MraY. The comparison of activity of mutants to their binding to UDP-MurNAc-pentapeptide suggests that D117 is important in binding and D196 is critical for catalysis. Dependence of MraY activity on the concentration of Mg^{2+} and pH were also investigated. In total, this work provides important insight into the mechanism of an important antibiotic target.

INTRODUCTION

The majority of bacteria contain the peptidoglycan (PG) layer as a major constituent of their cell wall, which protects the cells from internal turgor pressure and helps them to maintain their cell shape (Lovering et al., 2012). In addition, a functional peptidoglycan layer is required for effective cell division (Vollmer et al., 2008). Because the peptidoglycan layer is unique to bacteria, many enzymes involved in this pathway have been the site of action of antibacterial agents, including clinically important β-lactam antibiotics (Bugg, 1999). All of the enzymes in the peptidoglycan biosynthesis pathway have been structurally characterized in one species or another. However, mechanistic understanding of some of them lags behind largely due to the absence of substrate(s)-bound structures. One of them is MraY, the first integral membrane protein that catalyzes the reaction of making the first lipid-bound building block (Lipid I) in the PG biosynthesis pathway. Apo and inhibitor-bound MraY crystal structures are currently available (Chung et al., 2016, 2013; Hakulinen et al., 2017), but they cannot clearly resolve the binding mode of the two substrates, UDP-MurNAc-pentapeptide and undecaprenyl phosphate (C_{55}-P).

Decades of efforts by many groups were devoted to elucidating the mechanism of MraY using biochemical tools. In 1969, Heydanek et al. first proposed a two-step mechanism of MraY, which involves forming a covalent MurNAc-pentapeptide-phosphoenzyme intermediate and the subsequent release of the enzyme by a nucleophilic attack of C_{55}-P (Heydanek, Struve & Neuhaus, 1969). In 2004, Bouhss et al. first reported that MraY can be purified to homogeneity and its catalytic reaction requires Mg^{2+} (Bouhss,

Crouvoisier, Blanot & Mengin-Lecreulx, 2004). With pure MraY in hand, Al-Dabbagh et al. proposed an alternative one-step mechanism, which requires the formation of a ternary complex of the enzyme and two substrates for the reaction to occur (Al-Dabbagh et al., 2008). More recently, exchange reaction results of MraY from *Bacillus subtilis* using isotopes were in agreement with the one-step mechanism (Al-Dabbagh et al., 2016; Yao Liu et al., 2016).

Here, we propose MraY from *Hydrogenivirga sp*. also undergoes the one-step mechanism based on the evidence from a direct exchange reaction. We fit a continuous activity assay result into bisubstrate reaction models and the double-reciprocal plot indicates an ordered bisubstrate rather than ping pong bisubstrate mechanism. In addition, we discuss, via extensive mutagenesis and activity studies, which of the conserved residues in the active site may have potential roles in catalysis. Lastly, residues critical to substrate binding were determined by fluorescence anisotropy.

RESULTS

dUMP exchange reaction indicates both substrates are required concomitantly for the MraY catalytic reaction

As depicted in **Figure 5.1.A**, both one-step and two-step mechanisms have been proposed for the catalytic reaction of MraY. In a one-step mechanism, the MraY enzyme along with its two substrates, UDP-MurNAc-pentapeptide and C_{55}-P, are all required to form a ternary complex for catalysis, resulting in the formation of the two products, Lipid I and UMP. More specifically, it is currently believed that an invariant active site residue on MraY deprotonates a hydroxyl group of the phosphate group on C_{55}-P to make an oxyanion, which acts as a nucleophile to attack the β-phosphate group of UDP-MurNAc-pentapeptide to form both of the products. On the other hand, in a two-step mechanism, UDP-MurNAc-pentapeptide is first needed to form a covalent intermediate with the enzyme and forms UMP as the first product. This is followed by the nucleophilic attack of an oxyanion of C_{55}-P on the intermediate to produce Lipid I.

In order to prove which catalytic path MraY takes, we designed and performed an experiment that would prove the substrate requirements for the reaction. We first heterologously overexpressed MraY from a thermophilic species, *Hydrogenivirga sp*. in *Escherichia coli* and purified it with extensive washing steps in a cation column to completely remove endogenous C_{55}-P. Purified *Hy*MraY was then incubated with UDP-MurNAc-pentapeptide, dUMP, and $MgCl_2$ in the presence and absence of C_{55}-P at room temperature. The reactions were stopped by boiling, butanol was added to extract the hydrophobic substrate/product, and the aqueous layer was isolated and subject for LC-MS analysis. The rationale behind this

experiment is that when a reaction has all substrate(s) required for catalysis, the reaction will reach equilibrium, with the reverse reaction incorporating dUMP into UDP-MurNAc-pentapeptide. In other words, the UMP leaving group from the forward reaction is replaced by dUMP in the reverse reaction to generate dUDP-MurNAc-pentapeptide. We first confirmed that dUMP acts as a product inhibitor to a similar extent as UMP supporting that it behaves similarly in the reaction (**Fig. 5.1.B**).

At equilibrium, in the presence of excess dUMP, LC-MS analysis of the aqueous layer revealed that both substrates were required for dUMP incorporation into dUDP-MurNAc-pentapeptide (**Fig. 5.1.C**). In the first four reactions **Figure 5.1.C**, background incorporation was seen suggesting contaminating C_{55}-P; therefore, it was critical to remove endogenous C_{55}-P by extensive washing (reactions boxed in yellow).

HyMraY uses an ordered Bi–Bi mechanism.

In order to measure the activity of *Hy*MraY, UDP-MurNAc-pentapeptide was first treated with dansyl-chloride to attach a dansyl group (DNS) to N^{ε} of m-DAP, the third amino acid of the pentapeptide which had been shown to not affect catalysis (Stachyra, Dini, Ferrari, Bouhss, Van Heijenoort, et al., 2004). Each reaction mixture contained *Hy*MraY, UDP-MurNAc-pentapeptide-DNS, C_{55}-P, and $MgCl_2$ in the buffer with DM. Reactions were incubated at 37°C for 30 min, diluted, and the fluorescence was measured across 380–700 nm (**Fig. 5.2.A left**). Alternatively, butanol (BuOH) was incubated with the reaction sample and the BuOH layer was pipetted from the aqueous layer to obtain only the hydrophobic

product (**Fig. 5.2.A right**). In this reaction, as seen previously a simplified substrate C_{20}-P could be substituted for C_{55}-P (**Fig. 5.2.A**) (Y. Zheng, Struck & Young, 2009).

A continuous fluorescence-based assay was developed to measure the initial rate of reaction in the linear region of the Michaelis-Menten curve. The reaction mixture with purified *Hy*MraY, dansylated UDP-MurNAc-pentapeptide, and C_{55}-P was initiated with the addition of $MgCl_2$ in a plate reader and fluorescence from dansylated Lipid I was monitored for the first eight min. For one set of experiments, while the concentration of UDP-MurNAc-pentapeptide-DNS varied, the concentration of C_{55}-P was kept constant to be 7.5 µM. Two more sets of experiments were performed for 20 µM and 50 µM C_{55}-P. The concentrations of UDP-MurNAc-pentapeptide-DNS and its corresponding initial velocity (v_0) were subsequently used to draw a double-reciprocal plot with $1/v_0$ vs. 1/[UDP-MurNAc-pentapeptide-DNS]. In the double-reciprocal plot, there is a crossover point in the second quadrant met by three lines, each formed when concentration of C_{55}-P was kept constant (**Fig. 5.2.B**). According to Leskovac, this indicates the enzyme going through ordered Bi–Bi mechanism (Leskovac, 2003) rather than ping pong Bi–Bi mechanism.

Mutagenesis studies of HyMraY to identify residues critical for catalysis.

Al-Dabbagh et al. first proposed one-step mechanism for MraY from *B. subtilis* along with the invariant D98 playing a role in deprotonating the phosphate group of C_{55}-P (Al-Dabbagh et al., 2008). Based on the *in vivo* complementation assay, they showed 14 mutants including D98N are essential for activity (Al-Dabbagh et al., 2008). To confirm this analysis, we generated mutants of *Hy*MraY that correspond to the residue of D98 (*B. subtilis*) in *Hy*MraY

(D117) to either alanine (A), asparagine (N), or glutamate (E) (**Fig. 5.3.AB**). The level of expression and protein yield of the mutants after purification was similar to those of the wild type. All of the mutants D117A, D117N, and D117E were then subject to the activity assay using the quenched fluorescence-based method (**Fig. 5.2.A**). All showed significant to complete loss of activity. Compared to the control without Mg^{2+}, D117E showed some minimal activity (**Fig. 5.3.B**). This indicates that a longer carboxylic acid side chain could support catalysis, although significantly less efficiently.

In addition to D117, four other invariant aspartate residues located in the putative active site were mutated (**Fig. 5.3.A**). All four residues (D118, D193, D196, and D265) were each mutated to either asparagine or glutamate and they expressed and purified similarly to wild type. D118N had no activity, but D118E retained more than half of the wild-type activity (**Fig. 5.3.B**). This is similar to the D117 mutants, although the activity was less sensitive to the long Glu side chain. Neither the D265N or D265E mutations showed significant loss of activity (**Fig. 5.3.B**), suggesting that D265 is not as critical as D117 and D118. Amer and Valvano reported a potential catalytic role of the corresponding residues of D117 and D118 in *Hy*MraY (D90 and D91) and another invariant D156 in *Ec*WecA, a similar phosphotransferase to MraY, due to the failure in functional complementation *in vivo* and a large reduction in transferase activity *in vitro* (Amer & Valvano, 2002). Based on their multiple sequence alignment, D156 as well as D159 of *Ec*WecA are part of the conserved region, 150-NAFNMV**D**GI**D**GL-161, across WecA homologs (Amer & Valvano, 2002). The corresponding aspartate residues in *Hy*MraY are D193 and D196, which are completely conserved across MraY homologs, were again mutated to Asn or Glu (**Fig. 5.6**). For the case

of the D193 mutants, while the longer side chain of D193E lost activity, the amide side chain of D193N retained some activity (**Fig. 5.3.C**). Both D196N and D196E were completely dead (**Fig. 5.3.C**), and thus D196 is critical in activity of *Hy*MraY.

We further mutated K121 and K133 that are highly conserved across MraY orthologs and the side chains are facing toward the putative active site in the published *Aa*MraY (same numbering as *Hy*MraY) and *Cb*MraY (K97 and K111) structures. The initial motivation to investigate functional roles of these lysine residues was because Schaefer et al. proposed conserved arginine residues play a role as general bases to deprotonate a phosphate group of the lipid substrate in the LCP complex, which is a wall teichoic acid precursor transferase (Schaefer, Owens, Kahne & Walker, 2018). Also, lysine residues were proposed to have a role in positioning the phosphate group of C_{55}-P for its nucleophilic attack on UDP-GlcNAc in a human ortholog of MraY, DPAGT1, involved in the N-linked glycosylation pathway (Dong et al., 2018). Both K121A and K133A retained more than half of the wild-type activity, suggesting they are not catalytically essential (**Fig. 5.3.D**).

