STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF ENZYMES CENTRAL TO BACTERIAL CARRIER LIPID SYNTHESIS AND RECYLING

by

SEAN DOUGLAS WORKMAN

B.Sc. (Hons., Co-op), The University of British Columbia, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Biochemistry and Molecular Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Structural and functional characterization of enzymes central to bacterial carrier lipid synthesis and recycling.

submitted by Sean Workman in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Biochemistry and Molecular Biology

Examining Committee:

Dr. Natalie Strynadka, Biochemistry and Molecular Biology

Supervisor

Dr. Lawrence McIntosh, Biochemistry and Molecular Biology

Supervisory Committee Member

Dr. Calvin Yip, Biochemistry and Molecular Biology

Supervisory Committee Member

Dr. Katherine Ryan, Chemistry

University Examiner

Dr. Michael Murphy, Microbiology and Immunology

University Examiner

Abstract

In bacteria, the carrier lipid undecaprenyl phosphate ($C_{55}P$) is used as a scaffold for the synthesis of bacterial cell wall polymers such as peptidoglycan. $C_{55}P$ is synthesized as undecaprenyl pyrophosphate ($C_{55}PP$) by undecaprenyl pyrophosphate synthase (UppS) and must be dephosphorylated by an as yet unknown mechanism before it can be used in cell wall biosynthesis. Individual subunits of cell wall polymers are assembled in the cytoplasm on $C_{55}P$ before being flipped to the periplasmic face of the membrane where they are polymerized into the existing structure, releasing $C_{55}PP$ as a by-product. The resultant $C_{55}PP$ must be recycled to $C_{55}P$ before being used in another round of cell wall polymer biosynthesis. The major protein responsible for recycling in *Escherichia coli* is undecaprenyl pyrophosphate phosphatase (UppP). The ongoing synthesis and recycling of undecaprenyl phosphate by UppS and UppP, respectively, are required for the survival and pathogenesis of bacteria; thus, both enzymes represent attractive targets for the development of therapeutics. In this thesis, UppP was structurally and functionally characterized, and inhibition of UppS by novel inhibitors was investigated. The X-ray crystallographic structure of the polytopic integral membrane protein membrane protein UppP was solved to 2.0 Å resolution using lipid cubic phase (LCP) crystallization. The crystal structure revealed an unexpected membrane topology and three-dimensional structure that suggests a potential role for UppP as a $C_{55}P$(P) lipid flippase and allowed for the rationalization of previously published site-directed mutagenesis results. An ordered monoolein molecule in the active site of the enzyme allowed us to model a $C_{55}PP$ and propose a catalytic mechanism for $C_{55}PP$ dephosphorylation. The crystal structure of UppS from *Bacillus subtilis* was solved in apo- and inhibitor bound states using X-ray crystallography, allowing us to rationalize two novel inhibitors' superior efficacy against *B. subtilis* UppS *versus Staphylococcus aureus* or *E. coli* orthologues.

The inhibitors bind in the hydrophobic tunnel into which the nascent C_{55}PP product of UppS grows. Additionally, a crystal structure of *B. subtilis* UppS in complex with clomiphene provided a clearer structural basis of its inhibition of UppS and provides a basis for the rational design of improved UppS inhibitors.

Lay Summary

Some bacteria are able to cause infections that can lead to serious illness and sometimes death. The prevalence of bacteria that are resistant to current antibiotic therapies is on the rise and some strains of bacteria are becoming resistant to our last line of defense drugs. In order to develop new, more effective antibiotics, novel drug targets must be characterized and the effect of new antibiotics on established drug targets must be understood. Most drug targets are proteins, and understanding their structure allows us to understand their function. This thesis structurally characterizes a protein that is involved in the construction of the bacterial cell wall, a process that has been targeted with great historical success, and characterizes the mechanism of action for two novel antibiotics against a protein that is an established drug target. This research could lead to the development of much needed new antibiotics.

Preface

A portion of Chapter 1 of this thesis was published as part of an invited review article (Workman S.D., Strynadka N.C.J. 2020. A slippery scaffold: synthesis and recycling of the bacterial cell wall carrier lipid. *J Mol Biol*). Chapter 1 was written by me and revised by Dr. Natalie Strynadka.

A portion of Chapter 2 of this thesis was published as a research article (Workman S.D., Worrall L.J., Strynadka N.C.J. 2018. Crystal structure of an intramembranal phosphatase central to bacterial cell-wall peptidoglycan biosynthesis and lipid recycling. *Nat Commun.* 9:1159). Cloning, protein expression, purification, crystallization, and data collection was carried out by me. Structure solution and refinement was carried out by me with the assistance of Dr. Liam Worrall. Cross-linking and enzyme kinetics experiments were carried out by me. Figures were generated by me and Dr. Liam Worrall. The manuscript was written by me, Dr. Liam Worrall, and Dr. Natalie Strynadka, with revisions by Dr. Natalie Strynadka.

Chapter 3 is being prepared for submission as a research article. Cloning, protein expression, protein purification, crystallization, structure solution and refinement were done by me. Synthesis of MAC-0547630 and JPD-447 inhibitors was carried out by Dr. Jonathan Day in the laboratory of Dr. Michael Organ at the University of Ottawa. Checkboard analysis of inhibitor synergy was carried out by Dr. Maya Farha in the laboratory of Dr. Eric Brown at McMaster University. Figures were generated by me. The manuscript was written by me and revised by Dr. Natalie Strynadka.

Chapter 4 was written by me and revised by Dr. Natalie Strynadka.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

β-OG	n-Octyl-β-D-glucoside
C_5PP	Isopentyl pyrophosphate
C_{10}PP	Geranyl pyrophosphate
C_{15}PP	Farnesyl pyrophosphate
C_{15}(S)PP	Farnesyl thiopyrophosphate
C_{55}OH	Undecaprenol
C_{55}P	Undecaprenyl phosphate
C_{55}PP	Undecaprenyl pyrophosphate
CHAPS	(3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate)
CLS	Canadian Light Source
CMC	Critical micelle concentration
Cryo-EM	Cryogenic electron microscopy
DAP	Diaminopimelic acid
Ddl	D-Ala-D-Ala ligase
DDM	n-Dodecyl-β-D-maltoside
DM	n-Decyl-β-D-maltoside
DGPP	Diacylglycerol pyrophosphate
DMAPP	Dimethylallyl diphosphate
DolP	Dolichyl phosphate
DolPP	Dolichyl pyrophosphate
DSPG	Distearoyl phosphatidylglycerol
DSS	Disuccinimidyl suberate
EI	Electrospray ionization
FICI	Fractional inhibitory concentration index
FSEC	Fluorescence-detection size exclusion chromatography
EGS	Ethylene glycol bis(succinimidyl succinate)
GDN	Glyco-diosgenin
GFP	Green fluorescent protein
Glc*N*Ac	*N*-acetylglucosamine
HDX-MS	Hydrogen-deuterium exchange mass spectrometry
HiLiDe	High concentrations of lipid and detergent
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
IDA	Iminodiacetic acid
IMAC	Immobilized metal affinity chromatography
IMISX	*In meso in situ* crystallography
Kdo	3-Deoxy-d-manno-oct-2-ulosonic acid
LB	Lysogeny broth
LCP	Lipid cubic phase
LDAO	n-Dodecyl-N,N-dimethylamine-N-oxide
LLO	Lipid linked oligosaccharide
LPS	Lipopolysaccharide
MD	Molecular dynamics
MEP	2-*C*-Methyl-$_D$-erythritol 4-phosphate

MFS	Major facilitator superfamily
MR	Molecular replacement
MRSA	Methicillin resistant *Staphylococcus aureus*
Mur*N*Ac	*N*-acetylmuramic acid
NBD	7-nitro-2,1,3-benzoxadiazol
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
OA	O-antigen
OD	Optical density
PA	Phosphatidic acid
PAP2	Phosphatidic acid phosphatase 2
PBP	Penicillin binding protein
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PE	Phosphatidylethanolamine
PEG	Polyethylene glycol
PG	Peptidoglycan
RMSD	Root-mean-square deviation
RNAP	RNA polymerase
SAR	Structure activity relationship
SAXS	Small angle X-ray scattering
SDS	Sodium dodecyl sulfate
SMA	Styrene maleic acid
TLC	Thin layer chromatography
TA	Teichoic acid
TM	Transmembrane
VRSA	Vancomycin resistant *Staphylococcus aureus*
WTA	Wall teichoic acid

Acknowledgements

I would like to thank both my grandparents and my parents for supporting me and allowing me the privilege of being in a position where I could pursue my scientific passions. I would like to give special thanks to my father for introducing me to life sciences at such a young age.

I am very grateful to my supervisor, Dr. Natalie Strynadka, for giving me the opportunity to work in her lab and for providing me with the freedom and resources to achieve my goals. She always gave me the push I needed when things seemed hopeless, and she believed in me at some of the toughest times in my life.

I thank my committee members for their feedback on my work, and for ensuring that I stayed on the right path.

The members of the Strynadka lab, both past and present, have all contributed to my success in one way or another, be that through helping me with an experiment or just letting me bounce ideas off of them. Nathanael Caveney and Guillermo Cabellero, my coffee crew, made the late nights and weekend days at the lab so much easier. It has been an honour to have Bronwyn Lyons as my "science little sister" – I will always be grateful for her organization of lab outings. Dr. Chrisitan Lizak taught me everything I know about working with membrane proteins and there is no way I would have achieved my goals without his sage advice. I find it hard to believe how much work Dr. Liam Worrall managed to get done while I was in the lab, because I can't seem to remember a day going by without asking for his advice about something! Finally, I have to thank my "lab mom" Marija Vuckovic for being the glue that holds the lab together, for always trying to keep me thinking positive, and for being an example of the type of kind, helpful person that I aspire to be.

When I moved to Vancouver, I didn't know anybody and felt a little bit lost. Dr. Craig Kerr invited me to play on the dodgeball team he helped organize and I immediately had a circle of friends. Having fun and blowing off steam with "Microballology" kept me sane when things were starting to feel too tough. So thank you to Dr. Craig Kerr, Dr. Eric Brown, Dr. Keith Mewis, Dr. Jan Burian and Dr. Michael Jones for allowing me to come and be your team's all-star. The later additions of Udit Dalwadi, Bill Rees, Daniel Lisko, Zach Morse, Blair Hardman, Jeff To made things all the more fun.

I was fortunate to receive funding from NSERC and from UBC during the course of my studies.

Finally, I would like to thank Dr. Martin Boulanger for giving me the opportunity to discover my love for structural biology in his lab at the University of Victoria and to Dr. Michelle Parker for being such an excellent mentor during that time.

For Melissa, for being my light when things seemed most dark.

1 Introduction

1.1 Overview of the bacterial cell wall

The bacterial cell wall is a complex, polymeric structure that allows bacteria to survive and thrive in their often harsh environments. Peptidoglycan (PG), common to both Gram-negative and Gram-positive bacteria, is an essential component of the bacterial cell wall that is responsible for maintaining cell shape and preventing rupture due to osmotic pressure[1]. The mesh-like PG sacculus is an elastic structure composed of strands of alternating *N*-acetylmuramic acid (Mur*N*Ac) and *N*-acetylglucosamine (Glc*N*Ac) (Fig. 1.1) monomers that are connected by β-1,4 glycosidic bonds. Individual PG strands are crosslinked by short "stem" pentapeptides that are covalently attached to Mur*N*Ac through its D-lactoyl group[1]. The most commonly observed stem sequence in nascent PG of Gram negative bacteria such as *Escherichia coli* is L-Ala-γ-D-Glu-*meso*-DAP (or L-Lys)-D-Ala-D-Ala (DAP, diaminopimelic acid), with the terminal D-Ala being cleaved during the crosslinking reaction that yields the mature macromolecule (Fig 1.1). In most cases, the crosslinking (transpeptidation) of individual PG strands occurs through the carbonyl group of the D-Ala at position 4 and the amino group of the DAP at position three on an adjacent strand, either directly or in some cases mediated by a short peptide bridge[1].

Although the PG sacculus surrounds the exterior of the cell, biosynthesis of PG is initiated in the cytoplasm and involves several inner membrane-associated steps (Fig. 1.1). First, the well characterized Mur pathway acts to sequentially synthesize a UDP-Mur*N*Ac-pentapeptide precursor molecule[2]. In the first membrane associated step, the integral membrane protein MraY catalyzes the transfer of phospho-Mur*N*Ac-pentapeptide onto an undecaprenyl phosphate ($C_{55}P$) carrier lipid, generating lipid I (Fig 1.1). Subsequently, the membrane associated glycosyltransferase MurG modifies lipid I through the transfer of Glc*N*Ac from a UDP-Glc*N*Ac

donor to generate lipid II, the mature PG precursor. Lipid II is then flipped from the cytoplasmic leaflet of the membrane to the periplasmic leaflet through action of a dedicated flippase, MurJ, where polymerization of PG is carried out by the combined glycosyltransferase and transpeptidase activities of penicillin binding proteins (PBPs)[2]. Very recently, evidence for an additional role of the SEDS family of enzymes in glycopolymerization similarly coupled to PBP transpeptidation activity has also come to light [3,4].

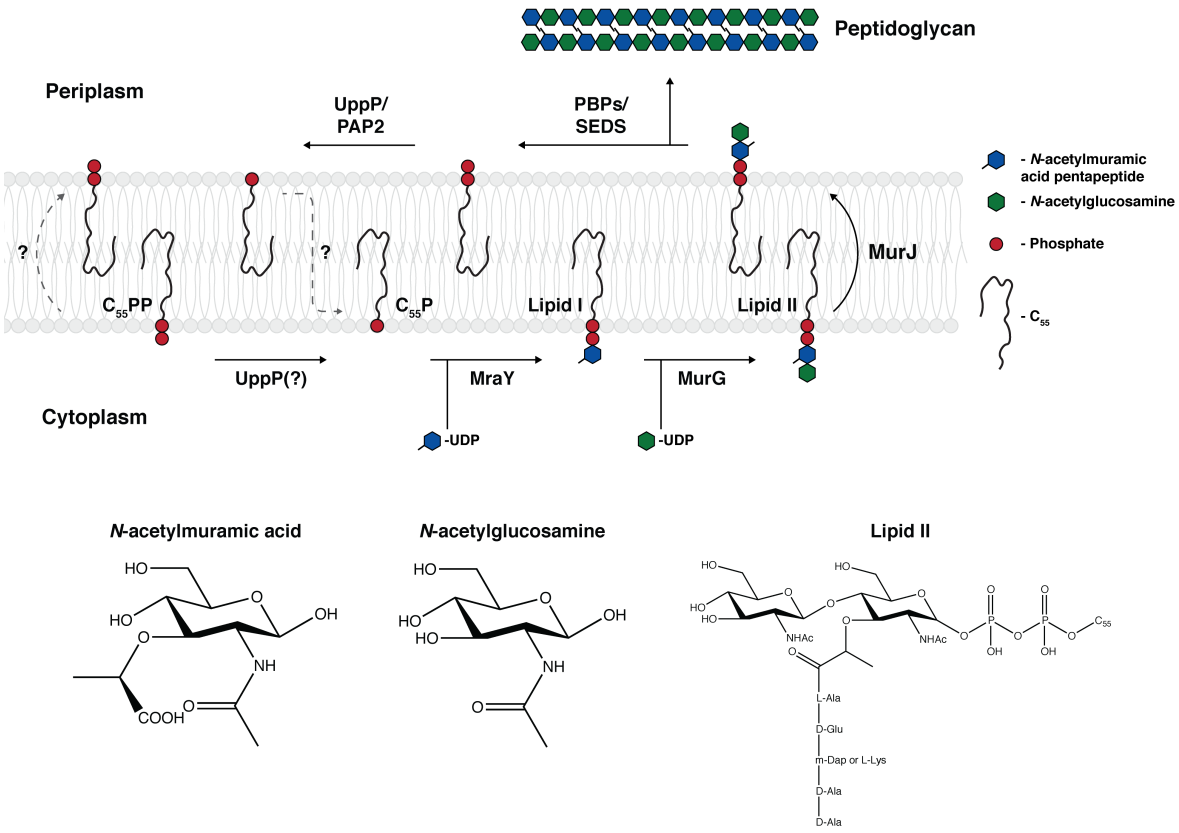


Figure 1.1 Schematic drawing of the peptidoglycan biosynthesis pathway.
The carrier lipid precursor ($C_{55}PP$) is dephosphorylated, by UppP or an as yet unknown mechanism, generating the active carrier lipid molecule ($C_{55}P$). MraY modifies the carrier lipid through the addition of an *N*-acetylmuramic acid pentapeptide to generate lipid I. Lipid I is modified by MurG through the addition of an *N*-acetylglucosamine to generate lipid II. Lipid II is flipped to the periplasmic leaflet of the inner membrane by MurJ, where the combined glycosyltransferase and transpeptidase activities of the PBPs and SEDS protein incorporate the nascent PG strand in to the existing sacculus. The by-product of PG polymerization is $C_{55}PP$,

which then needs to be recycled by UppP and/or PAP2 phosphatase enzymes before the carrier lipid can be used in a subsequent round of PG biosynthesis.