Among the members of *N*-acetylhexosamine-1-phosphate transferase superfamily, MraY sequences uniquely have a conserved 13-mer, 320-MA**P**I**HHH**FELKGW-332, in the loop between TM9 and TM10 (Anderson, Eveland & Price, 2000). Since the members are specific to a variety of soluble substrate with a nucleotide donor, this sequence analysis led us to consider that this loop is potentially involved in soluble substrate recognition. We mutated H324, H325, and H326 to alanine. We found that the activity of H326A was

comparable to that of the wild type and H325A retained about half of the wild-type activity, whereas H324A resulted in a loss of activity (**Fig. 5.3.EF**).

The activity of HyMraY is dependent on Mg^{2+} concentration and pH

The effects of magnesium and pH on the activity of *Hy*MraY and mutants were tested. Al-Dabbagh et al. reported the activity for the *Bs*MraY wild type was optimal at 30–60 mM Mg^{2+}, while activity was abolished at Mg^{2+} concentrations higher than 100 mM (Al-Dabbagh et al., 2008). They suggested some mutants of *Bs*MraY (H45R, D174N and D177N) became active at 250 mM Mg^{2+}, suggesting these residues are possibly involved in the binding of the metal ion. Here we assayed the wild type, D193N, D193E, D196N, and D196E of *Hy*MraY in the presence of increasing concentrations of $MgCl_2$ (0–250 mM) to investigate whether the lack of activity can be attributed to impaired Mg^{2+} binding. However, no activity could be recovered for D193E, D196N, or D196E mutants and the wild-type *Hy*MraY activity was optimal at 150 mM Mg^{2+} (**Fig. 5.4.A**). Furthermore, the effect of pH was examined in the range 6.5–9.8 for the wild type, D117N, and D118N. Al-Dabbagh et al. reported D98N had a distinct pH profile compared to the wild type, recovering maximal activity at pH 9.0–9.4 (Al-Dabbagh et al., 2008). Based on this result, the authors proposed that D98 is involved in deprotonating C_{55}-P, which is essential for activity and a one-step mechanism. However, no recovery in enzymatic activity was observed for the mutants in our assay (**Fig. 5.4.B**).

Altered activity of HyMraY mutants is correlated with change in binding of UDP-MurNAc-pentapeptide

To measure substrate binding to *Hy*MraY and various mutants, a fluorescence anisotropy assay was used taking advantage of changes in the rotational tumbling of a fluorescently labeled UDP-MurNAc-pentapeptide. Binding of UDP-MurNAc-pentapeptide-DNS to MraY leads to a decrease in rotational rate and an increase in fluorescence anisotropy. The fluorescent tag of substrate UDP-MurNAc-pentapeptide-DNS does not interfere with enzymatic reaction as previously demonstrated (Stachyra, Dini, Ferrari, Bouhss, van Heijenoort, et al., 2004). Additionally, changes in fluorescence anisotropy have been successfully applied to other membrane proteins (Gerber et al., 2013; Rinken, Lavogina & Kopanchuk, 2018).

The fluorescence anisotropy assay yielded a dissociation constant K_d of 5.2 μM for UDP-MurNAc-pentapeptide-DNS binding to *Hy*MraY that was dependent on Mg^{2+} (**Fig. 5.5.A**) and was lost in the presence EDTA after reaching saturation (data not shown). No significant change in anisotropy is observed in a control experiment with an unrelated membrane protein (TatC) suggesting that substrate UDP-MurNAc-pentapeptide-DNS binds to *Hy*MraY specifically (data not shown). The results support that C_{55}-P is not required for binding of UDP-MurNAc-pentapeptide-DNS into the active site of *Hy*MraY.

Detergent micelles could complicate anisotropy measurements; therefore, to prevent this *Hy*MraY and the variants were exchanged into amphipathic polymers (amphipols or Apols). This allows for solubilization of the protein in the absence of detergents (Tribet et al., 1996). When assayed using the water-soluble lipid-substrate analog, neryl phosphate (C_{10}-P),

*Hy*MraY in amphipols remained active at levels comparable to *Hy*MraY in detergent (data not shown). This validates further study of *Hy*MraY in amphipols.

Anisotropy experiments were performed with all mutants described above to determine the effect on binding of dansylated-UDP-MurNAc-pentapeptide. As before, *Hy*MraY mutants in amphipols were titrated into a solution of UDP-MurNAc-pentapeptide-DNS and fluorescence anisotropy was measured at 550 nm emission. The results are generally in agreement to enzymatic activity of the mutants (**Fig. 5.5**). The H326A mutant, which is enzymatically equipotent to the wild type, also exhibits equivalent binding of UDP-MurNAc-pentapeptide-DNS (**Fig. 5.5.B**), so, despite its conservation H326 does not play a direct role in substrate binding or catalysis. For H325A, a slight increase in anisotropy correlating with protein concentration was observed (**Fig. 5.5.B**) in agreement with the reduced enzymatic activity of this mutant. H324A showed no binding of UDP-MurNAc-pentapeptide-DNS (**Fig. 5.5.B**). It is likely that H324A plays a direct role in substrate recognition. The enzymatically inactive mutants D117A, D118A and D265A also exhibit an anisotropy increase at higher protein concentrations indicating that UDP-MurNAc-pentapeptide-DNS weakly binds (**Fig. 5.5.B**).

For the rest of the mutants designed to remove charge of the residue with minimal steric change, varying results were seen. The D265N mutation binding is not dramatically reduced consistent with its retention of half wild-type activity (**Fig. 5.5.C**). Both D118N and D196N resulted in a loss of substrate binding in agreement with the loss of enzymatic activity (**Fig. 5.5.C**). Both D193N and D117N had a similar reduction in binding. For D193N, this was

consistent with the retention of some activity, while for D117N despite the ability to bind substrate the mutation led to a complete loss of activity (**Fig. 5.5.C**). All of the conserved aspartate residues were mutated to glutamate and binding to the dansylated substrate was measured. For all of the mutants (D117E, D118E, D193E, D196E, D265E), the level of activity and the degree of binding were well correlated (**Fig. 5.5.D**) suggesting that a component of the reduction in activity caused by introduction of a larger side chain is due to weaker binding of UDP-MurNAc-pentapeptide-DNS to the active site in *Hy*MraY.

DISCUSSION

In order to determine whether MraY undergoes a one-step or two-step mechanism, we designed and performed dUMP exchange reaction. In our first trial, in addition to Rx 2 and Rx 4, residual amount of UDP-MurNAc-pentapeptide with dUMP incorporated was detected from Rx 1 and Rx 3 (**Fig. 5.1.C**). We speculated that these signals could be coming from endogenous C_{55}-P that was extracted along with MraY from the *E. coli* membrane and stayed in detergent micelles during the process of purification. As predicted, an LC-MS analysis of purified protein sample detected some C_{55}-P present. This likely was a problem for previous experiments that had resulted in different outcomes. For subsequent expression, MraY was extensively washed to ensure removal of all C_{55}-P, confirmed by LC-MS. The dUMP exchange reaction with this sample showed that dUMP incorporation was only detected in Rx 6 and Rx 8 **(Fig. 5.1.C**) when both substrates of MraY are present. Thus, *Hy*MraY catalyzes its reaction via a one-step mechanism in agreement with results seen for *Bs*MraY (Al-Dabbagh et al., 2016; Yao Liu et al., 2016).

An enzymatic reaction that involves two substrates and two products in steady state can be first tested whether it undergoes ordered Bi–Bi or ping pong Bi–Bi mechanisms. Based on our double-reciprocal plot revealing an intersection point of three straight lines in the second quadrant (**Fig. 5.2.B**), we suggest ordered Bi–Bi mechanism and exclude ping pong Bi–Bi mechanism since we expect to see parallel lines for the latter. However, more data analysis is required to distinguish ordered from random Bi–Bi mechanism. Also, it

is critical to perform product inhibition experiments in order to confirm the mechanism (Leskovac, 2003).

We performed an extensive mutagenesis screening for invariant aspartate, lysine, and histidine residues that are positioned in the putative active site facing toward the cytoplasm. In contrast to previous mutagenesis studies, we established baselines by measuring activity of mutants in the absence of Mg^{2+}. Both mutations of D196 to asparagine and glutamate abolished activity (**Fig. 5.3.C**). Interestingly, the side chain of D196 faces away from the cavity of MraY in an apo MraY structure (PDB ID: 4J72), but it is positioned into the cavity in both of the inhibitor-bound MraY structures (PDB ID: 5CKR, 5JNQ). This suggests D196 can move to the position to catalytically act on the substrates once they are bound. Other mutants worth mentioning are the D117 and D118 mutants. While activity of both D117N and D118N was completely lost, D117E and D118E retained some of their activity. This means having a carboxylic acid in the side chain, even though there is one extra carbon, is critical for activity. Based on these findings, we postulate the carboxylic acid on D117 and D118 may be coordinating Mg^{2+} that is required for catalysis.

In order to identify residues that are important in substrate binding, the same *Hy*MraY mutants used to measure activity were subject to binding assays measured by fluorescence anisotropy. Although lipid substrate, C_{55}-P, was not required to monitor UDP-MurNAc-pentapeptide binding to the enzyme, Mg^{2+} was required for its binding (**Fig. 5.5.A**). In general, we observed similar effects of mutations on binding as on activity. For example, D196N and D196E displayed no activity and no binding. However,

the D117 mutants (D117A, D117N) that were inactive (**Fig. 5.3.B**) retained weak binding (**Fig. 5.5.BC**) to UDP-MurNAc-pentapeptide. This suggests the importance of D117 in binding of the substrate in the active site. Additional polarization experiments at higher protein concentrations should be conducted to establish whether maximal anisotropy of the wild type could eventually be achieved with mutants and K_d correctly calculated.

In order to confirm how catalysis is performed by MraY, determining a structure of MraY in complex with two substrates is crucial. Molecular details of the active site of MraY will be compared to the biochemical data and confirm which residues are critical in catalysis and substrate binding. Nonetheless, our dUMP exchange reaction results and the double-reciprocal plot provide new mechanistic insights that MraY from Gram-negative bacteria undergoes a one-step mechanism and, more specifically, ordered Bi–Bi mechanism. In addition, extensive mutagenesis studies were performed on activity and substrate binding in parallel and revealed some residues that need attention for structural studies and further designing of antibiotics.