While PG is a critical component, it is far from the only polymer that makes up the bacterial cell wall. In Gram-positive bacteria, PG can be modified by wall teichoic acids (WTAs)[5]. These anionic polymers are attached to the PG sacculus through the C6 hydroxyl of Mur*N*Ac and can contribute up to 60% of the total cell wall weight[5]. The exact composition of WTAs can vary greatly between species and even strains, but the disaccharide linkage unit and the polyol nature of the main chain are common to most WTAs characterized thus far. WTAs play a diverse number of roles in cell division, virulence, persistence and drug resistance. Lipopolysaccharide (LPS) is a major component of the outer membrane of Gram-negative bacteria, frequently playing a key role in pathogenesis and providing an effective permeability barrier against antibiotics[6]. LPS is typically composed of lipid A, a core oligosaccharide, and an O-antigen. In addition to WTAs and LPS, many bacteria produce capsular polysaccharides, membrane derived polysaccharides, and in the case of *Enterobacteriaceae*, enterobacterial common antigen[7–9]. Common to the synthesis of all of these cell wall polymers is the use of an extended polyprenyl carrier lipid for their assembly on the inner membrane[10].

1.2 The bacterial cell wall as a drug target

Synthesis of the bacterial cell wall has historically stood as an important and attractive target for the development of therapeutics because its ongoing synthesis is typically necessary for the survival of bacteria within their natural environs, and because the enzymes comprising the cell wall synthetic machinery are unique to bacteria[11]. Though in recent years a great deal of work

has been done to try and develop inhibitors of WTA and LPS biosynthesis, the PG biosynthetic pathway remains one of the most important and successful drug targets in history[11].

Since the serendipitous discovery of benzylpenicillin (or penicillin G) by Sir Alexander Fleming in 1929, many antibiotic classes have been discovered and developed; however, none come close to the success of the β-lactams. Following the introduction of penicillin G to the clinic in the 1940s, a number of penicillin derivatives were developed to try to overcome β-lactamase mediated resistance and to attain better activity against Gram-positive bacteria[11]. Subsequent discovery and development of the cephalosporins (1950s), carbapenems (1970s), and monobactams (1980s) in addition to potent β-lactamase inhibitors have been the result of the ongoing arms race between humans and bacterial pathogens. Up to 65% of antibiotics that are prescribed belong the β-lactam family of antibiotics and despite their immediate success, their mechanism of action was not deduced until 1965[12,13]. Crosslinking of PG strands is critical for the strength and viability of bacteria, and β-lactams act by disrupting the penicillin binding protein (PBP) mediated transpeptidation reaction that creates these crosslinks. This is achieved through the β-lactams' eponymous central ring – specifically its chemical and stereochemical resemblance to the D-Ala-D-Ala terminus of the PG stem peptide in the manner of a substrate analog. When subjected to β-lactam treatment, the catalytic Ser residue of the PBP transpeptidase domain carries out a nucleophilic attack on the amide carbonyl carbon, leading to the formation of a tetrahedral oxyanion transition state that collapses to form a stable acyl-enzyme intermediate[14]. Inhibition of PBPs in this manner leads to a reduction in PG crosslinks and dysregulation of PG metabolism, ultimately resulting in lysis of the cell.

While inhibition of the periplasmic transpeptidation step of PG synthesis by β-lactam inhibitors has proven to be incredibly successful, it is not the only point in the synthesis of the

bacterial cell wall that is targeted by antibiotics in the clinic. The early (soluble) cytoplasmic steps of cell wall biosynthesis can be inhibited by D-cycloserine, which targets alanine racemase (Alr) and D-Ala-D-Ala ligase (Ddl), or by fosfomycin, which targets MurA[11] in the first committed step of lipid II biosynthesis (enoylpyruvate transfer from PEP to the 3' position of UDP-N-acetylglucosamine). There are currently no clinically approved drugs that target PG synthesis enzymes at the cytoplasmic face of the inner membrane. Tunicamycin, a uridyl peptide, has long been known to act as a competitive inhibitor of MraY phosphotransfer, but its interference with mammalian glycoprotein biosynthesis and consequent toxicity precludes its use as a drug in humans[11]. More recently discovered uridyl peptides (i.e. liposidomycins, caprazamycins, etc.) do not display the same off target effects and CPZEN-45, a derivative of caprazamycin, is currently being developed as a novel anti-tuberculosis drug [15]. At the periplasmic face of the inner membrane, the glycosyltransferase domain of PBPs stands as a promising target for the development of novel therapeutics, but as with the enzymes associated with the cytoplasmic face of the inner membrane, there are no clinically approved therapeutics. Moenomycin, a phosphoglycolipid natural product produced by *Streptomycetes*, has been shown structurally to act as a substrate mimic of lipid II paving the way for mechanistic understanding and novel inhibition strategies [16]. Although in of itself, moenomycin has demonstrated nM potency against several Gram-positive bacteria and is used widely as a cattle feed supplement, the compound was deemed unsuitable for human clinical use due to poor absorption and bioavailability[11].

As alluded to above, there are a series of antibiotic agents that target the substrates of PG synthesis, rather than directly interacting with biosynthetic enzymes. The glycopeptide family of antibiotics includes vancomycin, long considered a drug of last resort for the treatment of Gram-

positive infections; however, the rise of vancomycin resistance in the clinic has brought about a push for the development of next generation vancomycin derivatives with telavancin being a prominent example recently approved for the treatment of hospital acquired *Staphylococcus aureus* pneumonia [17]. Glycopeptides broadly inhibit PG synthesis by binding directly to the D-Ala-D-Ala moiety of lipid II or nascent peptidoglycan strands, though later generation derivatives also act to permeabilize membranes and bind modified stem peptides. Lantibiotics, such as nisin, show potential and also act by binding to lipid II, but bind at a distinct site from that of the glycopeptides[18]. Finally, the product of the PBP transglycosylase step, undecaprenyl pyrophosphate ($C_{55}PP$), can be targeted by bacitracin, a cyclic antibiotic that binds to the pyrophosphate moiety of $C_{55}PP$ in a Zn^{2+} dependent manner[19,20]. Bacitracin is limited to topical use in combination with other antibiotics due to its systemic toxicity.

As more stable and efficacious antibiotic therapies have been developed, so too have the strategies by which pathogens overcome them. Expression of efflux pumps, alternative β-lactam insensitive PBPs, and progressively more potent β-lactamase enzymes all call for the development of therapeutic agents that target different points in the biosynthesis of the bacterial cell wall[21]. In recent years, the enzymes responsible for the synthesis and recycling of $C_{55}P$, the universal carrier lipid for bacterial cell wall polymer biosynthesis, have emerged as promising targets for the development of novel therapeutics[10,22–24].

Synthesis of $C_{55}P$ occurs in the cytoplasm and is mediated by undecaprenyl pyrophosphate synthase (UppS), a *cis*-prenyltransferase. Prenyltransferase inhibitors are used for the treatment of cancers and parasites, among other diseases, and in recent years there has been an explosion of interest in the development of UppS inhibitors for the treatment of bacterial infections[21-30]. Recycling of $C_{55}P$ from $C_{55}PP$ is carried out by several integral membrane proteins belonging to

the phosphatidic acid phosphatase 2 (PAP2) family of enzymes, as well as undecaprenyl pyrophosphate phosphatase (UppP, also known as BacA for bacitracin resistance protein A)[10]. Bacitracin has long stood as a proof of concept for the notion of inhibiting bacterial cell wall biosynthesis at the carrier lipid recycling step and, very recently, a series of papers described the development of a phenylthiazole class of inhibitors that effectively inhibit both UppS and UppP[28,29,39]. In addition to *in vitro* inhibition of purified enzymes, the inhibitors have been shown to be effective *in vivo* against drug resistant Gram-positive pathogens; however, they show limited potency against Gram-negative pathogens.

1.3 Metabolism of the bacterial cell wall carrier lipid

The translocation of sugars and glycan chains across membranes using extended polyprenyl phosphate carrier lipids is a highly conserved process across all kingdoms of life[10,40,41]. In both prokaryotes and eukaryotes, these lipids play a central role in protein glycosylation. There are over 125 described congenital disorders of glycosylation (CDG) in mammals that cover all major glycosylation pathways and many of these disorders arise from deficiencies in the biosynthesis of dolichyl phosphate (DolP)[42]. Deficient DolP biosynthesis can lead to cerebellar ataxia, ocular malfunctions, and even early childhood death[43]. While bacteria use these polyprenyl phosphate lipids for protein glycosylation as well, their primary role is as the universal carrier lipid for the biosynthesis of bacterial cell envelope polysaccharides such as peptidoglycan (PG)[44], wall teichoic acids (WTA)[5], and lipopolysaccharide (LPS) O-antigen[45], amongst others. These cell envelope polysaccharides are vital for the survival of bacteria, as well as for their pathogenesis; thus, the metabolism of the carrier lipid has long been held as a promising target for the development of novel therapeutics, and in recent years there

have been a number of papers published that describe inhibitors of both the synthesis and recycling of the most common bacterial carrier lipid, undecaprenyl phosphate ($C_{55}P$).

Biosynthesis of $C_{55}P$ occurs in the cytoplasm through the sequential condensation of isopentyl pyrophosphate (C_5PP) with farnesyl pyrophosphate ($C_{15}PP$), generating undecaprenyl pyrophosphate ($C_{55}PP$)(Fig. 1.2A,B). Before being used in any cell envelope polysaccharide biosynthetic pathway, $C_{55}PP$ must first be converted to its active $C_{55}P$ form[10]. It should be noted that while the 55-carbon chain length is the most common bacterial lipid carrier length, there are examples of shorter chains such as decaprenyl phosphate ($C_{50}P$) in mycobacterial species[46]. Upon conversion to $C_{55}P$, sugar moieties are transferred to the carrier lipid by membrane embedded or membrane associated enzymes using nucleotide-activated sugar precursors (Fig. 1.2B). Once the cell envelope polysaccharide has been appropriately assembled on $C_{55}P$, the lipid-linked oligosaccharide (LLO) is flipped by a specific LLO flippase from the cytoplasmic leaflet of the plasma membrane to the periplasmic leaflet[47]. The length of these polyprenyl phosphate lipids is greater than the width of a typical bacterial plasma membrane, and the lipids are thought to likely kink into a U- or L-shaped conformation in order to be accommodated[48]. There have been some biophysical studies that have shown that the presence of these lipids and their derivates, presumably due to their length and unique geometry, can increase the fluidity and ion permeability of model membranes, and their presence could very well have an effect on translocation *in vivo*[49–51].

Once flipped to the periplasmic leaflet of the plasma membrane, the glycan moieties of the LLOs become accessible to substrate specific enzymes that catalyze their covalent transfer to specific glycan acceptors in the periplasmic space[10]. In most cases, the glycosyltransfer reaction results in the release of $C_{55}PP$ as a byproduct, which then needs to be recycled to its

$C_{55}P$ form and translocated back to the cytoplasmic leaflet of the membrane for a subsequent round of cell envelope polysaccharide biosynthesis (Fig. 1.2B)[10]. In some cases, such as in the biosynthesis of WTAs, a phosphate is transferred along with the glycan moiety, resulting in the generation of $C_{55}P$ rather than $C_{55}PP$, obviating the need for a recycling step[5]. Due to the metabolic cost of producing $C_{55}PP$, its recycling back to $C_{55}P$ following each round of bacterial cell envelope polysaccharide biosynthesis constitutes an important step in the metabolism of all bacteria and has been studied extensively in recent years. While the number and identity of intramembranal phosphatases varies from species to species, there have been four genes identified in *Escherichia coli* that encode enzymes that display $C_{55}PP$ phosphatase activity[52–54].

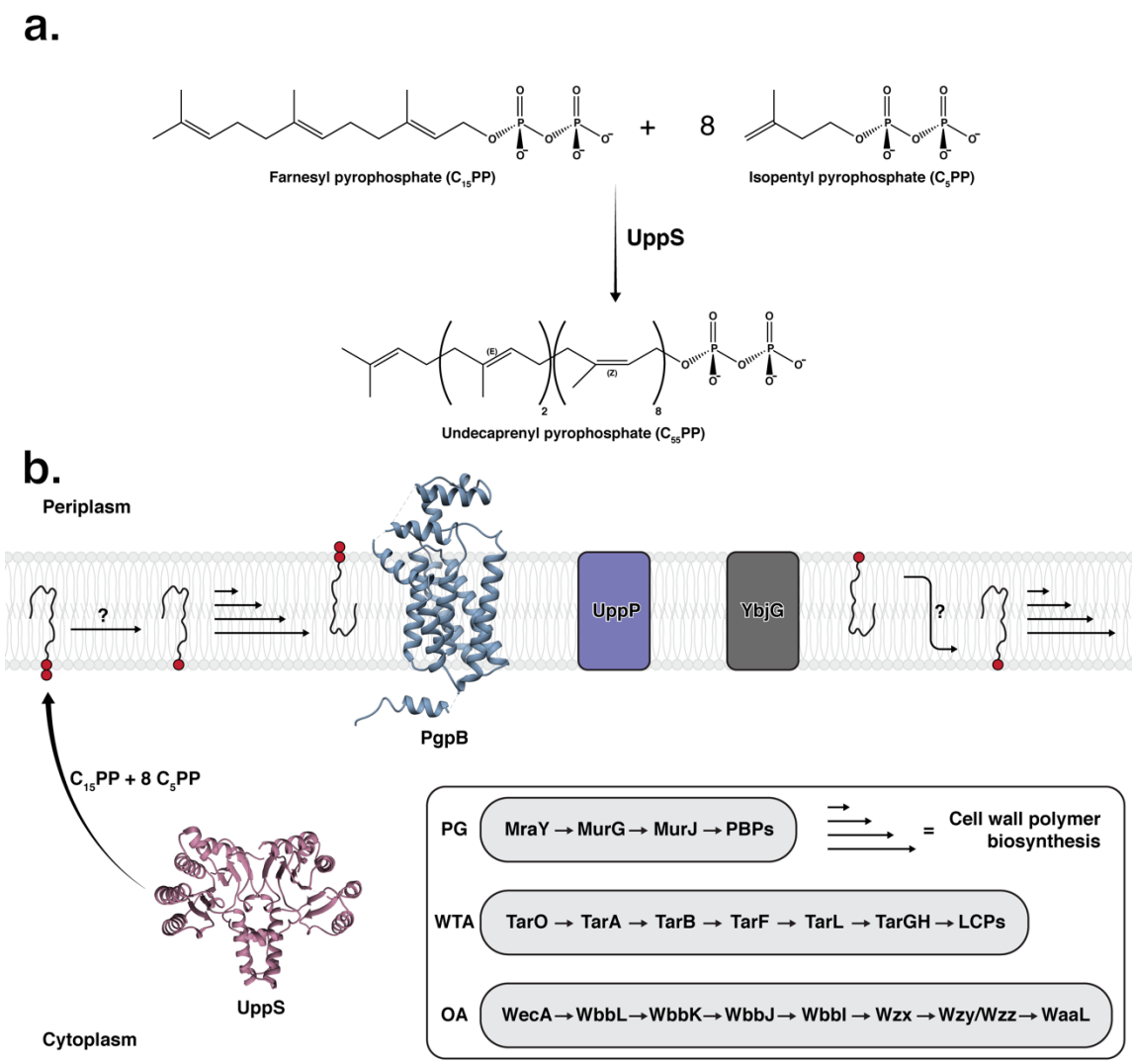


Figure 1.2 Metabolism of the bacterial cell wall carrier lipid.
a. The overall reaction catalyzed by UppS to generate C_{55}PP **b.** Schematic view of the metabolism of the bacterial cell wall carrier lipid. *De novo* synthesis of C_{55}PP occurs in the cytoplasm before partitioning into the cytoplasmic leaflet of the inner membrane. C_{55}PP is dephosphorylated to C_{55}P by an as yet unknown mechanism before acting as a scaffold for the assembly of cell wall polymers such as PG, WTAs, or OA (enzyme catalyzed biosynthetic steps shown in inset). Upon translocation to the periplasmic leaflet of the inner membrane by LLO flippases and incorporation of the polymer subunit into the existing cell wall, C_{55}PP is released as a by-product and recycled back to C_{55}P by UppP and the PAP2 phosphatases (PgpB and YbjG). Finally, C_{55}P is translocated back to the cytoplasmic leaflet of the inner membrane by an unknown mechanism where it can used in another round of cell wall polymer synthesis.

What remains elusive is the identity of the flippase that could translocate the carrier lipid back to cytoplasmic leaflet following its transformation from C_{55}PP to C_{55}P in the periplasm. It

has been known for some time that the rate of translocation of polyprenyl lipids between leaflets of a lipid bilayer is too slow to support a model of uncatalyzed translocation, with biosynthesis of peptidoglycan in *E. coli* turning over approximately 5,000 molecules of $C_{55}P$ per second in the form of lipid II[55–57]. Such a high rate of consumption for just one of the many cell envelope polysaccharides implies that recycling must indeed play an important role. It has been speculated that the LLO flippases responsible for translocating cell envelope polysaccharide precursors could play a role in the translocation; however, there has yet to be any work directly showing this. In addition to the questions regarding the identity of the putative $C_{55}P$ flippase, there is the fact that $C_{55}PP$ must be dephosphorylated to $C_{55}P$ in the course of *de novo* carrier lipid synthesis. All characterized $C_{55}PP$ phosphatases have been shown to act on the periplasmic face of the plasma membrane, leaving the possibilities that an as yet unidentified phosphatase is acting on the cytoplasmic face of the plasma membrane, or that a flip-flop type mechanism of $C_{55}P(P)$ regulation exists in which *de novo* synthesized $C_{55}PP$ is somehow translocated to the periplasmic leaflet of the membrane for conversion into $C_{55}P$ before being flipped back to the cytoplasmic membrane for use in cell envelope polysaccharide biosynthesis.

1.3.1 Biosynthesis of undecaprenyl pyrophosphate

Isoprenoids are a large class of naturally occurring chemicals which are derived from the universal five-carbon precursor isopentyl pyrophosphate (C_5PP) or its isomer dimethylallyl diphosphate (DMAPP). These precursors can be synthesized via two independent pathways: the classical mevalonate pathway, which is used exclusively in animals and archaea, or the alternative 2-*C*-Methyl-$_D$-erythritol 4-phosphate (MEP) pathway common to most bacteria [58]. Upon synthesis of these precursor molecules, a family of enzymes known as prenyltransferases

first catalyze their condensation into geranyl pyrophosphate (C_{10}PP), which can be used to synthesize various monoterpenes, and subsequently farnesyl pyrophosphate (C_{15}PP), which can be used as a building block for the synthesis of important molecules such sterols, sesquiterpenes, and carotenoids, among others[59]. These prenyltransferases catalyze the condensation reaction of an allylic primer molecule of varying length with differing numbers of C_5PP and can be classified based on the stereochemical outcome of the reaction that they catalyze as either *cis*- or *trans*- prenyltransferases (Fig. 1.3). In addition to their stereospecificity, each prenyltransferase acts upon a specific allylic primer molecule and carries out a specific number of condensation reactions. For example, farnesyl pyrophosphate synthase (FppS), which is the prototypical *trans*-prenyltransferase, catalyzes the condensation of DMAPP and two C_5PP molecules to generate *all-trans*-C_{15}PP[60]. Until recently it was thought that all prenyltransferases carried out "head-to-tail" 1'-4 condensation reactions and that the length of the product generated was determined by the size of a central hydrophobic tunnel into which the product is polymerized; however, in recent years there have been several "head-to-middle" prenyltransferases that can generate branched polyprenyl products, such as lavandulyl diphosphate synthase (LppS) and isosesquilavandulyl diphosphate synthase (Mcl22)[61–63]. Interestingly, the crystal structure of Mcl22 revealed that rather than having the polyprenyl product polymerize into a central hydrophobic tunnel, it instead occupies a surface pocket that is oriented 90° to the hydrophobic tunnel[63].

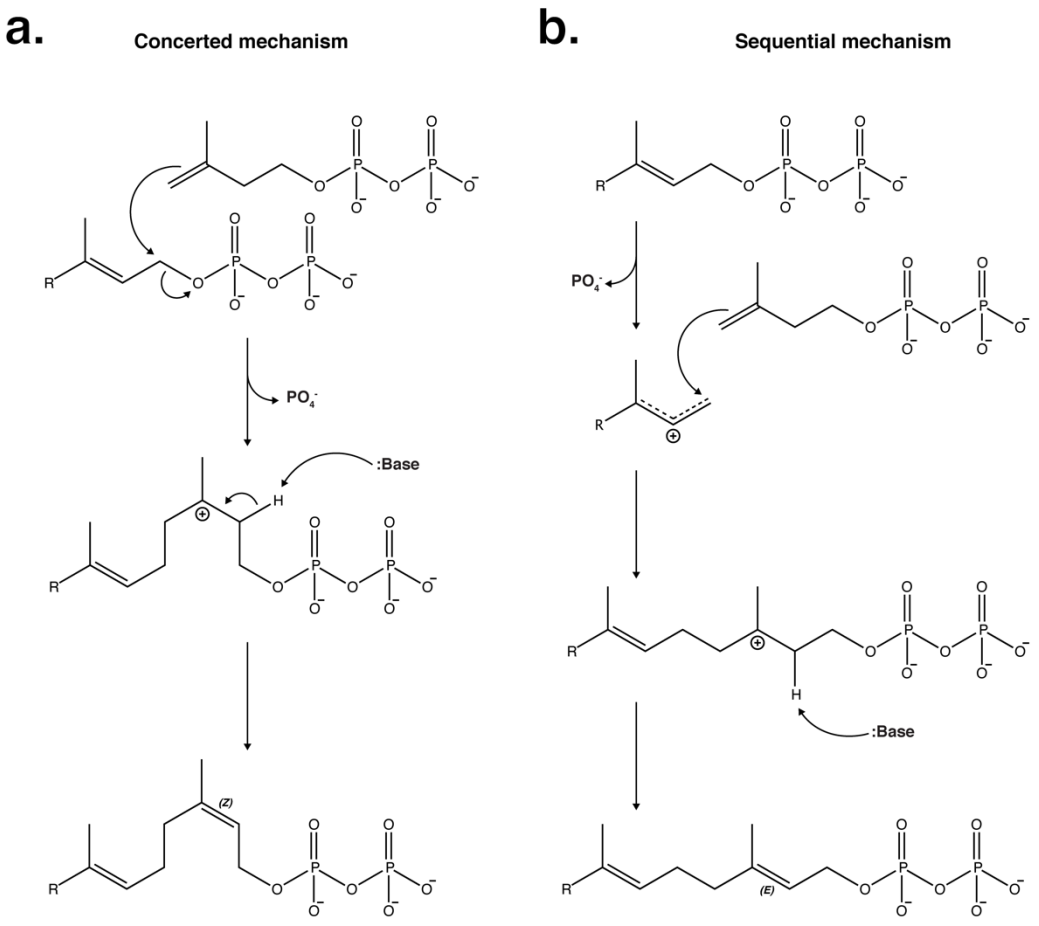


Figure 1.3 Catalytic mechanisms of prenyltransferase enzymes.

In bacteria, the biosynthesis of C_{55}PP is carried out by undecaprenyl pyrophosphate synthase (UppS), a *cis*-prenyltransferase that catalyzes the condensation of one molecule of C_{15}PP with eight molecules of C_5PP, generating di-*trans*-octa-*cis*-undecaprenyl pyrophosphate (Fig. 1.2A)[60]. The use of such a carrier lipid in the biosynthesis of bacterial cell envelope polysaccharides was proposed as far back as 1965 when Anderson *et al.* and Wright *et al.* identified lipid linked intermediates in peptidoglycan and O-antigen synthesis, respectively[64,65]. By 1967, Allen *et al.* had isolated an enzyme from *Micrococcus lysodeikticus* that could produce long polyprenyl chains using C_{15}PP and C_5PP as substrates, and it was subsequently shown by multiple groups that the product was C_{55}PP[66,67]. Throughout

the 1970s and 1980s, a number of UppS homologues were partially purified and characterized, demonstrating the stereospecificity of the enzyme[68], as well as its dependence on a divalent cation for activity (with a preference for Mg^{2+})[69,70]; however, it wasn't until 1998 that the gene encoding UppS from *Micrococcus luteus* (*Ml*UppS) was cloned, providing the first amino acid sequence of a *cis*-prenyltransferase and showing they share very little sequence homology with previously characterized *trans*-prenyltransferases[71]. It had previously been proposed that UppS was a membrane associated enzyme due to the apparent requirement of detergents during its purification[70]. Although cloning and overexpression of *Ml*UppS in *E. coli* allowed for purification to be achieved to near homogeneity in the absence of surfactants, it is still likely that UppS exists in close proximity to the plasma membrane *in vivo* as *in vitro* experiments have shown that the presence of Triton X-100 increases the dissociation rate of C_{55}PP[72,73]. While Triton X-100 can spuriously increase reaction rates of various unrelated enzymes *in vitro* by increasing stability or preventing surface adsorption, spatial proximity to the cytoplasmic face of the plasma membrane would make sense as it would allow for rapid partitioning of the hydrophobic C_{55}PP out of the cytoplasmic milieu.

Since it was originally cloned from *M. luteus*, a number of bacterial homologues of UppS have been identified via sequence homology, cloned, and biochemically characterized. In addition, structural characterization in the form of a number of crystal structures in both apo- and substrate-bound forms have allowed further valuable insight into the molecular mechanisms by which these enzymes work. The *E. coli* homologue of UppS (*Ec*UppS), the most thoroughly characterized variant, is a 28 kDa protein that forms a prominent dimer in solution (Fig. 1.4A). The core of the protein is composed of a six-stranded β-sheet organized in a β3-β4-β2-β1-β5-β6 topology, surrounded by helices α1-4 and α7. The majority of conserved residues map to β1-4

and α2/3, which collectively form a 30 Å long hydrophobic tunnel into which the nascent C_{55}PP product grows during rounds of C_5PP incorporation (Fig. 1.4B)[74]. The dimer interface is primarily composed of a coiled-coil structure formed between α5 of each monomer, though there does not seem to be any functional implications of dimer formation[74]. The entrance of the hydrophobic tunnel is decorated with a number of conserved arginine residues that have been implicated in the coordination of C_5PP (Arg194, Arg200) and C_{15}PP (Arg30, Arg77), in addition to an absolutely conserved aspartate (Asp26) that has been shown to be critical for catalysis (Fig. 1.4C)[75]. This conserved entrance comprises the active site of UppS, with C_{15}PP occupying the so-called S1 position and C_5PP occupying the so-called S2 position (Fig. 1.4B)[76]. The S1 and S2 nomenclature refers to the binding sites of sulfate ions that were proposed to act as pyrophosphate mimics in the original crystal structure of *Ec*UppS[74,77]. When C_{15}PP is bound in the S1 position, there is a conformational reorientation of α3 towards its hydrocarbon tail, allowing for a number of hydrophobic residues to make contact with the allylic substrate (Fig. 1.4B)[77]. In addition to the shift of α3, a flexible loop connecting β2 to α3 that cannot be modeled in apo- crystal forms of UppS becomes ordered when the S1 position is occupied by substrate.

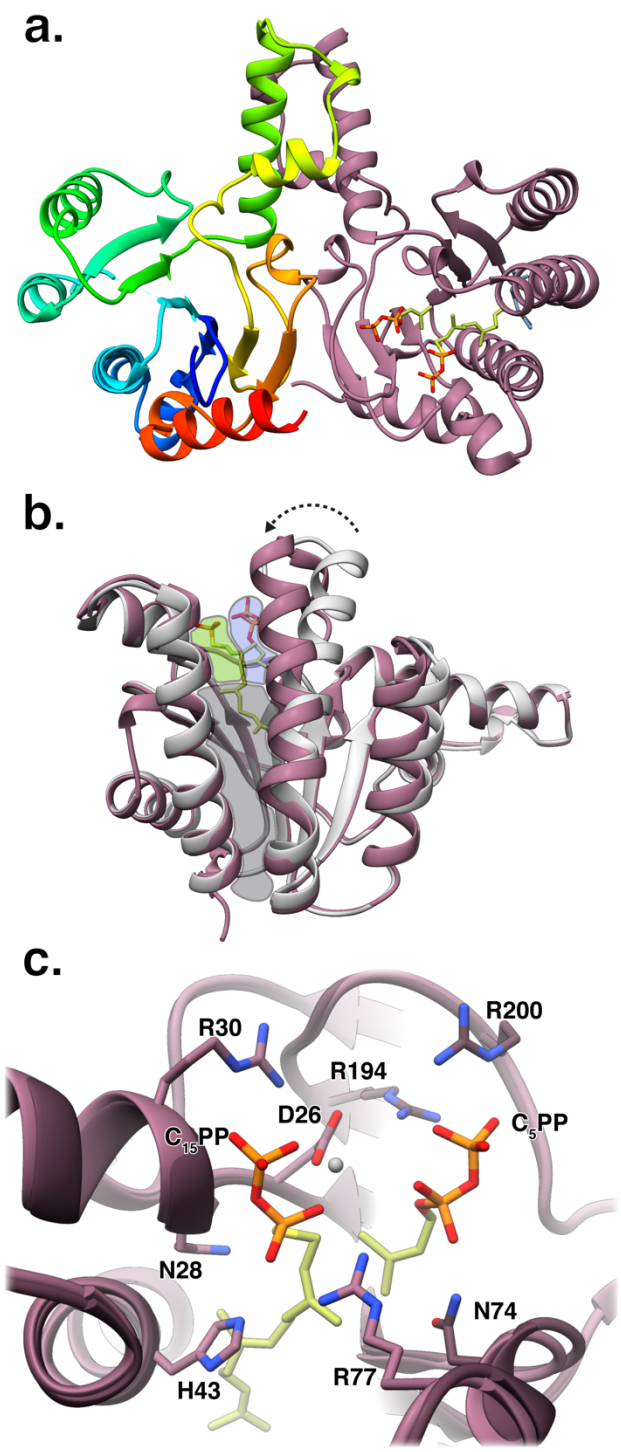


Figure 1.4 Structures of *Escherichia coli* undecaprenyl pyrophosphate synthase.
a. Cartoon representation of the *Ec*UppS dimer with one monomer colored rainbow from the N- (blue) to C-terminus (red) and the other in pink. Bound C_{15}PP and C_5PP (yellow) are shown in stick. **b.** Comparison of apo- (light grey) and substrate bound (pink) *Ec*UppS structures. Relative positions of the S1 (green) and S2 sites (blue), as well as the hydrophobic tunnel (grey) are shown by overlays. Reorientation of helix α3 upon substrate binding is indicated by the dashed arrow. **c.** Composite model of the *Ec*UppS active site generated separately from thio-C_{15}PP (PDB ID 1X06,

light) and D26A mutant C_5PP (PDB ID 1X09, dark) bound structures. Conserved active site residues are shown in stick, C_{15}PP/Asp26 coordinated Mg^{2+} shown as grey sphere.