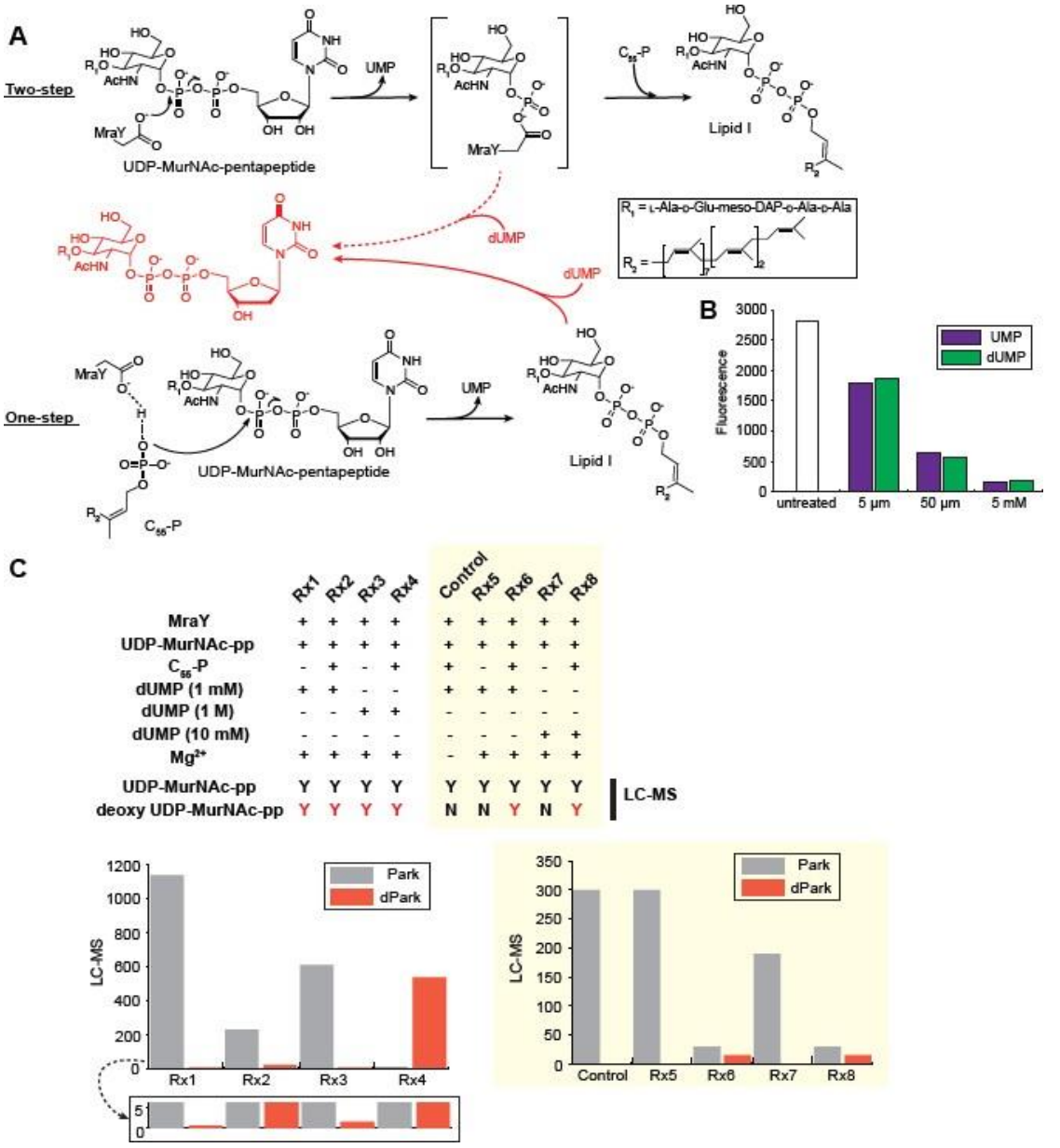

Figure 5.1. Mechanistic scheme and dUMP exchange reaction. A, Two proposed mechanisms of MraY. Two-step mechanism requires UDP-MurNAc-pentapeptide first to form a covalent MurNAc-pentapeptide-phosphoenzyme intermediate and releases UMP, followed by a nucleophilic attack by C_{55}-P for Lipid I formation. One-step mechanism requires both of the substrates for the reaction to go forward and produce both UMP and Lipid I. Since the catalytic reaction by MraY is reversible, depending on at which point of reaction we added dUMP, we could monitor incorporation of dUMP back to UDP-MurNAc-pentapeptide, which is shown in red. B, Inhibition of *Hy*MraY by UMP and dUMP. C, LC-MS results of dUMP exchange reaction. The first four reactions (Rx 1–4) with different compositions listed had deoxy UDP-MurNAc-pentapeptide detected everywhere (inset). After more stringent purification of *Hy*MraY, four reactions were repeated (Rx 5–8) with new dUMP concentration (10 mM) in addition to a negative control without $MgCl_2$ added (shaded in yellow). The LC-MS result for Rx 5–8 shows that only in the presence of both of the substrates, deoxy UDP-MurNAc-pentapeptide was detected. Park: UDP-MurNAc-pentapeptide is also called Park's nucleotide; dPark: deoxy Park's nucleotide. **Lada Klaić, Ph.D. designed and performed experiments.**

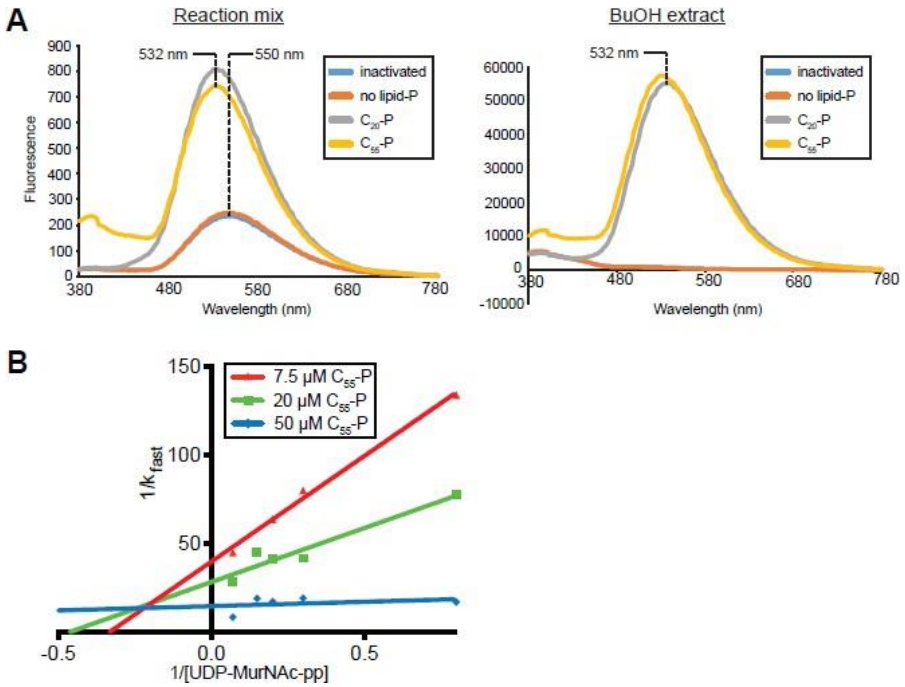

Figure 5.2. Continuous fluorescence activity assay of *Hy*MraY. A left, An example of a fluorescence trace from reaction mixture containing purified *Hy*MraY and two substrates (either native C_{55}-P or C_{20}-P and dansylated UDP-MurNAc-pentapeptide). As compared to traces from inactivated *Hy*MraY or the absence of lipid substrate, the peaks from reaction with lipid substrates shifted from 550 nm to 532 nm. A right, The same reaction as in A left, but only the butanol layer was subject to fluorescence measurement. Product (dansylated Lipid I) was extracted into this organic layer, whereas unreacted soluble UDP-MurNAc-pentapeptide-DNS stayed in the aqueous layer. All of the samples mentioned above were excited at 340 nm and scanned from 380 to 780 nm. B, A double-reciprocal plot of *Hy*MraY. Using the method shown in A right, activity of reactions in a range of UDP-MurNAc-pentapeptide-DNS concentrations was measured and reaction rate was determined for each concentration. Repeat this using three different concentrations of C_{55}-P. Then, a double-reciprocal plot was drawn and it reveals one intersection point of three lines in the second quadrant. **Lada Klaić, Ph.D. designed and performed experiments.**

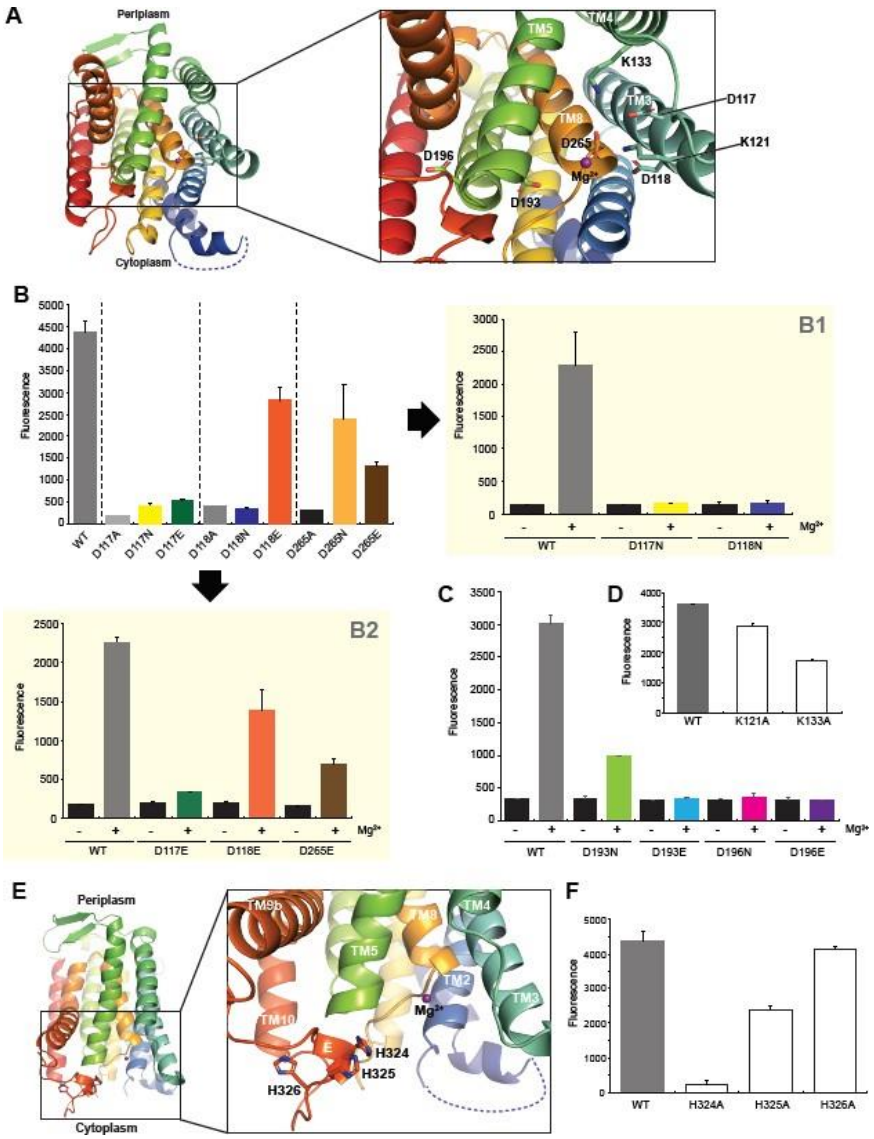

Figure 5.3. Mutagenesis and fluorescence activity assay of *Hy*MraY. A, Chain A of MraY from *A. aeolicus* (PDB ID: 4J72) was drawn in cartoon representation and colored in rainbow with the N-terminus in blue and C-terminus in red. The cytoplasm and periplasm were indicated in the bottom and top of the structure, respectively. The inset has the magnified view of the cavity with conserved aspartate and lysine residues that were mutated in this study. The side chains of those residues are shown in stick representation and the Mg^{2+} ion in purple sphere. PyMOL was used to visualize the *Aa*MraY structure. B, Fluorescence activity assay was performed for the wild-type and mutants of D117, D118, and D265. Activity of the wild type and all of the mutants were measured in triplicates, except for D117A, D118A, and D265A measured only once. B1, Fluorescence measurement of product was performed for the wild type, D117N, and D118N in the absence and presence of $MgCl_2$ to confirm the baseline. Activity of all of the enzymes here was measured in triplicates. B2, Fluorescence was measured for activity of the wild type, D117E, D118E, and D265E in the absence and presence of $MgCl_2$. Activity of all the *Hy*MraY enzymes here was measured in triplicates. C, Fluorescence was measured for activity of the wild type, D193N, D193E, D196N, and D196E with or without $MgCl_2$. Activity of the wild type and all of the mutants here were measured in duplicates. D, Activity of K121A and K133A was measured along with the wild type. Activity of the wild type and two mutants were measured in triplicates. E, Same cartoon representation of *Aa*MraY (PDB ID: 4J72) as in A, but tilted to show the HHH motif on loop E. F, Fluorescence of dansyl group on product was measured for the wild type, H324A, H325A, and H326A. Activity of the wild type and all three mutants were measured in duplicates. The error bars represent standard deviation in all panels above. **Lada Klaić, Ph.D. designed and performed experiments.**