The overall reaction catalyzed by UppS is common to all head-to-tail prenyltransferases: the C4-atom of C_5PP carries out a nucleophilic attack on the C1-atom of the allyl substrate that bears the pyrophosphate leaving group, followed by the stereospecific removal of a proton from the C_5PP C2-atom to generate a new double bond[78]. Prenyltransferase enzymes catalyze reactions using either a concerted or sequential mechanism (Fig. 1.3)[79]. In the concerted mechanism both nucleophilic attack and release of pyrophosphate occur simultaneously, whereas in the sequential mechanism, the allylic substrate is first ionized, leading to the release of pyrophosphate to form a carbocation intermediate that then undergoes nucleophilic attack. *Ec*UppS was shown to act as a *cis*-prenyltransferase by the use of two substrate analogs that were either deuterated at the C1 position ([1,1-$^2H_2$]-C_{15}PP), or had a proton replaced with a fluorine in the C2 position (2-fluoro-C_{15}PP)[73]. Using octaprenyl pyrophosphate synthase (OppS), a known *trans*-prenyltransferase, and the 2-fluoro-C_{15}PP substrate, the authors demonstrated that the fluorine's presence in the C2 position dramatically slowed its activity, consistent with the electron withdrawing effects of the C2-flourine slowing down the formation of a carbocation intermediate. In contrast, UppS was able to use the 2-fluoro-C_{15}PP as an alternative substrate with little to no effect on the C_5PP condensation rate. In addition, no secondary kinetic isotope effect (KIE) was measured when [1,1-$^2H_2$]-C_{15}PP was used as the substrate with UppS, suggesting that there is no discrete formation of a C_{15}PP carbocation intermediate[73]. In *trans*-prenyltransferases, a pair of DD*XX*D motifs (where X is any residue) coordinate the Mg^{2+} ions that facilitate the elimination of the pyrophosphate group from the allylic substrate[80]; however, in *cis*-prenyltransferases, the mechanism for catalysis was less clear. Fluorescence binding

studies demonstrated that Mg^{2+} was not required for binding of C_{15}PP, whereas C_5PP binding and enzyme activity absolutely depended on the presence of the metal ion[81].

Crystal structures of both wild-type and D26A *Ec*UppS in complex with C_5PP or C_5PP and C_{15}(S)PP (a catalytically inactive substrate analog), provided an explanation of the role of Mg^{2+} and the conserved Asp residue (Fig. 1.4C)[77]. In the wild-type *Ec*UppS C_5PP/C_{15}(S)PP ternary complex structure, the entire C_{15}(S)PP molecule in the S1 position can be modeled, but poor density of the C_5PP in the S2 position only allowed for modelling of a single phosphate. In this wild-type structure, the Mg^{2+} ion is clearly coordinated by the pyrophosphate moiety of C_{15}(S)PP and the carboxylate side chain of Asp26[77]. In the D26A *Ec*UppS C_5PP/C_{15}(S)PP ternary complex, the entire C_{15}(S)PP can be modeled, but the density of C_5PP is again poor, only allowing for the modelling of the hydrocarbon tail therein; however, in the D26A *Ec*UppS C_5PP complex, both the S1 and S2 positions are occupied by C_5PP molecules with the Mg^{2+} coordinated by the C_5PP in the S2 position[77]. These crystal structures, along with the fluorescence binding studies, allow us to envision a catalytic cycle for UppS where C_{15}PP binds to the S1 position prior to Mg^{2+} coordinated C_5PP at the S2 position and with the carboxyl group of Asp26 assisting the migration of Mg^{2+} from C_5PP to C_{15}PP. It is only once the Mg^{2+} has been transferred to C_{15}PP that the pyrophosphate group can leave and the attack by C_5PP can occur. Following the sequential addition of eight molecules of C_5PP to C_{15}PP, the full length C_{55}PP is proposed to sterically push against the flexible loop that becomes ordered upon substrate binding, reversing the conformational change that brought α3 in closer proximity to the hydrocarbon tail of C_{15}PP and allowing for product release[77]. Finally, the length of the polyprenyl lipid product appears to be controlled by the size of the hydrophobic tunnel into

which it is polymerized, with experiments showing that a L137A mutation, which reduces the steric bulk at the base of the tunnel, leads to the synthesis of longer chain products[74].

Because C_{55}PP plays such a central role in the biosynthesis of all bacterial cell envelope polysaccharides, it is unsurprising that UppS is an essential enzyme and has thus presented an appealing target for the development of novel antibiotics. There have been several classes of small molecule inhibitors of UppS that have been identified by rational design based on inhibitors of closely related enzymes in eukaryotes, as well as high-throughput *in vitro* and *in silico* screening approaches[31,35,37,82]. Crystal structures of *Ec*UppS and *Streptococcus pneumoniae* UppS (*Sp*UppS), as well as the closely related *Mycobacterium tuberculosis* decaprenyl diphosphate synthase (*Mt*DppS), in complex with different inhibitors have revealed multiple binding sites at the entrance of the hydrophobic tunnel where inhibitors compete with C_{15}PP and C_5PP binding, as well as in the hydrophobic tunnel itself where the inhibitors block the elongation of the C_{55}PP product[26,31–35,38,83,84]. Recent reports of phenylthiazole inhibitors that are effective against a broad range of pathogens (including MRSA and VRSA) that appear not to promote the development of resistance are of particular interest[28,29,39].

1.3.2 Dephosphorylation of undecaprenyl pyrophosphate

In addition to the discovery that a polyprenyl phosphate lipid was used as a universal carrier lipid in the biosynthesis of bacterial cell envelope polysaccharides, it was found quite early on that the carrier lipid was synthesized in its pyrophosphate form, and that the pyrophosphate form was released as a product from enzymes that catalyzed the incorporation of sugar moieties into the existing cell envelope[85]. It was clear that dephosphorylation of C_{55}PP would need to occur both in the course of *de novo* synthesis of C_{55}P, as well as in the recycling

of $C_{55}PP$ into $C_{55}P$, and in 1972, Goldman and Strominger reported the successful purification and biochemical characterization of an enzyme that displayed $C_{55}PP$ phosphatase activity, though its identity remained unknown[86]. Even earlier, in 1967, Siewert and Strominger demonstrated that the dephosphorylation of $C_{55}PP$ into $C_{55}P$ could be directly inhibited by the antibiotic mixture known as bacitracin (produced by *Bacillus licheniformis*), and in 1971 demonstrated that its inhibitory effects were achieved by a direct complex formed between bacitracin A (the most potent of the cyclic peptides that made up the antibiotic mixture), $C_{55}PP$, and a metal ion, thus sequestering $C_{55}PP$ from phosphatases[19,87]. In 1993, this knowledge was used to great effect when Cain *et al*. identified a gene that, when overexpressed, conferred resistance to bacitracin[88]. The gene, designated *bacA,* was initially thought to encode a membrane-bound undecaprenol ($C_{55}OH$) kinase based on sequence similarity to previously reported bacterial lipid kinases.

BacA, now referred to as UppP (for undecaprenyl pyrophosphate phosphatase), was shown to act as $C_{55}PP$ phosphatase rather than a $C_{55}OH$ kinase by El Ghachi *et al.* in 2004[52]. The authors confirmed that overproduction of UppP leads to increased bacitracin resistance, as well as a 280-fold increase in specific activity of $C_{55}PP$ phosphatase activity in membrane extracts compared to control cells carrying empty plasmid. Mass spectrometry confirmed that the product generated by DDM purified UppP acting on $C_{55}PP$ was of the expected mass of $C_{55}P$ (845.6 Da), and an MraY-coupled assay showed that 100% of the C55P could be converted to lipid I in the presence of MraY and UDP-MurNAc-pentapeptide. Curiously, it was also shown that a functional copy of the *uppP* gene was not required for growth of *E. coli* in LB media, despite its deletion resulting in a 75% decrease in overall $C_{55}PP$ phosphatase activity compared to wild-type *E. coli*. Contemporaneous reports showed that despite its apparent disposability

when grown in LB media, Δ*uppP S. aureus* and *S. pneumoniae* strains showed attenuated virulence in a mouse model of infection, and Δ*uppP M. tuberculosis* strains exhibited impaired biofilm formation[22,23]. More recent reports have gone on to show that a functional *uppP* gene is also required for biofilm formation in *Streptococcus mutans* and colonization of the stinkbug gut by *Burkholderia* symbionts[24,89]. The residual 25% C_{55}PP phosphatase activity in Δ*uppP E. coli* suggested that there must be one or more additional enzymes that could act to convert the pyrophosphate form of the carrier lipid to its monophosphate form.

Work done in 1992 had identified a membrane-bound phosphatase in *Sulfolobus acidocaldarius* that was proposed to act as a DolPP phosphatase and Edman degradation had been used to determine the N-terminal sequence of the protein[90]. Using this information, El Ghachi *et al.* identified the gene encoding this phosphatase and determined that it encoded a 220 amino acid integral membrane protein that belonged to the phosphatidic acid phosphatase 2 (PAP2) superfamily of enzymes[53]. From there, they carried out a BLAST search and identified BcrC from *B. licheniformis* and *B. subtilis* (at the time known as YwoA), as well as YbjG from *E. coli* as top hits. At the time, the function of BcrC had not been deduced and it was proposed to play an accessory role to the BcrAB ABC transporter in *B. licheniformis* that acts as a bacitracin efflux pump; however, it has since been shown to act as a C_{55}PP phosphatase as its overproduction results in a 600-fold increase in cellular C_{55}PP phosphatase activity with a concomitant increase in bacitracin resistance[91]. YbjG, the overproduction of which also results in increased bacitracin resistance, and BcrC, are both members of the PAP2 superfamily of enzymes. Two additional genes encoding PAP2 enzymes, *lpxT* (formerly *yeiU*, and of unknown function at the time) and *pgpB* (encoding a phosphatidyl glycerol phosphatase), were also identified in *E. coli*, and the absence of any sequence homology between UppP and these PAP2

enzymes clearly indicated that they belonged to separate families. Bioinformatic analysis indicated that *ybjG, lpxT, and pgpB* all encoded integral membrane proteins. As with *uppP*, individual disruptions of chromosomal copies of *ybjG*, *lpxT*, and *pgpB* could easily be obtained without affecting the growth of *E. coli*; however, inactivation of *uppP*, *ybjG*, and *pgpB* in combination could not be obtained, suggesting that the triple mutation had a lethal effect. When a *ybjG*/*pgpB* double mutant was grown with a temperature dependent plasmid carrying the *uppP* gene, the cells were able to grow normally, but a shift to the restrictive temperature for *uppP* expression resulted in a buildup of PG nucleotide precursors and cell lysis[53]. This result was consistent with the notion that the PAP2 enzymes identified in *E. coli* could contribute to the 25% residual C_{55}PP phosphatase activity previously observed in Δ*uppP E. coli* and indicated that the C_{55}PP phosphatase activity of LpxT was not sufficient to allow cell growth[52,53]. Indeed, overexpression of *ybjG*, *lpxT*, and *pgpB* genes resulted in 5-, 10-, and 70-fold increases in C_{55}PP phosphatase activity, respectively, with accompanying increases in bacitracin resistance, as was seen with *bcrC* in *Bacillus* species[53]. These data indicated that both UppP and members of the PAP2 superfamily were able to catalyze the dephosphorylation of C_{55}PP into C_{55}P. UppP and its central role in C_{55}PP phosphatase activity will be explored in depth in the next section of this introduction.

The PAP2 superfamily of enzymes is distributed widely throughout all three kingdoms of life, though homology between family members is mostly restricted to three sequence motifs that typify the family, designated C1 (K-X_6-RP), C2 (PSGH), and C3 (SR-X_6-H-X_3-D)[92]. This family contains both soluble and integral membrane proteins and is comprised of various types of enzymes including nonspecific acid phosphatases, a number of lipid phosphatases, vanadium peroxidases, and other enzymes of uncharacterized activity[93–95]. The three PAP2 enzymes in

E. coli that have been shown to play a role in C_{55}PP phosphatase activity have been characterized to differing degrees. YbjG's role in C_{55}PP dephosphorylation was probed using a fosmidomycin sensitivity assay[54]. Fosmidomycin specifically targets *de novo* C_{55}PP synthesis by inhibiting the formation of MEP, the precursor molecule of both IPP and DMAP. It was reasoned that in comparison to the parental strain, *E. coli* mutants deficient in recycling C_{55}PP phosphatase activity would display increased sensitivity to subinhibitory fosmidomycin concentrations. It was shown that Δ*pgpB* mutants displayed no phenotype in the presence of fosmidomycin, whereas Δ*uppP*, Δ*ybjG*, mutants and a double Δ*uppP*/Δ*ybjG* mutant displayed significant growth defects in comparison to the parental strain. From this, the authors proposed that YbjG is likely to be primarily involved in C_{55}PP recycling, while working under the assumption that UppP acts as the C_{55}PP phosphatase in *de novo* synthesis. LpxT was identified as a C_{55}PP phosphatase in the same study as YbjG and PgpB, but a subsequent study using ^{32}P-labelled lipid A, a precursor of mature LPS, showed that LpxT actually acts as a lipid A 1-phosphotransferase to generate lipid A 1-diphosphate using C_{55}PP as its phosphate donor[96]. Identification of C_{55}PP as the phosphate donor was carried out *in vitro* by using purified LpxT and Kdo_2-[4'-^{32}P]-lipid A as an acceptor. Confirmation of C_{55}PP as the phosphate donor was confirmed *in vivo* by using bacitracin to inhibit formation of the lipid A 1-diphosphate species in wild-type bacteria.

PgpB, the third PAP2 enzyme shown to dephosphorylate C_{55}PP in *E. coli*, has been more thoroughly characterized. A kinetic analysis of the C_{55}PP phosphatase activity of PgpB was carried out using purified PgpB, and it was shown to catalyze C_{55}PP dephosphorylation with a specific activity of 3.0 μmol min^{-1} mg^{-1}[97]. As previous work had shown that PgpB could remove both the β- and α-phosphates from diacylglycerol pyrophosphate (DGPP), its ability to remove the terminal phosphate from C_{55}P to generate undecaprenol (C_{55}OH) was investigated;

however, none was detected[97,98]. Optimal catalysis of C_{55}PP dephosphorylation was achieved between pH 6.5 and 7.5, though the reaction rate was reduced by less than 50% throughout the rest of the pH range tested. PgpB, while active on C_{55}PP, displays a broad substrate specificity with an increased specific activity observed when phosphatidic acid (PA) or DGPP are used as substrates rather than C_{55}PP (~2- and 10-fold, respectively). Interestingly, PgpB displayed a dose dependent increase in C_{55}PP phosphatase activity upon addition of phospholipids, perhaps reflecting a decrease in entropy brought on by interactions between the hydrocarbon tails of the phospholipids and C_{55}PP. Also of note is that, despite PgpB's apparent preference for DGPP based on specific activity, addition of DGPP to the reaction did not result in inhibition of C_{55}PP phosphatase activity, suggesting that C_{55}PP is a preferred substrate for the enzyme under the conditions tested[97].

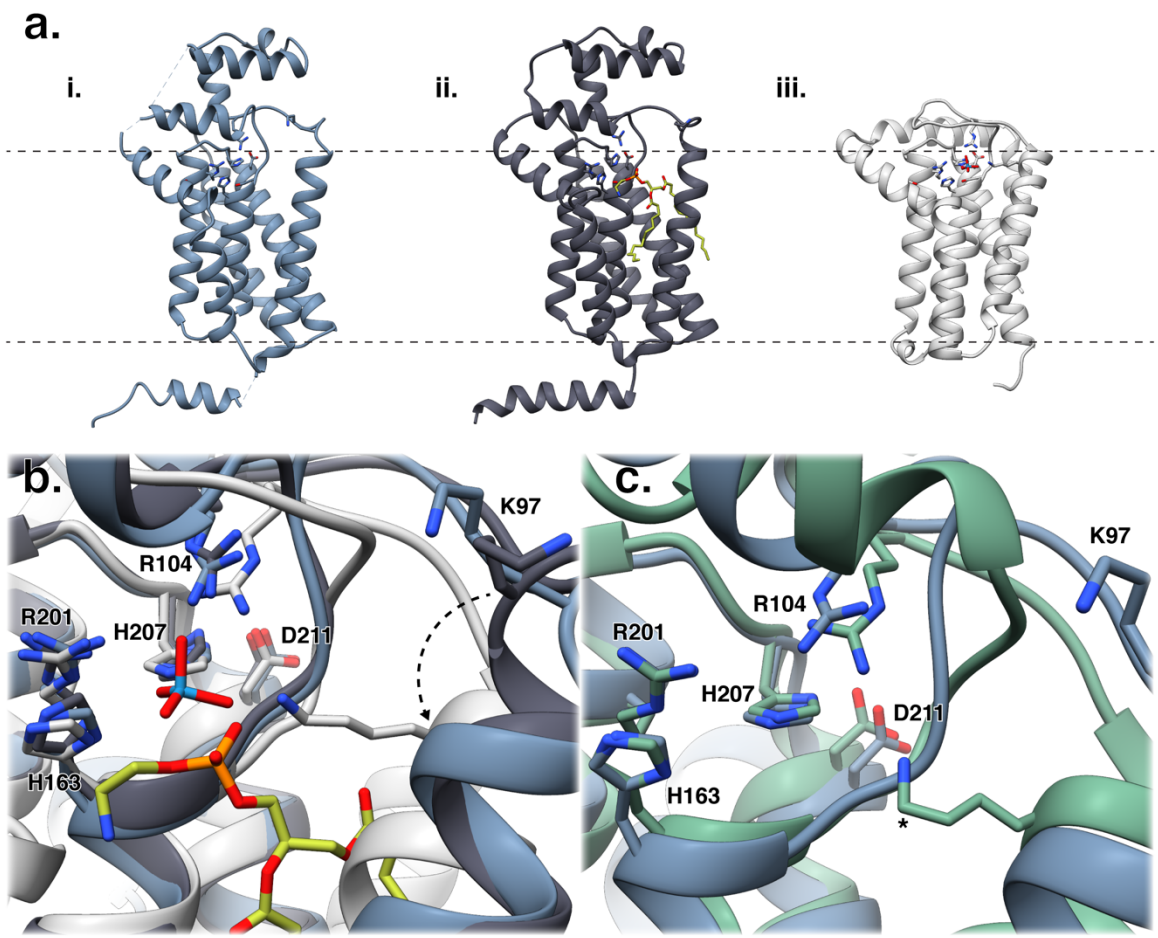


Figure 1.5 Structures of PAP2 family C_{55}PP phosphatases.
a. i. apo-*Ec*PgpB (blue, PDB ID 4PX7), ii. PE-bound *Ec*PgpB (slate, PDB ID 5JWY), iii. WO_4^{2-}-bound *Bs*UppS (light grey, PDB ID 5JKI). The C1/C2/C3 motif residues are shown in stick. Membrane boundary is indicated with a dashed line. **b.** Overlay of PAP2 family C_{55}PP phosphatase active sites. WO_4^{2-} is shown in light blue/red, PE is shown in yellow. The deviation in position of *Bs*PgpB Lys97 (*Ec*PgpB numbering) is indicated with a dashed arrow. **c.** Comparison of integral membrane *Ec*PgpB and soluble *Sb*NSAP (mint, PDB ID 1EOI) active sites. *Sb*NSAP Lys97 (indicated by *) likely shows significant deviation from *Ec*PgpB because it was crystallized in complex with a transition state analog (molybdate, not shown).

As with the C_{55}PP synthase UppS, crystal structures of PAP2 enzymes have been indispensable in gaining an understanding of the mechanistic details of their action (Fig. 1.5A). The earliest PAP2 enzyme crystal structures were of a soluble vanadium-containing chloroperoxidase from *Curvularia inaequalis* (*Ci*CPO) and a non-specific acid phosphatase from *Shimwella blattae* (*Sb*NSAP)[93,95]. These structures showed that despite their distance in

primary sequence, the three motifs (C1, C2, and C3) that are common to all PAP2 enzymes are in close proximity and form the conserved active site. The conserved architecture of these enzymes was shown to extend to the integral membrane protein members of the PAP2 family when the crystal structure of PgpB from *E. coli* (EcPgpB) was solved (Fig. 1.5A)[99]. Despite sharing an overall sequence identity of only ~15%, overlay of the core helices from the soluble *Sb*NSAP with α1-helix and TMs 4, 5, and 6 of EcPgpB showed a 1.4 Å RMSD over 60 Cα atom pairs. In *Ec*PgpB, the C1 and C2 motifs are located in the periplasmic space on the end of TM3 and on the region connecting TM5 and TM6, respectively. The C3 motif is located just below the estimated membrane boundary at the N-terminus of TM4; thus, the catalytic His residues (H163 and H207), as well as other conserved residues shown to be critical for activity are all located near the membrane-solvent interface, oriented towards the periplasmic space. A subsequent *Ec*PgpB- phosphatidylethanolamine (PE) co-structure revealed that Lys97, Arg104, and Arg201 create a phosphate binding site (Fig. 1.5B)[100].

A deviation in the position of Lys97 in the *Ec*PgpB crystal structures relative to equivalent Lys residues in PAP2 enzymes that were crystallized in complex with phosphate mimics suggests that, similar to UppS, PgpB uses an induced fit mechanism of substrate binding (Fig. 1.5B). It is proposed that this would allow for the observed substrate diversity of PgpB and optimal catalysis to be achieved upon structural rearrangement for each substrate. A crystal structure of *B. subtilis* PgpB in complex with tungstate (WO_4) was solved using lipid cubic phase (LCP) crystallization and no such deviation of its Lys97 equivalent was observed (Fig. 1.5A,B)[101]. This is potentially due to the presence of a WO_4, a phosphate mimic, in the active site, though the authors did carry out molecular dynamics simulations that did not indicate that a structural rearrangement would occur upon substrate binding. More work is required to

determine whether such an induced fit mechanism exists or not. When the active site of *Ec*PgpB is superimposed with that of *Sb*NSAP, both the main-chain and side-chain atoms of the catalytic residues overlay with one another, suggesting a conserved two-step phosphate hydrolysis mechanism (Fig. 1.5C)[99,102]. For *Ec*PgpB, it was proposed that His207 and Asp211 form a charge-relay system in which Asp211 primes His207 for nucleophilic attack on the phosphorous center of a bound phosphate group, leading to the formation of a phosphoenzyme intermediate. In the second step of the reaction, His163 activates a water for the hydrolysis of the phosphoenzyme intermediate, returning the enzyme to its resting state (Fig. 1.6).

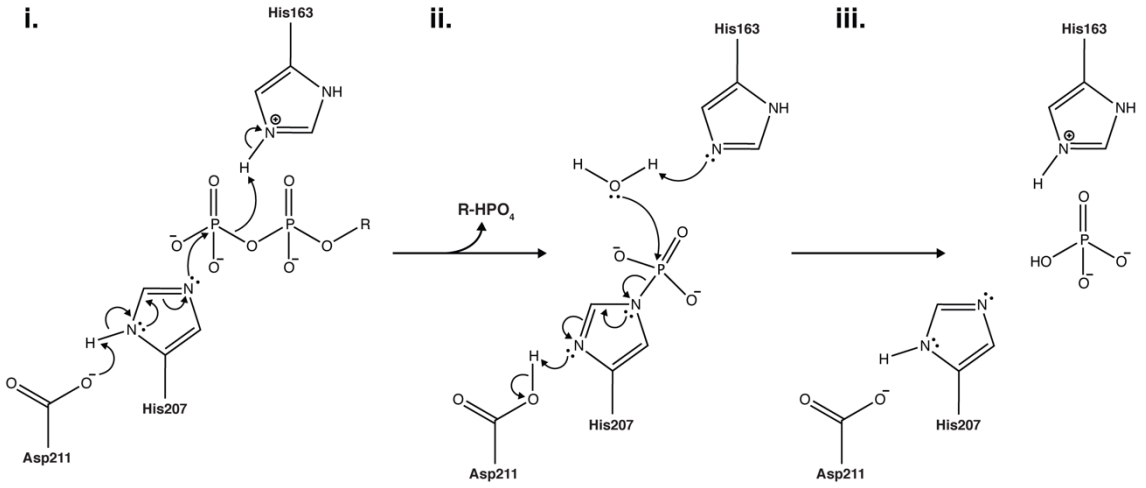


Figure 1.6 Catalytic mechanism of PAP2 family phosphatases.

In addition to C_{55}PP dephosphorylation, C_{55}P can be further metabolized to C_{55}OH[103]. While C_{55}OH has been known to make up a large portion of the undecaprenoid pool in a number of Gram-positive species such as *S. aureus*, it has never been detected in Gram-negative species. The enzyme responsible for converting C_{55}OH to C_{55}P has been known since 2003, when an *S. mutans* homologue of *E. coli* diacylglycerol kinase (DgkA) was identified and shown to exhibit C_{55}OH kinase activity, but until recently the identity of the C_{55}P phosphatase involved in C_{55}OH

generation was unknown[104]. Surprisingly, the *S. mutants* C_{55}OH kinase was shown to exhibit Mg^{2+} and ADP dependent phosphatase activity *in vitro*, with the same residues required for kinase activity implicated in phosphatase activity[105]. Importantly, it was also shown that expression of *S. mutans dgkA* in *E. coli* results in the *in vivo* production of C_{55}OH, as shown by TLC, HPLC, and mass spectrometry analysis. While not common to all bacteria, this mechanism of undecaprenoid sequestration provides Gram-positive bacteria with another regulatory mechanism for fine tuning the metabolism of the bacterial carrier lipid.

As stated above, it has been clear for some time that dephosphorylation of C_{55}PP would need to occur in the course of *de novo* synthesis as well as during the recycling of the lipid carrier following the incorporation of its glycan cargo into the bacterial cell envelope. Because *de novo* synthesis occurs in the cytoplasm, it is presumed that the newly synthesized C_{55}PP partitions into the cytoplasmic leaflet of the plasma membrane; however, the glycosyltransferase reactions that generate C_{55}PP during the course of cell envelope polysaccharide polymerization occur in the periplasm and presumably release the lipid carrier into the periplasmic leaflet of the plasma membrane. The identification of multiple enzymes in *E. coli* that can carry out C_{55}PP dephosphorylation allowed for the possibility that individual phosphatases could play distinct roles in *de novo* synthesis or recycling, with their active sites oriented towards the cytoplasm or periplasm respectively.

Before crystallographic data was available, the membrane topology of *E. coli* PAP2 proteins was probed by fusing PAP2 enzymes with reporter proteins that would be active only in the cytoplasm (e.g. GFP) or in the periplasm (e.g. β-lactamase). By creating a series of reporter protein insertions, or by serially truncating the proteins with a reporter fused to the C-terminus, it was demonstrated that PgpB, YbjG, and LpxT all appeared to have their active site residues

directed towards the periplasm, and in the case of *Ec*PgpB, subsequent crystal structures proved these predictions to be correct[54,97,99,100]. It was then proposed, in the absence of any topological data for UppP, that the PAP2 enzymes might act as C_{55}PP phosphatases in the course of recycling, while UppP would act on the cytoplasmic face of the membrane during *de novo* synthesis. This idea was shown to be incorrect when the membrane topology of UppP was determined using fused reporter proteins and it's conserved catalytic residues were also shown to be directed towards the periplasm[106]. This led to the hypothesis that there was either an as yet unidentified C_{55}PP phosphatase that could act specifically on the cytoplasmic face of the plasma membrane, or that *de novo* synthesized C_{55}PP required a "flip-flop" mechanism of dephosphorylation whereby the C_{55}PP would be flipped into the periplasmic leaflet of the membrane for dephosphorylation before being flipped back to cytoplasmic face of the membrane for use in cell envelope polysaccharide biosynthesis. It has been speculated that the lipid II flippase could play the role of C_{55}P(P) flippase or that the C_{55}PP phosphatases themselves could participate in the movement of the polyprenyl lipid's phosphate head group(s) across the membrane.

1.3.3 Undecaprenyl pyrophosphate phosphatase (UppP)

UppP is a 30 kDa polytopic integral membrane protein that is conserved throughout both Gram-negative and Gram-positive bacteria. The observation that *E. coli* UppP (*Ec*UppP) contributes 75% of the overall C_{55}PP phosphatase activity in *E.* coli is presumably indicative of UppP's importance throughout bacterial species[52]. Reflective of that is the fact that *Ec*UppP shares 63% identity with homologues in other Gram-negative species, and that homologues of *uppP* have been found in most sequenced bacterial genomes except for that of *Helicobacter pylori*,

which was recently shown to use a single PAP2 enzyme, HupA, as its main C_{55}PP phosphatase[107]. HupA accounts for more than 90% of C_{55}PP phosphatase activity in *H. pylori* but like UppP in *E. coli*, is dispensable for growth in LB media; however, HupA is required for stomach colonization in a mouse model infection, an observation that is consistent with the attenuated virulence and impaired biofilm formation observed in *S. pneumoniae* and *M. tuberculosis* Δ*uppP* mutants, respectively[22,23,107]. It is also interesting to note that while homologues of *uppP* are indeed found in almost all bacterial species, their contribution to C_{55}PP phosphatase activity can vary from species to species, and at different stages of bacterial growth. A recent study of carrier lipid metabolism *B. subtilis*, showed that *uppP* and *bcrC* are expressed at comparable levels during exponential growth, but that *uppP* expression is approximately 3-fold higher during stationary phase[108]. Cell envelope stress caused by bacitracin exposure led to upregulation of *bcrC* alone, rendering BcrC the more important C_{55}PP phosphatase during bacterial growth; however, Δ*uppP* mutants were deficient in sporulation, an important adaptive response to nutritional stress. It remains to be seen whether UppP is important for proper sporulation in all spore forming bacteria, but these data suggest that, despite UppP's remarkable conservation, other C_{55}PP phosphatases can be of greater importance during different stages of bacterial life cycles[108].

Due to its outsized contribution to the overall pool of C_{55}PP phosphatase activity in *E. coli*, *Ec*UppP has been studied in great depth. Bioinformatic analyses identified two highly conserved regions that have been implicated in substrate binding and catalysis[109]. Region I (residues 17-30) contains two conserved motifs: E/Q-X_3-E and P-V/I-SS-X_2-H. Region II (residues 170-178) contains a conserved PG-X-SRS-X_2-T motif. Expression and purification of *Ec*UppP allowed for biochemical characterization of its requirements for catalysis. Specifically, *Ec*UppP was shown to

display optimal activity at pH values between 5.0 and 7.5 and was shown to require a divalent metal cation for activity, with a marked preference for Ca^{2+} over the Mg^{2+} required for C_{55}PP synthesis[109]. Site-directed mutagenesis studies done by two separate groups identified a number of residues within the conserved regions that were implicated in the catalytic activity[106,109]. Unsurprisingly, mutation of the most strongly conserved residues in the regions had strong negative effects on the activity of *Ec*UppP, with E21A, S27A, H30A, S173A and R174A appearing to have the largest effects; however, the two groups came to differing conclusions about the identity of the catalytic residue, with one group proposing His30 and the other proposing Ser27. It should be noted that the group who proposed His30 as the catalytic residue did not investigate Ser27 as a possible candidate[109]. In addition to probing the effects of point mutations *in vitro*, Manat *et al.* tested the ability of different UppP mutants to complement a thermosensitive *E. coli* Δ*uppP* mutant strain *in vivo* and found a good correlation between the two experiments. All point mutants of UppP other than E21A, S27A, and R174A were able to restore growth to the thermosensitive strain, indicating that these three residues play a central role in UppP's activity, and that His30 is likely not the catalytic residue. The substrate specificity of *Ec*UppP was probed using either thin layer chromatography with radiolabeled C_{55}PP or a Malachite green based phosphate detection assay. The specific activity of *Ec*UppP with the isoprenyl phosphate lipids C_{55}PP or C_{15}PP as substrate were approximately the same (~ 10 μmol min^{-1} mg^{-1}), and while its efficiency was approximately halved, DGPP also appeared to act as a relatively good substrate.