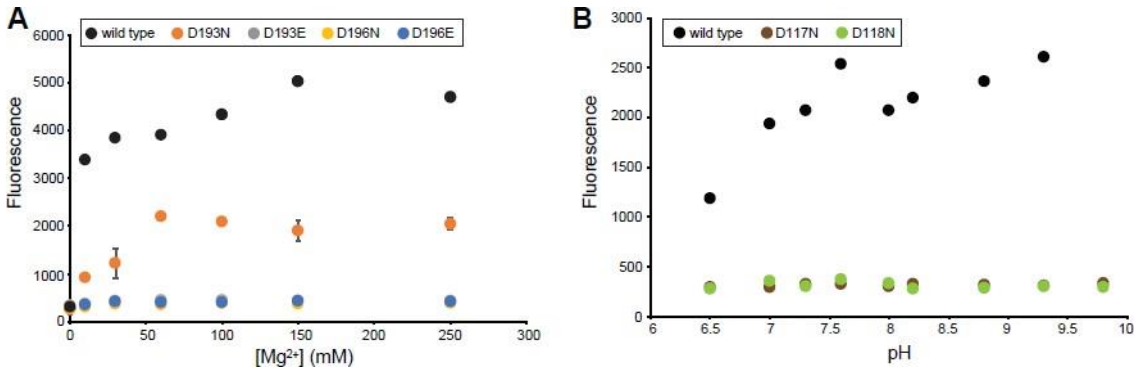

Figure 5.4. Mg^{2+} and pH dependence of *Hy*MraY activity. A, Activity of the wild type and four mutants of *Hy*MraY (D193N, D193E, D196N, D196E) was monitored in the range of 0–250 mM $MgCl_2$. While the wild-type activity was measured once, activity of all the mutants was measured in duplicates. The error bars with standard deviation were shown only for D193N because the other three mutants are inactive. B, Activity of the wild type and two mutants of *Hy*MraY (D117N, D118N) was measured in the range of pH 6.5–9.8. **Lada Klaić, Ph.D. designed and performed experiments.**

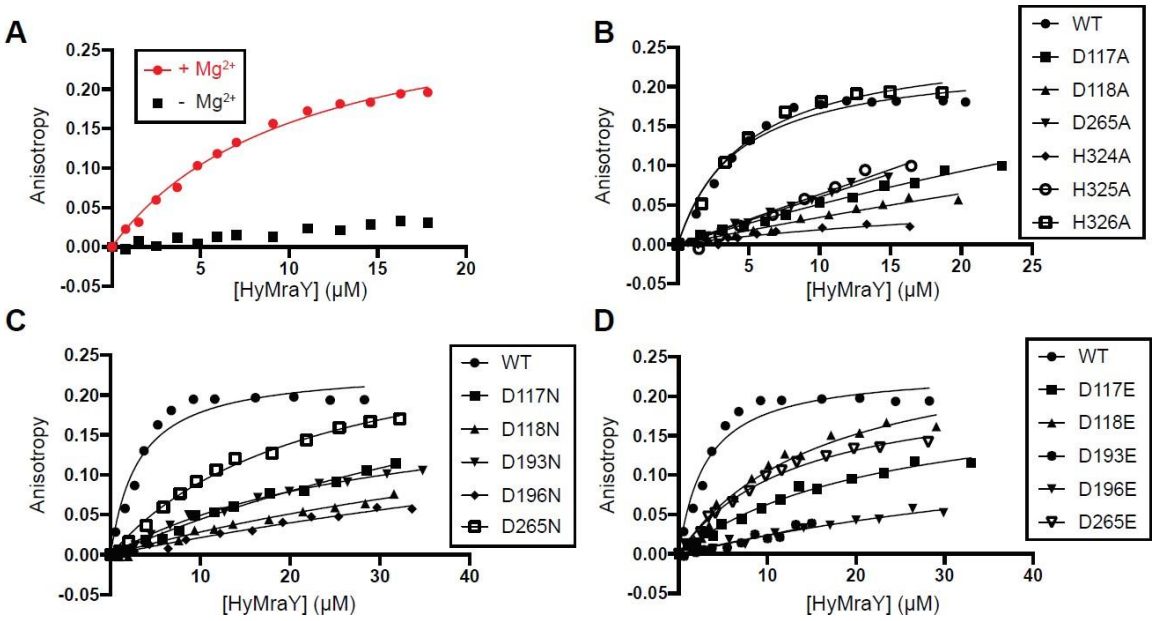

Figure 5.5. Florescence anisotropy assay for measuring binding of UDP-MurNAc-pentapeptide-DNS to *Hy*MraY. A, Fluorescence anisotropy measurement was performed for the wild-type *Hy*MraY with UDP-MurNAc-pentapeptide-DNS in the presence (red circle) and absence (black square) of Mg^{2+}. B, Anisotropy was measured for the alanine mutants of D117, D118, D265, D324, D325, and D326. C, Anisotropy was measured for the asparagine mutants of D117, D118, D193, D196, and D265. D, Anisotropy was measured for the glutamate mutants of D117, D118, D193, D196, and D265. All of the data points were moved to start from zero anisotropy and fitted using *one site specific binding* in Prism. **Lada Klaić, Ph.D. designed and performed experiments.**

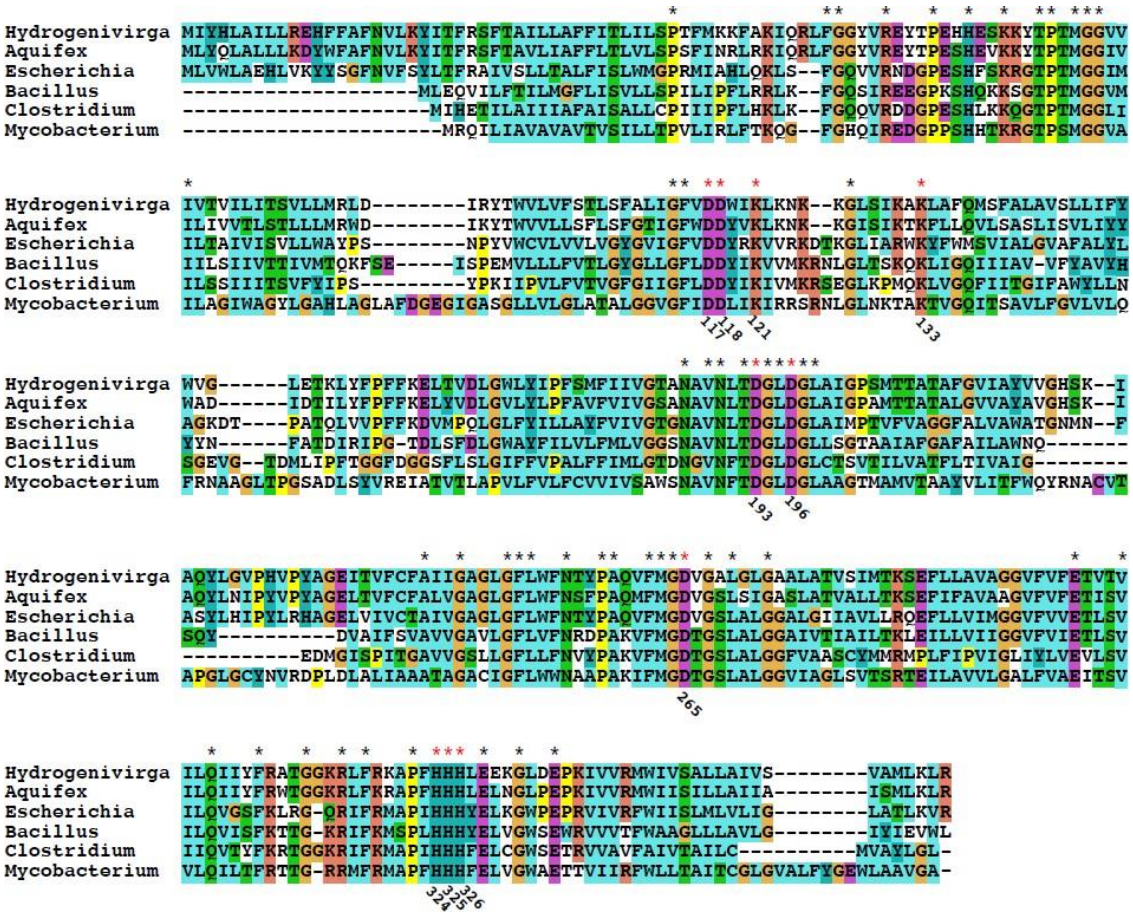

Figure 5.6. Multiple sequence alignment of orthologs of MraY. Sequences of MraY orthologs were from *Hydrogenivirga sp*., *Aquifex aeolicus*, *Escherichia coli*, *Bacillus subtilis*, *Clostridium bolteae*, and *Mycobacterium tuberculosis*. Multiple sequence alignment was performed using ClustalX. Each one-letter code of amino acids background colors based on the chemical properties of side chains. The background colors are following: blue represents Ala (A), Ile (I), Leu (L), Met (M), Phe (F), Trp (W), Val (V), Cys (C) (>60%); red represents Lys (K), Arg (R); magenta represents Glu (E), Asp (D); green represents Asn (N), Gln (Q), Ser (S), Thr (T); pink represents Cys (C) (>85%); orange represents Gly (G); yellow represents Pro (P); cyan represents His (H), Tyr (Y). The columns with asterisks on the top indicate full conservation. The *Hy*MraY invariant residues that were subject to mutagenesis studies have the asterisks colored red. The residue numbering on the bottom of the alignment is from *Hy*MraY.

MATERIALS AND METHODS

Expression and purification of HyMraY

The wild-type and mutant constructs of *Hy*MraY with a N-terminal hexahistidine tag were prepared in a pET22b vector. The plasmids were transformed into *E. coli* Nico21 (DE3) pLemo competent cells (New England Biolabs, Ipswich, MA). Cells were grown at 37°C in 2xYT media with 35 μg/mL chloramphenicol, 50 μg/mL ampicillin, and 0.5 mM L-rhamnose, while shaking at 225 rpm. At the optimal density at 600 nm (OD_{600}) of 0.4, the cells were transferred to 30°C. Protein expression was induced with 0.5 mM isopropyl β-D-thiogalactoside (IPTG, Anatrace, Maumee, OH) at OD_{600} = 0.6. After 4 hours of induction, cells were harvested by centrifugation at 4,000 rpm for 20 min at 4°C. The cell pellet was resuspended in 25 mM Tris-HCl pH 7.5, 1 M NaCl, 10% glycerol, 5 mM β-mercaptoethanol (βME) in the presence of protease inhibitors (PMSF, benzamidine, picrate) and lysed by passing through a microfluidizer for four times. Subsequently, the lysate was centrifugated at 12,000 rpm for 30 min at 4°C to remove insoluble cellular components and unbroken cells. The supernatant was then ultracentrifuged at 45,000 rpm for 30 min at 4°C to isolate membrane fraction.