1.4 Structural characterization of membrane proteins

Membrane proteins are a class of biologically and medically important molecules that allow for cells to interact with their external environment through signal transduction pathways,

transport of substrates across membranes, and cell-adhesion. It is estimated that that 20-30% of all genes encode for membrane proteins, and these proteins comprise more than half of all drug targets; however, they currently represent less than 1% of unique structures deposited in the Protein Data Bank (PDB)[110,111].

Structural characterization of membrane proteins has historically been an arduous process due to difficulties associated with obtaining sufficient amounts of purified, stable sample. The determination of early membrane protein structures relied on their natural abundance in biological sources[112–114]. Overexpression can increase the overall yield and allow for incorporation of affinity tags to ease purification, but it can also lead to aggregation or misfolding due to the limited space for insertion of proteins into the plasma membrane, which puts a limit on theoretical maximum yield[115]. Once an appropriate expression strategy has been devised, the membrane proteins must be solubilized from the membranes into which they were expressed before they can be purified. This is an often-challenging hurdle to overcome and can involve the screening of a great many detergents. Depending on the proposed method of structural characterization, detergent solubilized membrane protein can be reconstituted into a number of different membrane mimetics (i.e. liposomes, nanodiscs, peptidiscs, amphipols, saposin lipid nanoparticles[116,117]) or they can remain in detergent micelles; however, different detergents are often required for the extraction and structure determination steps[118]. Increasingly, amphipathic copolymers are being used for solubilization of membrane proteins, as they form “native” nanodiscs that are composed entirely of the lipids from the organism that was used to express the protein[119]. This technique has the advantage of never exposing membrane proteins to detergents (which can have delipidating effects), but the most commonly used copolymer – styrene maleic acid (SMA) – is limited by a narrow useful pH range and its sensitivity to the presence of divalent cations. New copolymers are

being developed to overcome the drawbacks of SMA and the field is growing at a rapid pace[120]. The screening of both expression and solubilization strategies can be expensive and time consuming. It for this reason that high-throughput screening techniques such as fluorescence-detection size exclusion chromatography (FSEC) have been developed[121].

The field of membrane protein structural biology has grown immensely since the first near-atomic resolution crystal structure of an integral membrane protein was determined in 1985[112]. Nuclear magnetic resonance (NMR) has enabled the determination of over 100 membrane protein structures with the majority of them utilizing solution-state NMR, though increasing contributions are being seen from solid-state NMR techniques[113]. The field has benefitted greatly from the recent "resolution revolution" in cryogenic electron microscopy (cryo-EM)[122] afforded through the new generations of field emission gun enabled microscopes and rapid throughput direct electron detectors. Single particle cryo-EM eliminates the need to grow crystals and allows for the study of membrane proteins stabilized in the wide array of membrane mimetics noted above and even in in liposomes[123]. These advances in cryo-EM technology are similarly advancing the possible resolution boundaries of cryo-electron tomography, allowing for some samples the possibility of moderate to near-atomic resolution study of macromolecular complexes *in situ*[124,125]. Despite these advances, the applicability of cryo-EM for structural characterization of small proteins (outside of a few published unique cases[126,127]) remains limited and, while the number of membrane protein structures solved by single particle cryo-EM will likely continue to rise, the vast majority of membrane proteins structures at high resolution thus far have been solved by X-ray crystallography.

Protein crystallization is an empirical process that regularly requires enormous amounts of screening to identify suitable conditions that are often difficult to reproduce[128]. In order for

crystallization to occur, a highly pure monodisperse sample must be concentrated to the point that supersaturation of the growth solution can be achieved during screening trials. At a basic level, protein crystallogenesis is nothing more than a controlled precipitation event (Fig. 1.7), with the formation of a crystal lattice depending on favourable intramolecular contacts that are often mediated by the solution in which the crystallization experiment is being conducted[128]. The crystallization of membrane proteins is much the same, but with the added confounding factor of the detergent micelle. The detergent micelle solubilizes membrane proteins by forming a torus around the hydrophobic transmembrane domains, allowing the polar portions of the proteins to protrude into the aqueous environment[129]. It is these protrusions that form crystal contacts in the formation of membrane protein crystals, as the detergent micelle is too dynamic to form ordered crystal contacts[129]. Thus, crystallization of membrane proteins requires screening of detergents as an additional parameter to find conditions suitable for crystallogenesis. Specific techniques such as bicelle and lipid cubic phase (LCP) crystallography have been developed in order to overcome the barrier that detergents represent for formation of a crystal lattice[130,131].

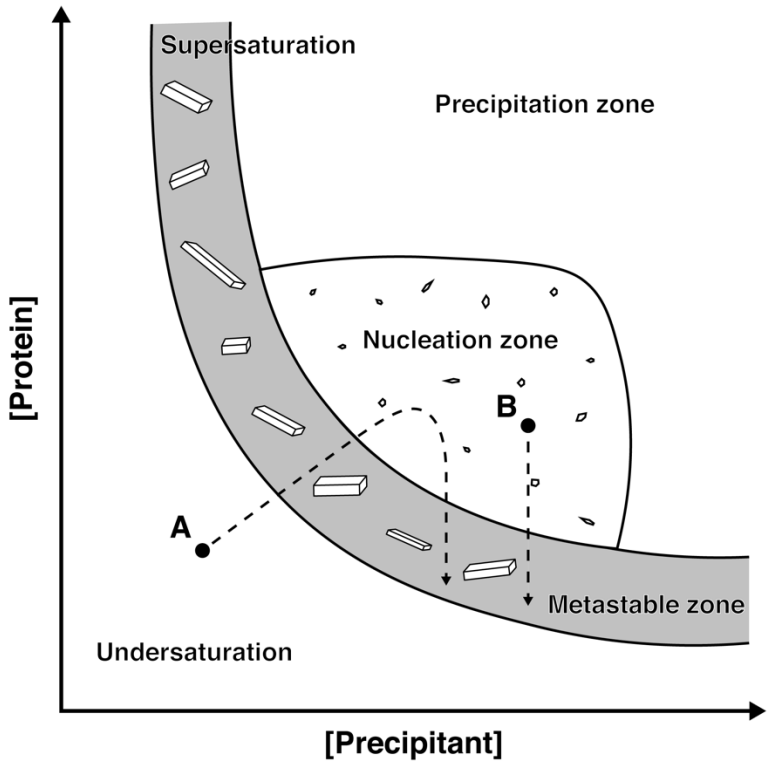


Figure 1.7 Schematic drawing a protein crystallization phase diagram.
The phase diagram is based on the two most commonly altered parameters in protein crystallization – the concentration of protein and precipitant. With vapour diffusion (A), the concentration of both protein and precipitant increase together until the nucleation zone is reached, the protein concentration in the drop decreases as crystals nucleate and grow. With LCP (B), or any batch crystallization technique, the crystallization experiment is set up such that the nucleation zone is the starting point, and protein concentration decreases as crystals nucleate and grow.

1.4.1 Expression of membrane proteins

Obtaining sufficient material for membrane protein crystallography is the initial barrier that must be overcome for structural characterization. While eukaryotic expression systems such as yeast or insect cells are most commonly used for the expression of eukaryotic proteins, expression of prokaryotic membrane proteins is generally carried out in *E. coli*, often exploiting the specialized C41 and C43 strains[118,132]. These strains were empirically selected for their ability to overproduce membrane proteins, and later work attributed this ability to decreased

expression of T7 RNA polymerase (RNAP) owing to mutations in the *lac*UV5 promoter by which T7 RNAP expression is controlled[133]. More recently, these specialized *E. coli* strains have been improved by the chromosomal deletion of the genes encoding AcrB and OmpF, which are the principal contaminants when membrane proteins are purified by immobilized metal affinity chromatography (IMAC)[134]. Both AcrB and OmpF crystallize at low concentration, leading to false positive crystallization experiments and many broken hearts. In recent years, *Lactococcus lactis* has emerged as a Gram-positive alternative to *E. coli* for membrane protein production, as it grows rapidly and to high density, does not require aeration (facilitating large scale fermentation), and the available promoter system is strong and tightly regulated[135]. In addition to expression systems, specific tags have been developed for the increased expression and integration of membrane proteins into the plasma membrane[136].

1.4.2 Detergent solubilization of membrane proteins

Solubilization of membrane proteins from the membrane into which they were expressed can be quite facile; however, doing so while maintaining the protein in a properly folded and active form can be a difficult task. The primary method of solubilizing proteins from membranes is by treatment with concentrated detergent solution[137]. Detergents are amphipathic molecules that contain a polar headgroup and a single hydrophobic alkyl chain tail and which, due to the disparity of cross section of the headgroup and the single acyl chain, spontaneously assemble into micelles with their hydrophobic tails buried in the core above a critical micelle concentration (CMC). The CMCs of different detergents vary immensely depending on their chemical structure, though the ionic strength, pH, and temperature of the solution in which they are dissolved can also have a significant effect.

Extraction of membrane proteins is generally carried out at a minimum of 10-20x CMC, while subsequent purification steps are regularly carried out at 2x CMC[118]. Detergents can be separated into three broad categories based on the structure of their headgroup as either ionic, zwitterionic, or non-ionic. Ionic detergents, such as sodium dodecyl sulfate (SDS), are generally quite effective at solubilizing membrane proteins, but are too harsh (often denaturing both lipid/protein and protein/protein interfaces) to maintain the sample in its properly folded state. For this reason, SDS and other ionic detergents, can be used as positive controls in screens for extraction detergents, but have limited further use in structural biology except for the most robustly folded proteins[118]. The zwitterionic detergents, including LDAO and CHAPS, are less denaturing than the fully ionic detergents and have successfully been used for both extraction and crystallization, but the non-ionic glycosidic detergents have by far been used most successfully[118]. Glycosidic detergents are characterized by a neutral carbohydrate polar head and an alkyl chain usually ranging from 7-14 carbon atoms. The most commonly used detergent for extraction is n-dodecyl-β-D-maltoside (DDM), and while it has also been used to great success in the crystallization stage, exchange into shorter chain glycosidic detergents such as n-decyl-β-D-maltoside (DM) or n-octyl-β-D-glucoside (β-OG) is often required for the formation of ordered crystals[118]. Development of new detergents is ongoing, and the introduction of the neopentyl glycol family of detergents, as well as a synthetic drop-in for digitonin (glyco-diosgenin, GDN) have been particularly exciting[138,139].

1.4.3 Fluorescence-detection size exclusion chromatography (FSEC)

Screening for appropriate expression and solubilization strategies can be extremely time consuming, even when only considering a single construct of a single protein. When homologues

from different species, truncations, and tag position are considered, the amount of work can increase exponentially. FSEC is a pre-crystallization screening strategy that allows for rapid assessment of all these factors by analyzing unpurified protein that has been expressed with an N- or C-terminally fused green fluorescent protein (GFP) tag[121]. By relying on the fluorescence signal from the GFP, nanogram amounts of protein can be used to measure relative expression or solubilization levels, determine optimal tag position, and identify oligomeric state in a given detergent. FSEC can also be applied to the optimization of membrane protein stability (FSEC-TS) by analyzing the thermostability of protein in the presence of different additives in the same manner that nano differential scanning fluorimetry (nanoDSF) and static light scattering (DSLS) are used[140]. FSEC and FSEC-TS have aided in the elucidation of numerous of X-ray crystal structures and is also being applied to the optimization of sample preparation for cryo-EM[141–148].

1.4.4 Lipid cubic phase (LCP) crystallization

The LCP or *in meso* method of crystallizing membrane proteins was initially described in 1996[130]. The authors wanted to devise a quasi-solid environment with viscoelastic properties similar to biological membranes, that could incorporate large amounts of protein without denaturation, and that could provide a structured, yet flexible matrix in which crystals could nucleate and grow. To that end, the liquid-crystalline LCP formed by two monoacylglycerols, monoolein and monopalmitolein, were tested for their ability to grow crystals of bacteriorhodopsin isolated from purple patches of *Halobacterium salinarum*[130]. Crystals grew in both mesophases, but only those from monoolein diffracted to high resolution; thus, monoolein has been the default choice for LCP crystallization ever since. Since that initial experiment, the field of LCP

crystallography has grown considerably and there are now over one hundred membrane protein structures deposited in the PDB that were solved using the *in meso* technique. While the monoolein mesophase may not be the perfect membrane mimetic, multiple enzymes have been shown to retain their activity when reconstituted in the cubic phase, suggesting that the conformational landscape that they are able to sample *in meso* is more biologically relevant than what might be seen in detergents[149,150].

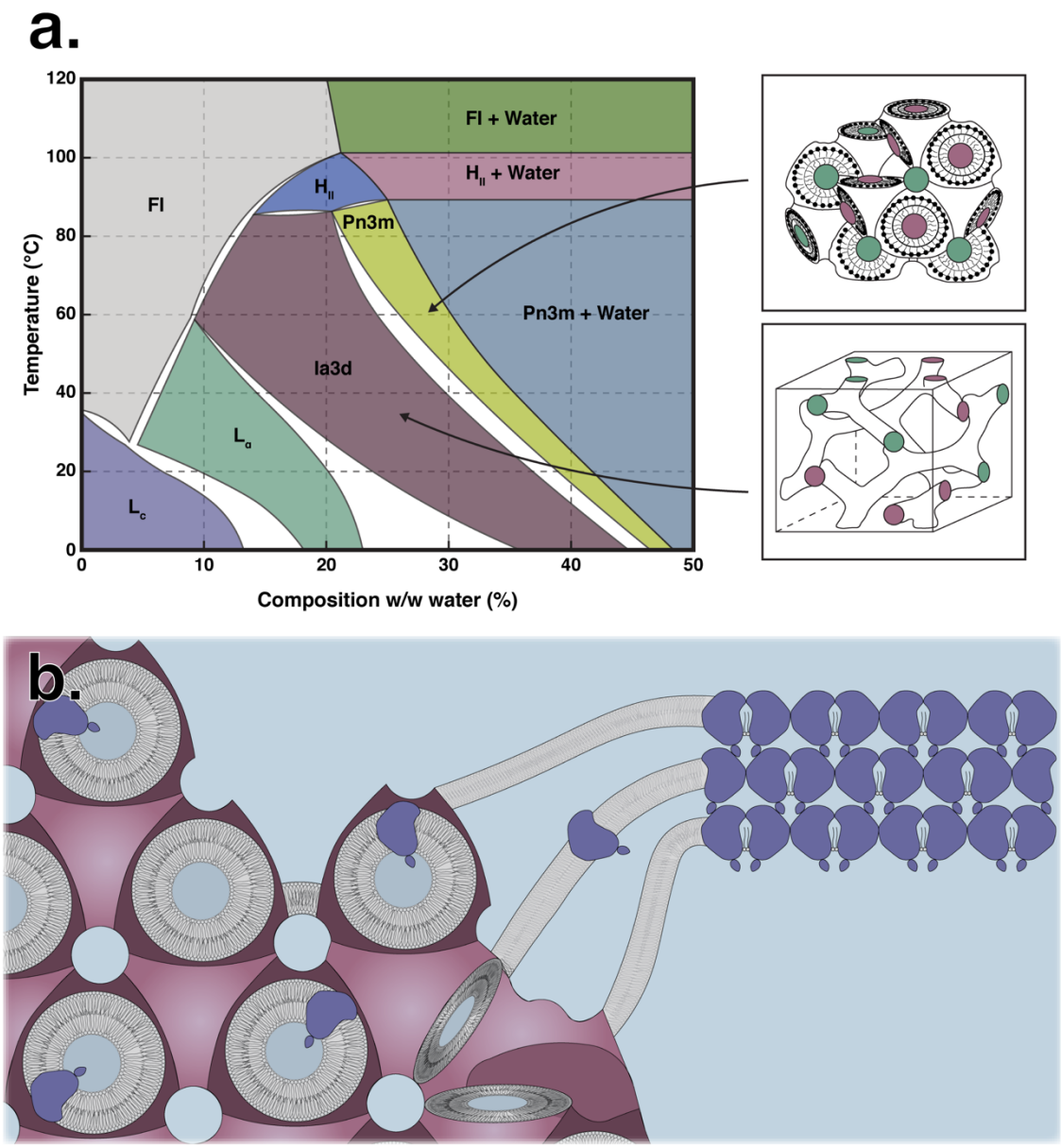


Figure 1.8 Mechanism of lipid cubic phase crystallization.
a. Temperature-composition phase diagram of the water-monoolein system. Insets show schematic representations of the Pn3m and Ia3d cubic phases. (FI, fluid isotropic phase. L_α, lamellar liquid crystalline phase. L_c, lamellar crystalline phase. H_{II}, hexagonal phase.) The generalized method of protein re-constitution into monoolein mesophase results in the formation of the Pn3m phase. Adapted from Caffrey, 2015[151]. **b.** Schematic representation of the protein crystallization in lipid cubic phase. Adapted from Caffrey, 2009[152].