The membrane pellet was resuspended in 10 mM HEPES pH 7.5, 1 M NaCl, 10% glycerol, 1% DM, 5 mM βME, and rocked for two hours at 4°C. The solubilized membrane components were collected as a supernatant after an ultracentrifugation at 45,000 rpm for 30 min at 4°C. The supernatant was supplemented with 10 mM imidazole and loaded onto pre-equilibrated 1 mL Ni-NTA agarose resin (Qiagen, Germantown, MD). The resin was

incubated overnight at 4°C and then was washed with 50 CV of 10 mM HEPES, pH 7.5, 1 M NaCl, 10% glycerol, 0.15% DM, 5 mM βME. The protein was washed and eluted with increasing concentrations of EDTA (10, 30, and 200 mM), each with 20 CV, in 10 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 0.15% DM, 5 mM βME. Fractions containing the protein, confirmed by a SDS-PAGE, were pooled, concentrated with a 50 kDa MWCO centrifugal filter (Amicon Ultra-4), and buffer exchanged into 20 mM HEPES pH 7.5, 50 mM NaCl, 10% glycerol, 0.15% DM, 5 mM βME. The protein was further purified by cation-exchange chromatography (Uno S6, Bio-Rad, Hercules, CA), eluted in the range of 270 - 650 mM NaCl, followed by size-exclusion chromatography (Superdex 200 10/300 GL; GE Healthcare, Little Chalfont, UK), eluted in 20 mM HEPES pH 7.5, 100 mM NaCl, 10% glycerol, 0.15% DM, 5 mM βME. The protein-containing fractions were pooled, concentrated with a 50 kDa MWCO centrifugal filter, flash frozen in liquid N_2, and stored at -80°C for further assays.

Isolation of UDP-MurNAc-L-Ala-γ-D-Glu-m-DAP-D-Ala-D-Ala (UDP-MurNAc-pentapeptide)

UDP-MurNAc-pentapeptide was isolated from *Bacillus cereus* (strain Frankland and Frankland, ATCC 14579 (Manassas, VA)) according to the protocol used in Kohlrausch et al., with some modifications. Briefly, *B. cereus* was grown in LB at 30°C to OD_{600} = 0.75. The cells were incubated for 15 min after 130 mg/L chloramphenicol was added. Incubation was continued for 60 min following the addition of 15 mg/L vancomycin. The cells were harvested by centrifugation at 4,000 rpm for 15 min at 4°C. The pellet was

resuspended in distilled water (0.1 g wet weight/mL) and stirred into the same volume of boiling water. After 15 min of boiling, suspension was allowed to cool to 4 °C and centrifugated at 45,000 rpm for 60 min at 4°C. The supernatant was flash frozen, stored at -80°C overnight, and lyophilized. Then, the lyophilized sample was dissolved in water (10 mL/1L culture) adjusted to pH 2 by addition of 20 % H_3PO_4. After centrifugation (4,000 rpm, 10 min, 4°C), the supernatant was dialyzed against 10 mM Tris-HCl pH 8.8, and separated via anion-exchange chromatography (Resource Q, Bio-Rad, Hercules, CA), eluted with a 0 - 1M NH_4Ac gradient. Fractions containing UDP-MurNAc-pentapeptide in the range of 360 - 440 mM NH_4Ac were pooled, lyophilized, and confirmed by high-resolution MS (Expected $[MH]^+$= 1426.22 Da; 1427.0 Da observed). Its concentration was determined by spectrophotometry by using $\varepsilon_{260} = 5.8 \times 10^4$ $L \cdot mol^{-1} \cdot cm^{-1}$.

Preparation of undecaprenyl phosphate

Undecaprenyl phosphate was obtained from Larodan Fine Chemicals AB (Malmö, Sweden).

Deoxy UMP exchange reaction and LC-MS analysis

*Hy*MraY for exchange reactions was prepared as above with slight modifications. Following the Ni-NTA column, pooled fractions were run on an Uno S6 column twice and with double volume of eluent in order to wash away all the endogenous undecaprenyl phosphate in the sample. After the gel-filtration column, purified *Hy*MraY (5 ug, 0.6 uM) was incubated with UDP-MurNAc-pentapeptide (5 µL, 12.5 µM;) and deoxy UMP (5 µL, 25µM) in 178 µL 50 mM Tris-HCl pH 8.0, 170 mM KCl, and 10 mM $MgCl_2$. Either undecaprenyl phosphate (5

µL, 28.5 µM in 4% DM) was added to reaction or equal volume of 4% DM was added into three reactions that were incubated at RT for 30 min, 4 hours, and 21 hours, respectively, followed by heat treatment (3 min, 100°C) to stop the reactions. Reactions were extracted with BuOH (200 µL) and water phase was washed with EtOAc (200 µL, four times) to remove detergent. Then, the samples were flash frozen, lyophilized, and analyzed by LC-MS as described below.

Samples were analyzed by LC-MS using a Waters UPLC/LCT Premier XE TOF mass spectrometer by electrospray ionization in the positive ion mode with a reversed-phase BEH C18 2.1x50 mm column. Mobile phases were water and acetonitrile with 0.1% formic acid. Chromatography was performed at a flow rate of 0.33 mL/min at 40 °C using a mobile phase gradient from 10% acetonitrile to 90% acetonitrile in 8 minutes. The mass spectrometer settings were: capillary voltage = 3kV, cone voltage = 65, source temperature = 120 °C, and desolvation temperature = 350 °C.

UMP/dUMP inhibition of MraY enzymatic reaction

Deoxy UMP and UMP were purchased from Sigma-Aldrich (St. Louis, MO). 5 µM, 50 µM, and 5 mM of UMP or dUMP was added to the reaction mixture in a final volume of 200 µL, which contains 20 mM HEPES pH 7.5, 100 mM NaCl, 5% glycerol, 40 mM $MgCl_2$, 0.15% DM, 33 µM undecaprenyl phosphate, and 475 nM purified *Hy*MraY. Reactions were initiated by the addition of 1 µM of UDP-MurNAc-pentapeptide-DNS and incubated at 37°C for 30 min with gentle mixing. Reactions were terminated by heat treatment (3 min, 100 °C).

Fluorescence was measured (excitation 340 nm, emission 380–700 nm) using a 96-well plate reader (TECAN).

Synthesis of UDP-MurNAc-L-Ala-γ-D-Glu-m-DAP(Nε-dansyl)-D-Ala-D-Ala (UDP-MurNAc-pentapeptide-DNS)

UDP-MurNAc-pentapeptide-DNS was synthesized by the reaction of UDP-MurNAc-pentapeptide with dansyl-chloride as described (Weppner & Neuhaus, 1977) with some modifications. Briefly, UDP-MurNAc-pentapeptide (14 mg, 0.012 mmol) was dissolved in 4 mL of 1:1 (volume-to-volume) mixture of 0.25 M $NaHCO_3$ and acetone and dansyl-chloride (72 mg, 0.24 mmol) was added. The solution was stirred for 12 hours in the dark. Acetone was evaporated, and the precipitate was removed by filtration. Filtrate was dialyzed against 10 mM Tris-HCl pH 8.8 and purified via anion-exchange chromatography, eluted in the range of 0 - 1M NH_4Ac gradient. Fractions containing UDP-MurNAc-pentapeptide-DNS were flash frozen, lyophilized, and confirmed by high-resolution MS (Expected $[MH]^+$= 1426.22 Da; 1427.0 Da observed).

Synthesis of phytol-P (C_{20}-P)

To a stirred solution of phytol (77.1 mg, 0.26 mmol) and tetra-*n*-butylammonium dihydrogen phosphate (441 mg, 1.3 mmol) in dry chloroform (1.5 ml) was added in one portion a solution of trichloroacetonitrile (0.153 ml, 1.53 mmol) in dry chloroform (1.5 ml). The reaction mixture was stirred at room temperature for two and a half hours. After evaporation of the solvent, the residue was extracted with upper phase of equilibrium 1-butanol/water mixture

(15 ml) and the solution was washed with the lower phase of the same mixture (10 ml, three times). The 1-butanol solution was co-evaporated with heptane at room temperature to give crude product. The crude product was dissolved in n-propanol : water : 25 % ammonia mixture (1:1:0.02, 10 ml) and passed through a column of Dowex 50Wx8 (NH_4^+ form) equilibrated with the same solvent mixture. Removal of tetra-*n*-butylammonium cation which is essential for the following anion-exchange separation, was verified by TLC and ESI-MS and the Dowex 50Wx8 chromatography was repeated till complete removal was achieved (three times). The effluent was concentrated and n-propanol was added and evaporated from the residue (5ml, twice). The residue was dissolved in chloroform : methanol (2:1) and applied to a DEAE column (AcO^- form) equilibrated with the same solvent mixture. Elution of the column with 0→30 mM ammonium acetate in methanol separated undesired phosphor-ester side product and yielded phytol phosphate. Undesired phytol-diphosphate was eluted with 200 mM ammonium acetate. Fractions containing desired phytol phosphate were passed over Sephadex LH-20 equilibrated with methanol. Product was isolated and analyzed by ESI-MS and ^{1}H and ^{32}P NMR.

Site-directed mutagenesis of HyMraY

All the site-directed mutations, including D117A/N/E, D118A/N/E, D265A/N/E, D193N/E, D196N/E, H324A, H325A, H326A, K320M were prepared in a mixture solution containing Phusion® High-Fidelity PCR Master Mix with HF Buffer from New England Biolabs Inc. (Ipswich, MA), the DNA template of the wild-type *Hy*MraY and primers with or without 5% DMSO. The mixtures were initially heated up to 98°C for 30 sec, entered a 35-cycle of

denaturation at 98°C for 10 sec, annealing to primers either at 55 or 58°C for 20 sec, amplification at 74°C for 3 min, and extra 10 min at 74°C. The size of the mutant DNAs was checked on a 1% agarose gel and then they were treated with *DpnI* at 37°C for 3 hrs. Subsequently, the mutations were confirmed byDNA sequencing analysis.

Stopped and continuous fluorescent activity assay of HyMraY

The activity of the wild-type and mutant MraY was determined using either stopped or continuous fluorescent assay as described by Stachyra et al. (Stachyra, Dini, Ferrari, Heijenoort, et al., 2004). Reaction mixtures were prepared in a final volume of 200 µL, which contain 20 mM HEPES pH 7.5, 100 mM NaCl, 5% glycerol, 40 mM $MgCl_2$, 0.15% DM, 33 µM undecaprenyl phosphate, and 475 nM purified *Hy*MraY. Reactions were initiated by the addition of 1 µM of UDP-MurNAc-pentapeptide-DNS and incubated at 37°C for 30 min with gentle mixing. Reactions were terminated by heat treatment (3 min, 100 °C). Fluorescence was measured (excitation 340 nm, emission 380–700 nm) using a 96-well plate reader (TECAN) for total reaction or water and butanol extracts separately following the separation of lipid-linked product soluble in butanol layer from unreacted UDP-MurNAc- pentapeptide-DNS in water layer.