A mechanism for LCP crystallization has been proposed on the basis of extensive small angle X-ray scattering (SAXS) and molecular dynamics (MD) experiments (Fig. 1.8)[153–155].

When homogenized at a ratio of 2:3 (protein solution:monoolein), the detergent solubilized protein is reconstituted into the bilayer of the highly curved, bicontinuous cubic phase with its polar regions protruding into solvent channels (Fig. 1.8B). Addition of a precipitant solution leads to an alteration of the LCP's lattice parameter and induces hydrophobic mismatch between the transmembrane domains of the protein and the cubic bilayer, leading to localized transitions from the cubic phase to the lamellar phase. It is in this lamellar phase that crystallogenesis occurs, with individual lamellae stacking one on top of the other, leading to Type I crystal formation (Fig. 1.8B). It is proposed that this lamellar phase also acts as a conduit between the growing crystals and the remaining bulk cubic phase through which protein can laterally diffuse and be incorporated into the growing face of the crystal[151]. It is also thought that in addition to precipitants, the detergent used during the membrane protein purification can promote the formation of local lamellar phase change, though the presence of too much detergent can prevent the formation of the initial cubic phase[151]. Also of note is that certain components in crystallization screens (e.g. polyethylene glycol) cause the cubic phase to "swell" and form what is called the sponge phase. The sponge phase retains the bicontinuous nature of the cubic phase, but loses its 3D periodicity and contains highly enlarged aqueous channels, and is thus suited for the LCP crystallization of membrane proteins that contain large extramembranal domains[151].

In recent years the ability to form "ultra-swollen" cubic mesophases without transition to the sponge phase has been demonstrated[156]. Other innovations in the field of LCP crystallography include the design and synthesis of novel monoacylglycerols as host lipids[157], pioneering of the *in meso in situ* crystallography (IMISX) technique[158,159], and the use of LCP injectors as a means to deliver microcrystals for serial femtosecond crystallography (SFX) using

X-ray free-electron laser (XFEL) sources[160] or serial microsecond crystallography (SMX) at synchrotron light sources[161].

1.5 Overview of thesis objectives

The research described in the following chapters was pursued with the goal of improving our fundamental understanding of bacterial cell wall biosynthesis through the structural and functional characterization of the polytopic integral membrane protein UppP, as well as with the goal of using structural biology to aid in the development of novel antibiotics that target cell wall carrier lipid metabolism. UppS and UppP play central roles in the metabolism of $C_{55}P$, a critical component for the synthesis of all bacterial cell wall polymers, and as such have been subjected to thorough biochemical characterization; however, a number of questions still remain. This thesis describes the years long effort to solve the X-ray crystal structure of *Escherichia coli* UppP, which we show unexpectedly adopts a unique inverted repeat topology that is commonly observed in transporters and ion channels. It has been suggested for some time that the $C_{55}PP$ phosphatases could play a role in facilitating the translocation of the recycled carrier lipid from the periplasmic leaflet of the inner membrane back to the cytoplasmic leaflet, and our structure provides the first empirical evidence that this might indeed be the case. This thesis also describes the characterization of small molecule inhibitors of UppS that were identified via a novel high-throughput screening platform. The inhibitors were found to bind in a similar location to previously characterized UppS inhibitors and show potential for use in combination therapies to combat drug resistant Gram-positive infections. The structural and functional studies contained herein contributes to our understanding of bacterial cell wall synthesis and provide a basis for structure guided design of UppP inhibitors.

2 Structural and functional characterization of an intramembranal phosphatase central to cell wall biosynthesis and carrier lipid recycling

2.1 Introduction

The translocation of sugars and glycan chains across membranes using long poly-prenyl phosphate lipids is a process that is highly conserved across all kingdoms of life. In bacteria, the most common carrier lipid is undecaprenyl phosphate ($C_{55}P$). As well as playing a role in protein glycosylation, $C_{55}P$ acts as the universal carrier lipid in the biosynthesis of peptidoglycan, wall teichoic acids, and many other major bacterial cell wall polymers[10]. During this process, nucleotide activated sugar moieties are transferred to $C_{55}P$ at the cytoplasmic face of the membrane bilayer. The resulting glycolipids are subsequently flipped to the periplasmic face of the plasma membrane by specialized glycolipid flippases such as MurJ and TagGH[47,162]. The lipid activated glycan moieties are then covalently transferred to specific glycan acceptors in the periplasmic space, in most cases resulting in the release of undecaprenyl pyrophosphate ($C_{55}PP$) as a byproduct[10,44]. In addition to being released as a byproduct of cell wall biosynthesis, $C_{55}PP$ is synthesized *de novo* by the cytosolic enzyme undecaprenyl pyrophosphate synthase (UppS). In both cases, $C_{55}PP$ must be dephosphorylated to $C_{55}P$ before it can be linked to a sugar or glycan. Disruption of the biosynthesis or recycling of $C_{55}PP$ halts peptidoglycan biosynthesis and subsequently results in cell lysis. While the synthesis of $C_{55}PP$ by UppS has been well characterized[60], its essential dephosphorylation to $C_{55}P$ remains poorly understood, as does the mechanism by which $C_{55}P$ would translocate from the periplasmic to cytoplasmic leaflet of the plasma membrane.

A 30 kDa polytopic integral membrane protein, undecaprenyl pyrophosphate phosphatase (UppP; also referred to in previous literature as BacA) was first identified in a screen for genes

that could confer resistance to the antibiotic bacitracin upon amplification[88]. While it was first proposed to function as an undecaprenol kinase, it has since been shown that UppP is rather a C_{55}PP phosphatase[52]. Knockout of *uppP* in *Escherichia coli* resulted in a 75 % decrease in C_{55}PP phosphatase activity, with little observed effect on apparent growth in liquid media[52]. A later study identified two phosphatidic acid phosphatase 2 (PAP2) family proteins, PgpB and YbjG, as the enzymes responsible for the residual C_{55}PP phosphatase activity with a *uppP*/*pgpB*/*ybjG* knockout lethal[53]. Despite the apparent redundancy in liquid media, effects of *uppP* knockouts *in vivo* are significant, with, for example, deficient *Staphylococcus aureus* and *Streptococcus pneumoniae* showing attenuated virulence in mouse models of infection[23] and *uppP* deficient *Mycobacterium tuberculosis* showing impaired biofilm formation[22]. These results, combined with the historical success of antibiotics targeting peptidoglycan biosynthesis[163], suggest that UppP could be a viable target for the development of therapeutics.

Bioinformatic and biochemical analyses of *E. coli* UppP (*Ec*UppP) have identified two conserved motifs containing residues that are implicated in the phosphatase activity of the protein; however, there is no consensus on the identity of *Ec*UppP's catalytic players, the mechanistic details of the dephosphorylation reaction, or on which face of the membrane dephosphorylation of C_{55}PP occurs. Site directed mutagenesis studies of *Ec*UppP carried out by Chang *et al.* and by Manat *et al.* proposed, alternatively, His30 or Ser27 as the central catalytic residue involved[106,164]. Both of these studies demonstrated that phosphatase activity was dependent on a divalent cation with a marked preference for magnesium and calcium[106,164]. Topology mapping of PgpB and YbjG, the primary sequences of which are completely distinct from UppP, suggested that the catalytic motif of the PAP2 phosphatases would be directed towards the periplasmic face of the bilayer, and confirmation of PgpB's topology was given by its recently

determined X-ray crystal structures from both *E. coli* and *Bacillus subtilis*[54,99,101]. The results of these studies suggested that the PAP2 phosphatases could act as recycling enzymes at the periplasmic face of the inner membrane, with UppP carrying out dephosphorylation during *de novo* synthesis of the carrier lipid in the cytoplasm. However, a recent topology analysis of *Ec*UppP has suggested both of the proposed catalytic residues are also directed towards the periplasm[106], again raising questions as to how cytoplasmic dephosphorylation during *de novo* synthesis might occur. Furthermore, while the specific glycolipid flippases involved in numerous biosynthetic pathways have been identified, and in some cases structurally characterized[165,166], there remains little information regarding the translocation of $C_{55}P$ to the cytoplasmic face of the inner membrane[10].

Here we present the X-ray crystallographic structure of *Ec*UppP at 2.0 Å resolution. Our high-resolution structure reveals a surprising membrane topology and overall architecture predominantly found in ion channels and transporters, gives insight into the active site and mechanistic details allowing for the observed intramembranal $C_{55}PP$ dephosphorylation and provides support for a potential role as a phosphatase activated $C_{55}P$ specific flippase.

2.2 Methods

2.2.1 Cloning, expression, and purification of UppP

The gene encoding full-length UppP from *Escherichia coli* K-12 (ATCC 10798) was cloned into a pET28a vector encoding an N-terminal hexahistidine tag and a thrombin cleavage site using restriction free cloning. UppP mutants were generated by quick change PCR.

UppP was overexpressed in *E. coli* C41 cells (Sigma), grown in ZYP-5052 autoinduction media supplemented with 100 μg/mL kanamycin at 37°C for 3 hours before lowering the

temperature to 27°C and allowing growth to continue overnight[167]. Cells were harvested by centrifugation and resuspended in lysis buffer containing 20 mM HEPES pH 7.5, 500 m NaCl, 10% glycerol, and 50 µL of Protease Inhibitor Cocktail Set III (Millipore). Resuspended cells were lysed 2x using an EmulsiFlex-C5 homogenizer (Avestin). Cellular debris was pelleted by centrifugation at 15,000 x *g* for 0.5 h and membranes were pelleted by centrifuging the resultant supernatant at 200,000 x *g* for 1.0 h. Membrane pellets were collected and homogenized in 20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol using a Dounce homogenizer. UppP was extracted from membranes by addition of *N*-dodecyl-β-D-maltopyranoside (DDM, Anatrace) to a final concentration of 1% (w/v) and allowing to stir for 1.0 h at 4°C. The extraction mixture was centrifuged at 200,000 x *g* for 0.5 h to remove insoluble material. The resulting supernatant was loaded onto a 5 mL Ni-NTA Superflow (Qiagen) column pre-equilibrated with 20 mM HEPES pH 7.5, 500 mM NaCl, 30 mM imidazole, 0.016% DDM, washed with 60 mM imidazole, and UppP was eluted with 250 mM imidazole. The purified protein was immediately desalted into 20 mM HEPES pH 8.0, 150 mM NaCl, 0.016% DDM, and the hexahistidine tag was removed by thrombin cleavage overnight. UppP was further purified by gel-filtration using a Superdex 200 Increase 10/300 GL column equilibrated with 20 mM HEPES pH 7.0, 150 mM NaCl, 0.016% DDM in order to remove aggregates and excess detergent before being concentrated for crystallization experiments in an Amicon Ultra-15 concentrator (Millipore) with a molecular weight cutoff of 50 kDa. Selenomethionine (SeMet) labelled UppP was overexpressed in *E. coli* C41 cells grown in PASM-5052 autoinduction labelling media[167]. The purification was carried out exactly as above with the exception that 2 mM β-mercaptoethanol was included throughout.

2.2.2 Crystallization of UppP

Vapour diffusion crystallization conditions for UppP were carried out in 96-well Intelli-Plates (Art Robbins Instruments) using a Crystal Phoenix (Art Robbins). Initial crystals were obtained in a condition containing 30% PEG 400, 300 mM $LiSO_4$, and 100 mM ADA pH 6.5 with UppP concentrated to 5 mg ml^{-1}. Using a grid-based optimization strategy based on the original condition, larger crystals were obtained in an optimized condition of 27% PEG 400, 400 mM $LiSO_4$, and 100 mM ADA pH 6.5. Efforts to further optimize the crystals included screening of the optimized condition combined with the Additive Screen (Hampton Research), pre-incubation of UppP with substrate or product, and limited proteolysis using trypsin.

Lipid cubic phase (LCP) crystallization trials were performed by mixing UppP concentrated to 12 mg ml^{-1} with monoolein (Sigma) in a 2:3 protein to lipid ratio (v/v). Crystallization trials were set up in 96-well glass Laminex sandwich plates (Molecular Dimensions) using an LCP Gryphon (Art Robbins Instruments). 100 nL LCP droplets were overlaid with 1 μL crystallization solutions and incubated at 20°C. Initial crystals were obtained using a solution of 40% PEG 200, 100 mM NaCl, 100 mM $LiSO_4$, and 100 NaCitrate pH 5. Optimized crystals grew in a solution containing 45-50% PEG 200, 0-150 mM NaCl, 200-400 mM $LiSO_4$, and 100 mM NaCitrate pH 4. In order to harvest the crystals, the cover glass of the Laminex plate around each well was manually scored with a tungsten carbide glass scribe (Molecular Dimensions) and gently removed from the well using tweezers. Crystals were looped from the mesophase using MicroMounts (MiTeGen) and were frozen directly in liquid nitrogen. SeMet substituted protein crystallized with the addition of 2 mM β-mercaptoethanol throughout purification. Mercury derivatized crystals were obtained using the same technique as above, but the concentrated UppP was incubated with a six-fold molar excess of ethyl mercury phosphate at

4°C for 30 minutes before preparation of LCP, and the final crystals were obtained in a solution containing 45% PEG 200, 150 mM $MgCl_2$, 400 mM $LiSO_4$, 100 mM NaCitrate pH 5.

2.2.3 Data collection and structure determination

X-ray diffraction data sets for both the native and mercury derivatized UppP crystals were obtained using Beamline 08ID-1 at the Canadian Light Source (CLS) in Saskatoon, Saskatchewan. Data were indexed and scaled using XDS[168]. Experimental phases were determined by autoSHARP[169] using SHELXD[170] for heavy atom substructure search, which identified a single mercury site at Cys165. Iterative cycles of model-building and refinement were performed with Coot[171] and Phenix[172]. Data collection and refinement statistics are given in Table 1. The final model has good stereochemistry (98.5% Ramachandran favored, 0.37% outliers), few clashes (clashscore of 3.69), and a MolProbity score of 1.16.