Continuous assay was used to determine kinetic parameters of reaction. Reaction mixtures were prepared similarly as above with some modifications. While each concentration of C_{55}-P (7.5, 20, 50 µM) was kept constant, concentrations of UDP-MurNAc-pentapeptide-DNS varied in the range between 1.27 and 13.15 µM. Each reaction that contains 500 nM *Hy*MraY was initiated by injecting $MgCl_2$ to final concentration of 40 mM in a TECAN platereader.

For the first eight minutes of reaction, fluorescence was measured (excitation 340 nm, emission 530 nm) in every one- or two-second intervals at about 27 °C. Magellan software was used.

Reconstitution of HyMraY into amphipathic polymers (amphipols)

After the wild-type and mutant *Hy*MraY were purified in the buffer containing 0.15 % DM, they were reconstituted in Amphipol A8-35 (Anatrace) via incubation in 1 : 5 (w/w) (protein : amphipols) and rocked for four hours at 4°C. Detergent was removed by addition of Bio-Beads SM-2 (Bio-Rad) at 20 g wet beads per 1g detergent and rocked overnight at 4°C followed by filtration. To ensue complete detergent removal, this sample was diluted with buffer to below the CMC and filtered in a 30 kDa MWCO centrifugal filter (Amicon Ultra-15). The retained portion was subsequently filtered via size-exclusion chromatography (Superdex 200 10/300 GL; GE Healthcare, Little Chalfont, UK), eluted in 20 mM HEPES pH 7.5, 100 mM NaCl, 10 % glycerol, 5 mM βME. The reconstituted *Hy*MraY in Apols was tested for activity and shows no change as compared to initial detergent conditions (Data not shown).

Quantification of UDP-MurNAc-pentapeptide-DNS binding to HyMraY using fluorescence anisotropy

All fluorescence anisotropy measurements were performed using the wild-type and mutant *Hy*MraY in Apols. The protein was titrated into a buffer containing 1 µM UDP-MurNAc-pentapeptide-DNS at 25°C. (20 mM HEPES pH 7.5, 100 mM NaCl, 40 mM $MgCl_2$, 10% glycerol, 0.15% DM or 10 mM HEPES pH 7.5, 10% glycerol, 150 mM NaCl, 0.03% DDM).

Fluorescence anisotropy was measured at 550 nm (Fluorolog-322, 340 nm excitation) at steady state (20 min) and recorded as the mean of four readings. Control measurements were done in Mg^{2+}-free conditions. Disassociation constants were fit to the data assuming a single binding site.

ACKNOWLEDGEMENTS

We are grateful to the Caltech CCE Multiuser Mass Spectrometry Laboratory for sample analysis and to the National Science Foundation (CRIF 0541745) for enabling the purchase of the Waters LC-MS system used in this work.

C h a p t e r 6

Concluding remarks

This thesis has focused on discussing three bacterial proteins that can serve as potential targets for designing new antibiotics. All three proteins catalyze enzymatic reactions involving one or more monosaccharide(s) and contribute their products to the assembly of more complicated glycoconjugates that play essential roles in the cell.

The protein, Gne described in Chapter 2 produces a GalNAc residue to synthesize three major glycoconjugates that decorate the surface of the human pathogen, *Campylobacter jejuni*. Since involvement of these carbohydrate structures in causing human gastroenteritis had been suggested, inhibiting *Cj*Gne is one strategy to attenuate or abolish pathogenesis of *C. jejuni*. Structural details at a near-atomic level and biochemical data of *Cj*Gne discussed here provide information on unique features of this protein compared to its related enzymes, which will be useful for designing an inhibitor compound.

Two other proteins, MraY and MurG, were discussed in Chapter 3 with the introduction of the peptidoglycan layer that is the major constituent of the bacterial cell wall, protecting the cell from internal osmotic pressure and maintaining the cell shape. MraY and MurG together produce a building block, Lipid II or a lipid-linked disaccharide pentapeptide, for the peptidoglycan. Chapter 3 focused on the structural investigation of MraY from thermophiles, *Hydrogenivirga sp.* and *Mycobacterium thermoresistibile*, and MurG from

Hydrogenivirga sp. I have taken advantage of thermal stability of these proteins for purification and crystallization. However, the long-term goal is to study MraY and MurG from *Mycobacterium tuberculosis*, which is the causative agent of human tuberculosis. Chapter 5 discussed the mechanistic perspective of *Hy*MraY based on the kinetic and biochemical data, which will be further informative once a substrate(s)-bound MraY structure is determined.

Lastly, Chapter 4 discussed the structural studies of MraY from *Escherichia coli* in complex with its novel type of inhibitor, lysis protein E from the phage ΦX174. Structure determination of this *Ec*MraY-E complex will enable us to visualize molecular interaction between *Ec*MraY and protein E. I made some improvement in resolution of X-ray datasets compared to what we previously had, but a near-atomic resolution has to be achieved to correctly build a model into electron density. Once this is accomplished, protein E can be further modified to become more potent and eventually used to treat patients with urinary tract infection caused by uropathogenic strains of *E. coli*.

In conclusion, I hope this thesis conveys the importance of fully understanding of a protein target for designing therapeutics for infectious diseases. Not only *in vitro* structural and biochemical analysis, but also *in vivo* functional analysis of a target protein are required to develop effective antibacterial agents. Interdisciplinary collaboration across fields such as structural biology, biochemistry, synthetic chemistry, microbiology will accelerate this process and we will be one step closer to combating antibiotic resistance.

BIBLIOGRAPHY

Abdi, E., Eleanor, W., Pamela, S., Gundogdu, O., Mills, D. C., Inglis, N. F., … Dorrella, N. (2012). Campylobacter jejuni outer membrane vesicles play an important role in bacterial interactions with human intestinal epithelial cells. *Infection and Immunity*. https://doi.org/10.1128/IAI.00161-12

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., … Zwart, P. H. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*. https://doi.org/10.1107/S0907444909052925

Adams, P. D., Pavel, V., Chen, V. B., Ian, W., Echols, N., Moriarty, N. W., … Thomas, C. (2010). research papers PHENIX : a comprehensive Python-based system for macromolecular structure solution research papers, 213–221. https://doi.org/10.1107/S0907444909052925

Al-Dabbagh, B., Henry, X., El Ghachi, M., Auger, G., Blanot, D., Parquet, C., … 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. https://doi.org/10.1021/bi8006274

Al-Dabbagh, B., Olatunji, S., Crouvoisier, M., El Ghachi, M., Blanot, D., Mengin-Lecreulx, D. & Bouhss, A. (2016). Catalytic mechanism of MraY and WecA, two paralogues of the polyprenyl-phosphate N-Acetylhexosamine 1-phosphate transferase superfamily. *Biochimie*, *127*, 249–257. https://doi.org/10.1016/j.biochi.2016.06.005

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*. https://doi.org/10.1016/S0022-2836(05)80360-2

Amer, A. O. & Valvano, M. a. (2002). Conserved aspartic acids are essential for the enzymic activity of the WecA protein initiating the biosynthesis of O-specific lipopolysaccharide and enterobacterial common antigen in Escherichia coli. *Microbiology (Reading, England)*, *148*(Pt 2), 571–582. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11832520

Anderson, M. S., Eveland, S. S. & Price, N. P. (2000). Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family. *FEMS Microbiology Letters*, *191*(2), 169–175. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11024259

Azad, G. K. & Tomar, R. S. (2014). Ebselen, a promising antioxidant drug: Mechanisms of action and targets of biological pathways. *Molecular Biology Reports*. https://doi.org/10.1007/s11033-014-3417-x

Bacon, D. J., Szymanski, C. M., Burr, D. H., Silver, R. P., Alm, R. A. & Guerry, P. (2001). A phase-variable capsule is involved in virulence of Campylobacter jejuni 81–176. *Molecular Microbiology*. https://doi.org/10.1046/j.1365-2958.2001.02431.x

Barreteau, H., Kovač, A., Boniface, A., Sova, M., Gobec, S. & Blanot, D. (2008). Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiology Reviews*. https://doi.org/10.1111/j.1574-6976.2008.00104.x

Bauer, A. J., Rayment, I., Frey, P. A. & Holden, H. M. (1992). The molecular structure of UDP-galactose 4- epimerase from Escherichia coli determined at 2.5 Å resolution. *Proteins: Structure, Function, and Bioinformatics*. https://doi.org/10.1002/prot.340120409

Bayburt, T. H., Grinkova, Y. V. & Sligar, S. G. (2002). Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Letters*. https://doi.org/10.1021/nl025623k

Beerens, K., Soetaert, W. & Desmet, T. (2015). UDP-hexose 4-epimerases: a view on structure, mechanism and substrate specificity. *Carbohydrate Research*, *414*, 8–14. https://doi.org/http://doi.org/10.1016/j.carres.2015.06.006

Bernatchez, S., Szymanski, C. M., Ishiyama, N., Li, J., Jarrell, H. C., Lau, P. C., … Wakarchuk, W. W. (2005). A single bifunctional UDP-GlcNAc/Glc 4-epimerase supports the synthesis of three cell-surface glycoconjugates in Campylobacter jejuni. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M407767200

Bernhardt, T. G., Roof, W. D. & Young, R. (2000). Genetic evidence that the bacteriophage phi X174 lysis protein inhibits cell wall synthesis. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.97.8.4297

Bernhardt, T G, Roof, W. D. & Young, R. (2000). Genetic evidence that the bacteriophage ΦX174 lysis protein inhibits cell wall synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(8), 4297–4302. https://doi.org/10.1073/pnas.97.8.4297

Bernhardt, Thomas 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*. https://doi.org/10.1046/j.1365-2958.2002.02984.x

Bhatt, V. S., Guo, C. Y., Guan, W., Zhao, G., Yi, W., Liu, Z. J. & Wang, P. G. (2011). Altered architecture of substrate binding region defines the unique specificity of UDP-GalNAc 4-epimerases. *Protein Science*. https://doi.org/10.1002/pro.611

Blango, M. G. & Mulvey, M. A. (2010). Persistence of uropathogenic Escherichia coli in the face of multiple antibiotics. *Antimicrobial Agents and Chemotherapy*. https://doi.org/10.1128/AAC.00014-10

Bouhss, A., Crouvoisier, M., Blanot, D. & Mengin-Lecreulx, D. (2004). Purification and characterization of the bacterial MraY translocase catalyzing the first membrane step of peptidoglycan biosynthesis. *The Journal of Biological Chemistry*, *279*(29), 29974–29980. https://doi.org/10.1074/jbc.M314165200

Bouhss, A., Trunkfield, A. E., Bugg, T. D. H. & 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

Brown, E. D. & Wright, G. D. (2016). Antibacterial drug discovery in the resistance era. *Nature*. https://doi.org/10.1038/nature17042

Bugg, T. D. H. (1999). 3.10 - Bacterial Peptidoglycan Biosynthesis and its Inhibition. In S. D. Barton, K. Nakanishi & O. B. T.-C. N. P. C. Meth-Cohn (Eds.) (pp. 241–294). Oxford: Pergamon. https://doi.org/https://doi.org/10.1016/B978-0-08-091283-7.00080-1

Cardone, G., Heymann, J. B. & Steven, A. C. (2013). One number does not fit all: Mapping local variations in resolution in cryo-EM reconstructions. *Journal of Structural Biology*. https://doi.org/10.1016/j.jsb.2013.08.002

Carson, M., Johnson, D. H., McDonald, H., Brouillette, C. & DeLucas, L. J. (2007). His-tag impact on structure. *Acta Crystallographica Section D: Biological Crystallography*. https://doi.org/10.1107/S0907444906052024

CDC. (2013). Antibiotic resistance threats in the United States, 2013. *United States Department of Health and Human Services*. https://doi.org/CS239559-B

Chen, S., McMullan, G., Faruqi, A. R., Murshudov, G. N., Short, J. M., Scheres, S. H. W. & Henderson, R. (2013). High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy*. https://doi.org/10.1016/j.ultramic.2013.06.004

Chester, N. & Marshak, D. R. (1993). Dimethyl sulfoxide-mediated primer Tm reduction: A method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Analytical Biochemistry*. https://doi.org/10.1006/abio.1993.1121

Chung, B. C., Mashalidis, E. H., Tanino, T., Kim, M., Matsuda, A., Hong, J., … Lee, S. (2016). Structural insights into inhibition of lipid I production in bacterial cell wall synthesis. *Nature*, *533*(7604), 557–560. https://doi.org/10.1038/nature17636

Chung, B. C., Zhao, J., Gillespie, R. a, Kwon, D.-Y., Guan, Z., Hong, J., … Lee, S.-Y. (2013). Crystal structure of MraY, an essential membrane enzyme for bacterial cell wall synthesis.

Science (New York, N.Y.), *341*(6149), 1012–1016. https://doi.org/10.1126/science.1236501

Cooper, D. R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M., Zawadzki, M. & Derewenda, Z. (2007). Protein crystallization by surface entropy reduction: Optimization of the SER strategy. *Acta Crystallographica Section D: Biological Crystallography*. https://doi.org/10.1107/S0907444907010931

Davies, J. & Davies, D. (2010). Origins and Evolution of Antibiotic Resistance O. *Microbiology and Molecular Biology Reviews*. https://doi.org/10.1128/MMBR.00016-10

Day, C. J., Semchenko, E. A. & Korolik, V. (2012). Glycoconjugates Play a Key Role in Campylobacter jejuni Infection: Interactions between Host and Pathogen. *Frontiers in Cellular and Infection Microbiology*. https://doi.org/10.3389/fcimb.2012.00009

Denisov, I. G. & Sligar, S. G. (2016). Nanodiscs for structural and functional studies of membrane proteins. *Nature Structural and Molecular Biology*. https://doi.org/10.1038/nsmb.3195

Denisov, I. G. & Sligar, S. G. (2017). Nanodiscs in Membrane Biochemistry and Biophysics. *Chemical Reviews*. https://doi.org/10.1021/acs.chemrev.6b00690

Dong, Y. Y., Wang, H., Pike, A. C. W., Cochrane, S. A., Hamedzadeh, S., Wyszyński, F. J., … Carpenter, E. P. (2018). Structures of DPAGT1 Explain Glycosylation Disease Mechanisms and Advance TB Antibiotic Design. *Cell*. https://doi.org/10.1016/j.cell.2018.10.037

Edwards, T. E., Liao, R., Phan, I., Myler, P. J. & Grundner, C. (2012). Mycobacterium thermoresistibile as a source of thermostable orthologs of Mycobacterium tuberculosis proteins. *Protein Science*. https://doi.org/10.1002/pro.2084

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Features and development of Coot . *Acta Crystallographica Section D Biological Crystallography*.

https://doi.org/10.1107/s0907444910007493

Fernández-Cruz, A., Muñoz, P., Mohedano, R., Valerio, M., Marín, M., Alcalá, L., … Bouza, E. (2010). Campylobacter bacteremia: Clinical characteristics, incidence, and outcome over 23 years. *Medicine*. https://doi.org/10.1097/MD.0b013e3181f2638d

Flores-Mireles, A. L., Walker, J. N., Caparon, M. & Hultgren, S. J. (2015). Urinary tract infections: Epidemiology, mechanisms of infection and treatment options. *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro3432

Frieden, T. (2013). Antibiotic Resistance Threats in the United States, 2013. *Brochure - US Centers for Disease Control and Prevention*. https://doi.org/CS239559-B

Fry, Ben N., Korolik, V., Ten Brinke, J. A., Pennings, M. T. T., Zalm, R., Teunis, B. J. J., … Van Der Zeijst, B. A. M. (1998). The lipopolysaccharide biosynthesis locus of Campylobacter jejuni 81116. *Microbiology*. https://doi.org/10.1099/00221287-144-8-2049

Fry, Benjamin N., Feng, S., Chen, Y. Y., Newell, D. G., Coloe, P. J. & Korolik, V. (2000). The galE gene of Campylobacter jejuni is involved in lipopolysaccharide synthesis and virulence. *Infection and Immunity*. https://doi.org/10.1128/IAI.68.5.2594-2601.2000

Gerber, S., Lizak, C., Michaud, G., Bucher, M., Darbre, T., Aebi, M., … Locher, K. P. (2013). Mechanism of bacterial oligosaccharyltransferase: In vitro quantification of sequon binding and catalysis. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M112.445940

Getz, E. B., Xiao, M., Chakrabarty, T., Cooke, R. & Selvin, P. R. (1999). A comparison between the sulfhydryl reductants tris(2- carboxyethyl)phosphine and dithiothreitol for use in protein biochemistry. *Analytical Biochemistry*. https://doi.org/10.1006/abio.1999.4203

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*. https://doi.org/10.1038/nmeth.1318

Guerry, P., Szymanski, C. M., Prendergast, M. M., Hickey, T. E., Ewing, C. P., Pattarini, D. L. & Moran, A. P. (2002). Phase variation of Campylobacter jejuni 81–176 lipooligosaccharide affects ganglioside mimicry and invasiveness in vitro. *Infection and Immunity*. https://doi.org/10.1128/IAI.70.2.787-793.2002

Guo, H., Li, L. & Wang, P. G. (2006). Biochemical characterization of UDP-GlcNAc/Glc 4-epimerase from Escherichia coli O86:B7. *Biochemistry*. https://doi.org/10.1021/bi0612770

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, (January). https://doi.org/10.1038/nchembio.2270

Han, J. C. & Han, G. Y. (1994). A procedure for quantitative determination of tris(2-carboxyethyl)phosphine, an odorless reducing agent more stable and effective than dithiothreitol. *Analytical Biochemistry*. https://doi.org/10.1006/abio.1994.1290

Hatahet, F., Boyd, D. & Beckwith, J. (2014). Disulfide bond formation in prokaryotes: History, diversity and design. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1844*(8), 1402–1414. https://doi.org/http://dx.doi.org/10.1016/j.bbapap.2014.02.014

Herzik, M. A., Wu, M. & Lander, G. C. (2019). High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nature Communications*. https://doi.org/10.1038/s41467-019-08991-8

Heydanek, M. G., Struve, W. G. & Neuhaus, F. C. (1969). Initial state in peptidoglycan synthesis. III. Kinetics and uncoupling of phospho-N-acetylmuramyl-pentapeptide translocase (uridine 5'-phosphate). *Biochemistry*, *8*(3), 1214–1221. https://doi.org/10.1021/bi00831a056

Hofreuter, D. (2014). Defining the metabolic requirements for the growth and colonization capacity of Campylobacter jejuni. *Frontiers in Cellular and Infection Microbiology*. https://doi.org/10.3389/fcimb.2014.00137

Ishiyama, N., Creuzenet, C., Lam, J. S. & Berghuis, A. M. (2004). Crystal structure of WbpP, a genuine UDP-N-acetylglucosamine 4-epimerase from Pseudomonas aeruginosa: Substrate specificity in UDP-hexose 4-epimerases. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M401642200

Jankute, M., Cox, J. A. G., Harrison, J. & Besra, G. S. (2015). Assembly of the Mycobacterial Cell Wall. *Annual Review of Microbiology*. https://doi.org/10.1146/annurev-micro-091014-104121

Jorda, J. & Yeates, T. O. (2011). Widespread Disulfide Bonding in Proteins from Thermophilic Archaea. *Archaea*. https://doi.org/10.1155/2011/409156

Kaakoush, N. O., Castaño-Rodríguez, N., Mitchell, H. M. & Man, S. M. (2015). Global epidemiology of campylobacter infection. *Clinical Microbiology Reviews*. https://doi.org/10.1128/CMR.00006-15

Kabsch, W. (2010). *XDS*. *Acta Crystallographica Section D Biological Crystallography*. https://doi.org/10.1107/S0907444909047337

Karlyshev, A. V., Ketley, J. M. & Wren, B. W. (2005). The Campylobacter jejuni glycome. *FEMS Microbiology Reviews*. https://doi.org/10.1016/j.femsre.2005.01.003

Kavanagh, K. L., Jörnvall, H., Persson, B. & Oppermann, U. (2008). Medium- and short-chain dehydrogenase/reductase gene and protein families: The SDR superfamily: Functional and structural diversity within a family of metabolic and regulatory enzymes. *Cellular and Molecular Life Sciences*. https://doi.org/10.1007/s00018-008-8588-y

Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *ELife*. https://doi.org/10.7554/eLife.18722

Lesk, A. M. (1995). NAD-binding domains of dehydrogenases. *Current Opinion in Structural Biology*. https://doi.org/10.1016/0959-440X(95)80010-7

Leskovac, V. (2003). *Comprehensive Enzyme Kinetics*.

Linton, D., Dorrell, N., Hitchen, P. G., Amber, S., Karlyshev, A. V., Morris, H. R., … Wren, B. W. (2005a). Functional analysis of the Campylobacter jejuni N-linked protein glycosylation pathway. *Molecular Microbiology*. https://doi.org/10.1111/j.1365-2958.2005.04519.x

Linton, D., Dorrell, N., Hitchen, P. G., Amber, S., Karlyshev, A. V, Morris, H. R., … Wren, B. W. (2005b). Functional analysis of the Campylobacter jejuni N-linked protein glycosylation pathway. *Molecular Microbiology*, *55*(6), 1695–1703. https://doi.org/10.1111/j.1365-2958.2005.04519.x

Liu, Yao, Rodrigues, J. P. G. L. M., Bonvin, A. M. J. J., Zaal, E. a., Berkers, C. R., Heger, M., … Egmond, M. R. (2016). New insight in the catalytic mechanism of bacterial MraY from enzyme kinetics and docking studies. *Journal of Biological Chemistry*, *291*(29), 15057–15068. https://doi.org/10.1074/jbc.M116.717884

Liu, Yijeng, Thoden, J. B., Kim, J., Berger, E., Gulick, A. M., Ruzicka, F. J., … Frey, P. A. (1997). Mechanistic roles of tyrosine 149 and serine 124 in UDP-galactose 4- epimerase from Escherichia coli. *Biochemistry*. https://doi.org/10.1021/bi970430a

Lovering, A. L., Safadi, S. S. & Strynadka, N. C. J. (2012). Structural Perspective of Peptidoglycan Biosynthesis and Assembly. *Annual Review of Biochemistry*. https://doi.org/10.1146/annurev-biochem-061809-112742

Mahdavi, J., Pirinccioglu, N., Oldfield, N. J., Carlsohn, E., Stoof, J., Aslam, A., … Ala'Aldeen, D. A. A. (2014). A novel O-linked glycan modulates Campylobacter jejuni major outer membrane protein-mediated adhesion to human histo-blood group antigens and chicken colonization. *Open Biology*, *4*(1), 130202–130202. https://doi.org/10.1098/rsob.130202

Matteï, P. J., Neves, D. & Dessen, A. (2010). Bridging cell wall biosynthesis and bacterial morphogenesis. *Current Opinion in Structural Biology*. https://doi.org/10.1016/j.sbi.2010.09.014

McCoy, A. J. (2006). Solving structures of protein complexes by molecular replacement with Phaser. In *Acta Crystallographica Section D: Biological Crystallography*. https://doi.org/10.1107/S0907444906045975

Mitachi, K., Yun, H. G., Kurosu, S. M., Eslamimehr, S., Lemieux, M. R., Klaić, L., … Kurosu, M. (2018). Novel FR-900493 Analogues That Inhibit the Outgrowth of Clostridium difficile Spores. *ACS Omega*. https://doi.org/10.1021/acsomega.7b01740

Nachamkin, I., Allos, B. M. & Ho, T. (1998). Campylobacter species and Guillain-Barrè syndrome. *Clinical Microbiology Reviews*.

Niesen, F. H., Berglund, H. & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*. https://doi.org/10.1038/nprot.2007.321

Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., … Barrell, B. G. (2000). The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. *Nature*, *403*(6770), 665–668. Retrieved from http://dx.doi.org/10.1038/35001088

Persson, B. & Kallberg, Y. (2013). Chemico-Biological Interactions Classification and nomenclature of the superfamily of short-chain dehydrogenases / reductases (SDRs). *Chemico-Biological Interactions*, *202*(1–3), 111–115. https://doi.org/10.1016/j.cbi.2012.11.009

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*. https://doi.org/10.1002/jcc.20084

Phillips, R., Ursell, T., Wiggins, P. & Sens, P. (2009). Emerging roles for lipids in shaping membrane-protein function. *Nature*. https://doi.org/10.1038/nature08147

Pope, J. E., Krizova, A., Garg, A. X., Thiessen-Philbrook, H. & Ouimet, J. M. (2007). Campylobacter Reactive Arthritis: A Systematic Review. *Seminars in Arthritis and Rheumatism*.

https://doi.org/10.1016/j.semarthrit.2006.12.006

Rinken, A., Lavogina, D. & Kopanchuk, S. (2018). Assays with Detection of Fluorescence Anisotropy: Challenges and Possibilities for Characterizing Ligand Binding to GPCRs. *Trends in Pharmacological Sciences*. https://doi.org/10.1016/j.tips.2017.10.004

Rodolis, M. T., Mihalyi, A., O'Reilly, A., Slikas, J., Roper, D. I., Hancock, R. E. W. & Bugg, T. D. H. (2014). Identification of a novel inhibition site in translocase MraY based upon the site of interaction with lysis protein E from bacteriophage ϕX174. *Chembiochem : A European Journal of Chemical Biology*, *15*(9), 1300–1308. https://doi.org/10.1002/cbic.201402064

Sakuraba, H., Kawai, T., Yoneda, K. & Ohshima, T. (2011). Crystal structure of UDP-galactose 4-epimerase from the hyperthermophilic archaeon Pyrobaculum calidifontis. *Archives of Biochemistry and Biophysics*, *512*(2), 126–134. https://doi.org/10.1016/j.abb.2011.05.013

Saliba, A. E., Vonkova, I. & Gavin, A. C. (2015). The systematic analysis of protein-lipid interactions comes of age. *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/nrm4080

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*. https://doi.org/10.1111/j.1574-6976.2008.00105.x

Schaefer, K., Owens, T. W., Kahne, D. & Walker, S. (2018). Substrate Preferences Establish the Order of Cell Wall Assembly in Staphylococcus aureus. *Journal of the American Chemical Society*, *140*(7), 2442–2445. https://doi.org/10.1021/jacs.7b13551

Scheres, S. H. W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*. https://doi.org/10.1016/j.jsb.2012.09.006

Schulz, J. M., Watson, A. L., Sanders, R., Ross, K. L., Thoden, J. B., Holden, H. M. & Fridovich-

Keil, J. L. (2004). Determinants of function and substrate specificity in human UDP-galactose 4'-epimerase. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M405005200

Shaw, M. P., Bond, C. S., Roper, J. R., Gourley, D. G., Ferguson, M. A. J. & Hunter, W. N. (2003). High-resolution crystal structure of Trypanosoma brucei UDP-galactose 4'-epimerase: A potential target for structure-based development of novel trypanocides. *Molecular and Biochemical Parasitology*. https://doi.org/10.1016/S0166-6851(02)00243-8

Shin, S.-M., Choi, J. M., di Luccio, E., Lee, Y.-J., Lee, S.-J., Lee, S. J., … Lee, D.-W. (2015). The structural basis of substrate promiscuity in UDP-hexose 4-epimerase from the hyperthermophilic Eubacterium Thermotoga maritima. *Archives of Biochemistry and Biophysics*, *585*, 39–51. https://doi.org/http://doi.org/10.1016/j.abb.2015.08.025

Silhavy, T. J., Kahne, D. & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor Perspectives in Biology*. https://doi.org/10.1101/cshperspect.a000414

Silver, L. L. (2013). Viable screening targets related to the bacterial cell wall. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.12006

Stachyra, T., Dini, C., Ferrari, P., Bouhss, A., van Heijenoort, J., Mengin-Lecreulx, D., … Le Beller, D. (2004). Fluorescence Detection-Based Functional Assay for High-Throughput Screening for MraY. *Antimicrobial Agents and Chemotherapy*, *48*(3), 897 LP-902. https://doi.org/10.1128/AAC.48.3.897-902.2004

Stachyra, T., Dini, C., Ferrari, P., Bouhss, A., Van Heijenoort, J., Mengin-Lecreulx, D., … Le Beller, D. (2004). Fluorescence Detection-Based Functional Assay for High-Throughput Screening for MraY. *Antimicrobial Agents and Chemotherapy*. https://doi.org/10.1128/AAC.48.3.897-902.2004

Stachyra, T., Dini, C., Ferrari, P., Heijenoort, J. Van, Mengin-lecreulx, D., Blanot, D., … Bouhss, A. (2004). Fluorescence Detection-Based Functional Assay for High-Throughput Screening for MraY Fluorescence Detection-Based Functional Assay for High-

Throughput Screening for MraY, *48*(3), 897–902. https://doi.org/10.1128/AAC.48.3.897

Szymanski, C. M., Ruijin, Y., Ewing, C. P., Trust, T. J. & Guerry, P. (1999). Evidence for a system of general protein glycosylation in Campylobacter jejuni. *Molecular Microbiology*. https://doi.org/10.1046/j.1365-2958.1999.01415.x

Szymanski, C. M., St. Michael, F., Jarrell, H. C., Li, J., Gilbert, M., Larocque, S., … Brisson, J. R. (2003). Detection of Conserved N-Linked Glycans and Phase-variable Lipooligosaccharides and Capsules from Campylobacter Cells by Mass Spectrometry and High Resolution Magic Angle Spinning NMR Spectroscopy. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M301273200

Tanaka, S. & Clemons, W. M. (2012). Minimal requirements for inhibition of MraY by lysis protein E from bacteriophage ΦX174. *Molecular Microbiology*, *85*(5), 975–985. https://doi.org/10.1111/j.1365-2958.2012.08153.x

Thoden, J. B., Frey, P. A. & Holden, H. M. (1996). Crystal structures of the oxidized and reduced forms of UDP-galactose 4-epimerase isolated from Escherichia coli. *Biochemistry*. https://doi.org/10.1021/bi952715y

Thoden, J. B., Frey, P. A. & Holden, H. M. (2002). Molecular Structure of the NADH/UDP-glucose Abortive Complex of UDP-galactose 4-Epimerase from Escherichia coli : Implications for the Catalytic Mechanism † , ‡ . *Biochemistry*. https://doi.org/10.1021/bi9601114

Thoden, J. B., Henderson, J. M., Fridovich-keil, J. L. & Holden, H. M. (2002). Structural Analysis of the Y299C Mutant of Escherichia coli, *277*(30), 27528–27534. https://doi.org/10.1074/jbc.M204413200

Thoden, J. B., Henderson, J. M., Fridovich-Keil, J. L. & Holden, H. M. (2002). Structural analysis of the Y299C mutant of Escherichia coli UDP-galactose 4-epimerase. Teaching an old dog new tricks. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M204413200

Thoden, J. B. & Holden, H. M. (1998). Dramatic Differences in the Binding of UDP-Galactose and UDP-Glucose to UDP-Galactose 4-Epimerase from Escherichia coli. *Biochemistry*, *37*(33), 11469–11477. https://doi.org/10.1021/bi9808969

Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L. & Holden, H. M. (2000). Crystallographic evidence for Tyr 157 functioning as the active site base in human UDP-galactose 4-epimerase. *Biochemistry*. https://doi.org/10.1021/bi000215l

Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L. & Holden, H. M. (2001). Human UDP-galactose 4-epimerase. Accommodation of UDP-N-acetylglucosamine within the active site. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M100220200

Tribet, C., Audebert, R. & Popot, J.-L. (1996). Amphipols: Polymers that keep membrane proteins soluble in aqueous solutions. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.93.26.15047

Urbaniak, M. D., Tabudravu, J. N., Msaki, A., Matera, K. M., Brenk, R., Jaspars, M. & Ferguson, M. A. J. (2006). Identification of novel inhibitors of UDP-Glc 4'-epimerase, a validated drug target for African sleeping sickness. *Bioorganic and Medicinal Chemistry Letters*. https://doi.org/10.1016/j.bmcl.2006.08.091

van Heijenoort, J. (2007). Lipid Intermediates in the Biosynthesis of Bacterial Peptidoglycan. *Microbiology and Molecular Biology Reviews*. https://doi.org/10.1128/MMBR.00016-07

Vollmer, W., Blanot, D. & De Pedro, M. A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*. https://doi.org/10.1111/j.1574-6976.2007.00094.x

Weppner, W. A. & Neuhaus, F. C. (1977). Fluorescent substrate for nascent peptidoglycan synthesis. Uridine diphosphate N acetylmuramyl (NC??) 5 dimethylaminonaphthalene 1 sulfonyl)pentapeptide. *Journal of Biological Chemistry*.

WHO. (2017). WHO publishes list of bacteria for which new antibiotics are urgently needed. https://doi.org//entity/mediacentre/news/releases/2017/bacteria-antibiotics-

needed/en/index.html

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., … Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*. https://doi.org/10.1107/S0907444910045749

World Health Organisation (WHO). (2017). WHO publishes list of bacteria for which new antibiotics are urgently needed. https://doi.org//entity/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/index.html

Young, K. T., Davis, L. M. & DiRita, V. J. (2007). Campylobacter jejuni: Molecular biology and pathogenesis. *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro1718

Young, N. M., Brisson, J. R., Kelly, J., Watson, D. C., Tessier, L., Lanthier, P. H., … Szymanski, C. M. (2002). Structure of the N-linked glycan present on multiple glycoproteins in the gram-negative bacterium, Campylobacter jejuni. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M206114200

Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y. & Agard, D. A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods*. https://doi.org/10.1038/nmeth.4193

Zheng, Y., Struck, D. K. & Young, R. (2009). Purification and functional characterization of φX174 lysis protein E. *Biochemistry*. https://doi.org/10.1021/bi900469g

Zhou, H.-X. & Cross, T. A. (2013). Influences of Membrane Mimetic Environments on Membrane Protein Structures. *Annual Review of Biophysics*. https://doi.org/10.1146/annurev-biophys-083012-130326

Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E. & Scheres, S. H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *ELife*. https://doi.org/10.7554/eLife.42166