Conservation analysis was carried out using Consurf[173]. Cleft analysis was carried out using the 3V channel finder[174]. Figures were created with Chimera[175] and PyMOL.

2.2.4 Kinetic assays

UppP phosphatase activity was measured using the EnzChek Phosphate Assay Kit (Thermo Fisher). The enzymatic assay reaction mixture contained 50 mM HEPES pH 7.0, 150 mM NaCl, 1 mM $CaCl_2$, and 0.016% DDM. The enzymatic activity of wild-type UppP was measured using 0.9-24 μM farnesyl pyrophosphate (C15-PP) with 25 nM UppP. For comparison of wild-type UppP to S27A and H30A mutants, steady-state activity was measured in the presence of 200 μM C15-PP with 25 nM wild-type UppP, 1000 nM S27A UppP, and 100 nM H30A UppP. In the presence of the coupled assay enzyme purine nucleoside phosphorylase (PNP), the release of

phosphate results in the conversion of 2-amino-6-mercapto-7-methylpurine riboside (MESG) to ribose 1-phosphate and 2-amino-6-mercapto-7-methylpurine, production of which can be measured at 360 nM. All assays were performed in triplicate. Initial velocity data were fitted to the Michaelis-Menten equation using Origin data analysis software.

2.2.5 Chemical cross-linking

Purified UppP was incubated at room temperature for 30 minutes with or without a 50-fold molar excess disuccinimidyl suberate (DSS) or ethylene glycol bis(succinimidyl succinate) (EGS). Tris(hydroxymethyl)aminomethane (Tris) was added to a final concentration of 50 mM to quench the crosslinking reaction. After incubating for a further 30 minutes, 10 μg samples of each reaction were loaded on a 12% SDS-PAGE gel for analysis.

2.2.6 Fluorescence-detection size exclusion chromatography (FSEC)

Plasmid encoding each potential construct or orthologue of UppP was transformed by heat-shock into either BL21, C41, C43, or Rosetta *E. coli*. Single colonies were picked from transformation plates and used to inoculate 5 mL LB starter cultures. 50 mL expression cultures were grown in 250 mL Erlenmeyer flasks with 225 rpm shaking. Expression media, inoculation ODs, IPTG concentrations, and temperatures were varied. From each 50 mL culture, a 2 mL aliquot was taken, and cells were collected by centrifugation at 10,000 rpm. Cell pellets were resuspended in a 5:1 ratio of lysis buffer (as in 2.2.1), and cell lysis was achieved by sonication on ice using a microtip. Cell debris were collected by centrifugation at 5,000 rpm, the supernatant was transferred to a 1.5 mL tube for ultracentrifugation at 100,000 x *g* for 1.0 h in order to collect a membrane pellet. Membrane pellets were resuspended in lysis buffer supplemented with 1.0%

DDM and tubes were subjected to rotation for 1.0.h at 4°C. The solubilization mixture was transferred to a fresh 1.5 mL tube for ultracentrifugation at 100,000 x *g* for 0.5 hour, and the resultant supernatant was injected onto an Agilent HPLC coupled to a SEC column at a flowrate of 0.5 mL/min.

2.3 Results

2.3.1 Expression screening

A prerequisite for the structural characterization of UppP by X-ray crystallography was the development of an expression and purification strategy that would yield milligram amounts of highly homogenous purification. To that end, we employed FSEC to screen a number of bacterial UppP orthologues for relative expression levels[121]. The candidate orthologues were chosen on the basis of availability of genomic DNA, as well as clinical significance of their bacterium of origin. A notable exception was the inclusion of *Aquifex aeolicus*, a hyperthermophilic bacterium that can tolerate temperatures up to 95° C. Orthologues from hyperthermophilic bacteria are often good candidates for the crystallization of integral membrane proteins due to their intrinsic thermal stability, as evidenced by the crystal structure of *A. aeolicus* MraY, another important bacterial cell biosynthetic protein, that had been crystallized just prior to the outset of this project[176]. Initial screening indicated that the *E. coli* orthologue was the most promising candidate from those tested, as its fluorescence signal indicted >2.5x greater expression than the next candidate. Following selection of the *E. coli* orthologue, the following parameters were tested: tag position, induction temperature, expression strain, induction optical density (OD), expression media, and induction length. All of the C-terminally GFP tagged constructs performed far worse than their N-terminally tagged counterparts and inducing at 37°C resulted in less yield overall and more protein

in the void volume of the column than inducing at room temperature. Surprisingly, the FSEC data indicated that both BL21 and Rosetta strains of *E. coli* performed slightly better than C41 and much better than C43, despite those strains being specifically selected for their ability to overexpress membrane proteins; however, the C41 strain was chosen due to the higher cell densities that were achieved in large scale trials. As previously reported for the Walker strains, induction of protein expression at or below OD_{6oo} of 0.5 resulted in better yields in C41 than induction from higher cell densities. Length of induction, perhaps unsurprisingly, correlated well with protein yield and, with no visible changes in monodispersity or degradation of the sample between the shorter and longer time points, we decided to use an overnight induction. One the largest differences observed during screening was between different expression media, with expression in Terrific Broth yielding >8x more protein than expression in LB.

FSEC mediated screening indicated that expressing the N-terminally His-tagged *E. coli* orthologue of UppP overnight at room temperature using C41 cells in Terrific Broth with an induction of OD_{600} of 0.5 was the best strategy moving forward (Fig. 2.1).

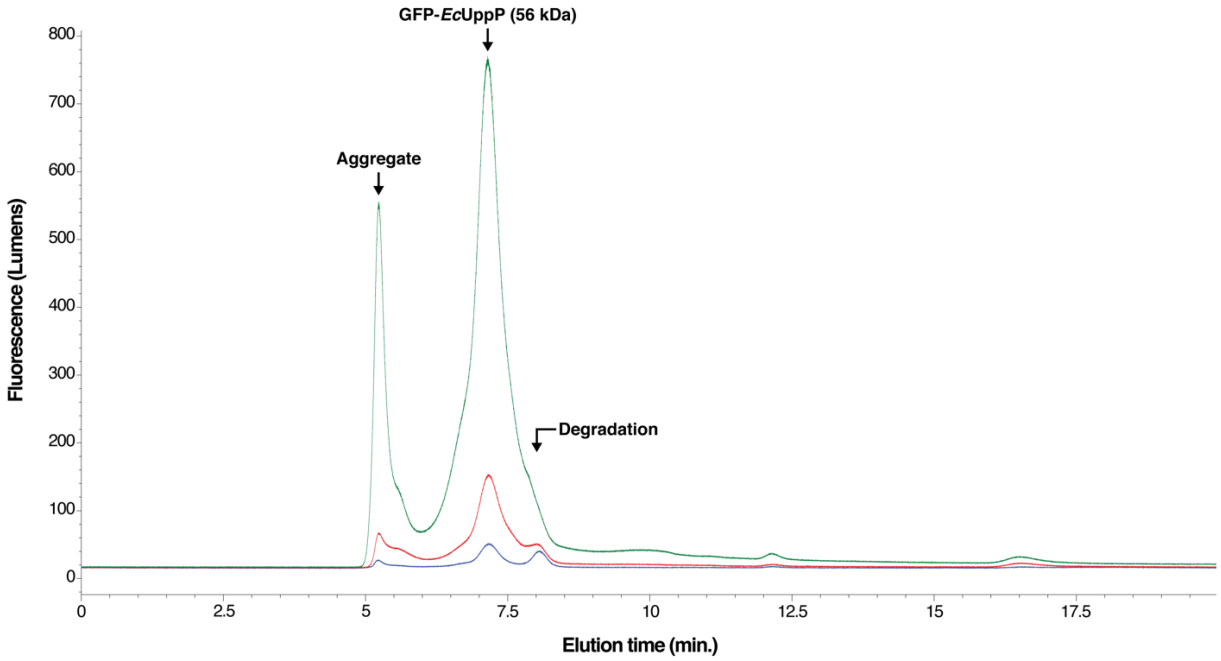


Figure 2.1 FSEC optimization of UppP expression.
Representative FSEC traces showing the yield from the initial expression strategy for *Ec*UppP (blue), after a single round of optimization (red), and after two rounds of optimization (green).

2.3.2 Optimization of purification

Detergents are used for the solubilization of membrane proteins from the membranes into which they are expressed; however, once a protein is extracted from the membrane, further delipidation may be undesirable as many membrane proteins require specific lipids for structural integrity or proper function. It is for this reason that membrane protein purifications should be carried out in as few steps as possible, and it is why a great deal of time was spent optimizing the IMAC purification of *Ec*UppP. The various parameters that were screened in optimization are summarized in Table 2.2. A Ni-NTA based resin was eventually used for IMAC after it was found that Ni-IDA resins did not give sufficient purity, and Co^{2+} based resins seemed to promote the formation of a spurious tetrameric oligomer on gel-filtration. Other than choosing the correct resin, incorporating a much harder post extraction spin and switching from a decahistidine tag to

hexahistidine tag were the factors that made the greatest difference in the purity of the sample prior to gel-filtration (Fig. 2.2a).

A gel-filtration step was incorporated into the purification scheme in order to separate monodisperse *Ec*UppP from aggregates, as well as empty DDM micelles (Fig. 2.2c). Due to *Ec*UppP's small size, it was not possible to concentrate the detergent-protein complex using a 100 kDa MWCO concentrator that would have allowed for empty micelles to flow through. The gel-filtration step allowed us to collect only the most concentrated fractions of monodisperse protein, minimizing the degree to which we would need to concentrate the final sample for crystallization trials.

Table 2.1 Strategies for the optimization of UppP IMAC purification.

Modification	Effect	Modification	Effect
Extraction time		Inclusion of membrane washes	
Addition of glycerol		C-terminal His-tag	- -
Addition of reducing agent	-	Hexa- vs. decahistidine tag	+
Harder low speed spin	+	Addition of MBP tag	+
Harder high speed spin		Ni^{2+} vs. Co^{2+} resin	+ +
Harder post-extraction spin	+ +	Ni-IDA vs. Ni-NTA	+

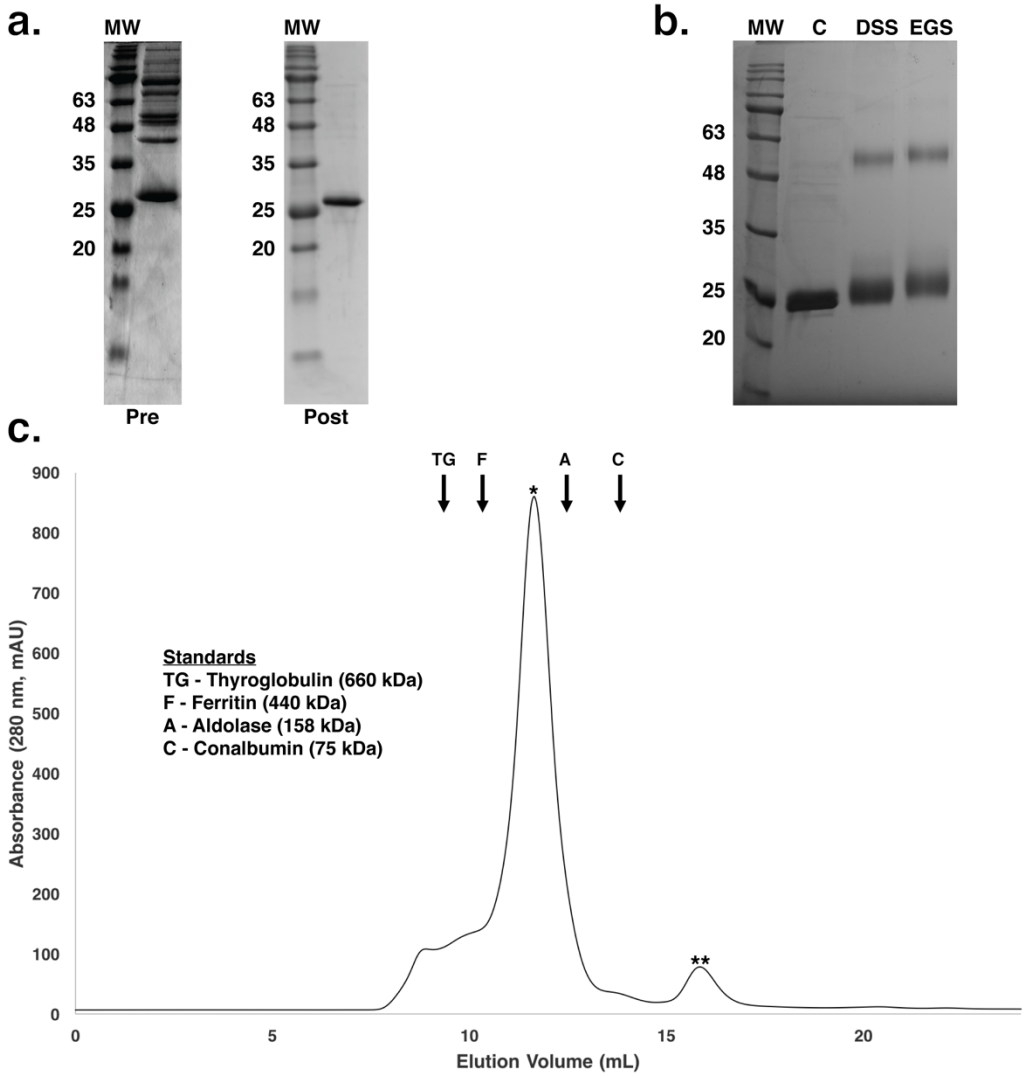


Figure 2.2 Optimization of IMAC purification and data supporting the dimerization of UppP.

a. SDS-PAGE analysis of IMAC purified UppP pre- and post-optimization of purification strategy **b.** SDS-PAGE analysis of UppP (C) incubated with disuccinimidyl suberate (DSS) or ethylene glycol bis(succinimidyl succinate) (EGS). Addition of either DSS or EGS results in the appearance of a single higher MW band at approximately 60 kDa, consistent with crosslinked UppP dimer **c.** Gel filtration chromatogram of Ni-NTA purified UppP using a Superdex 200 Increase 10/300 GL column. The elution volume of UppP (*) is consistent with the formation of a dimer (60 kDa) + two DDM micelles at room temperature (175 kDa). The small peak (**) eluting after UppP corresponds to the thrombin used in removal of the hexahistidine affinity tag.

2.3.3 Crystallization and structure solution of UppP

Extensive crystallization trials of detergent solubilized *Ec*UppP were carried out at a wide range of protein concentrations using the vapour diffusion method, both with the hexahistidine tag left intact or removed by thrombin cleavage, but the protein remained recalcitrant to crystallization. In a bid to express *Ec*UppP in a more physiological lipid environment, and with the aim of inducing expression in a gentler manner, the expression protocol was altered to use ZYP-5052 autoinduction media at 27°C[167]. The first purification using the altered expression protocol yielded our first *Ec*UppP crystals in a 0.4 ul drop containing 30% PEG 400, 300 mM $LiSO_4$, and 100 mM ADA pH 6.5 with UppP concentrated to 5 mg ml^{-1} **(**Fig. 2.3a,b**)**. The initial crystals were replicated in a larger drop (2 ul) and were subsequently optimized using a grid screening approach (Fig. 2.3c,d**)**. Optimized crystals were looped, directly frozen in liquid N_2 without the addition of a cryoprotectant, and sent for X-ray diffraction analysis at the CLS. The crystals yielded a highly anisotropic X-ray diffraction pattern with diffraction extending to ~6 Å in the best direction **(**Fig. 2.4**)**; thus, further optimization of the vapour diffusion crystals was pursued by screening of additives such as secondary detergents, alcohols, or soluble substrate mimetics **(**Fig. 2.3e,f**)**. While these additives did effect noticeable changes in crystal morphology, none of them improved the anisotropic diffraction that was observed as with the initial images, nor did the incorporation of various cryoprotectants prior to freezing.

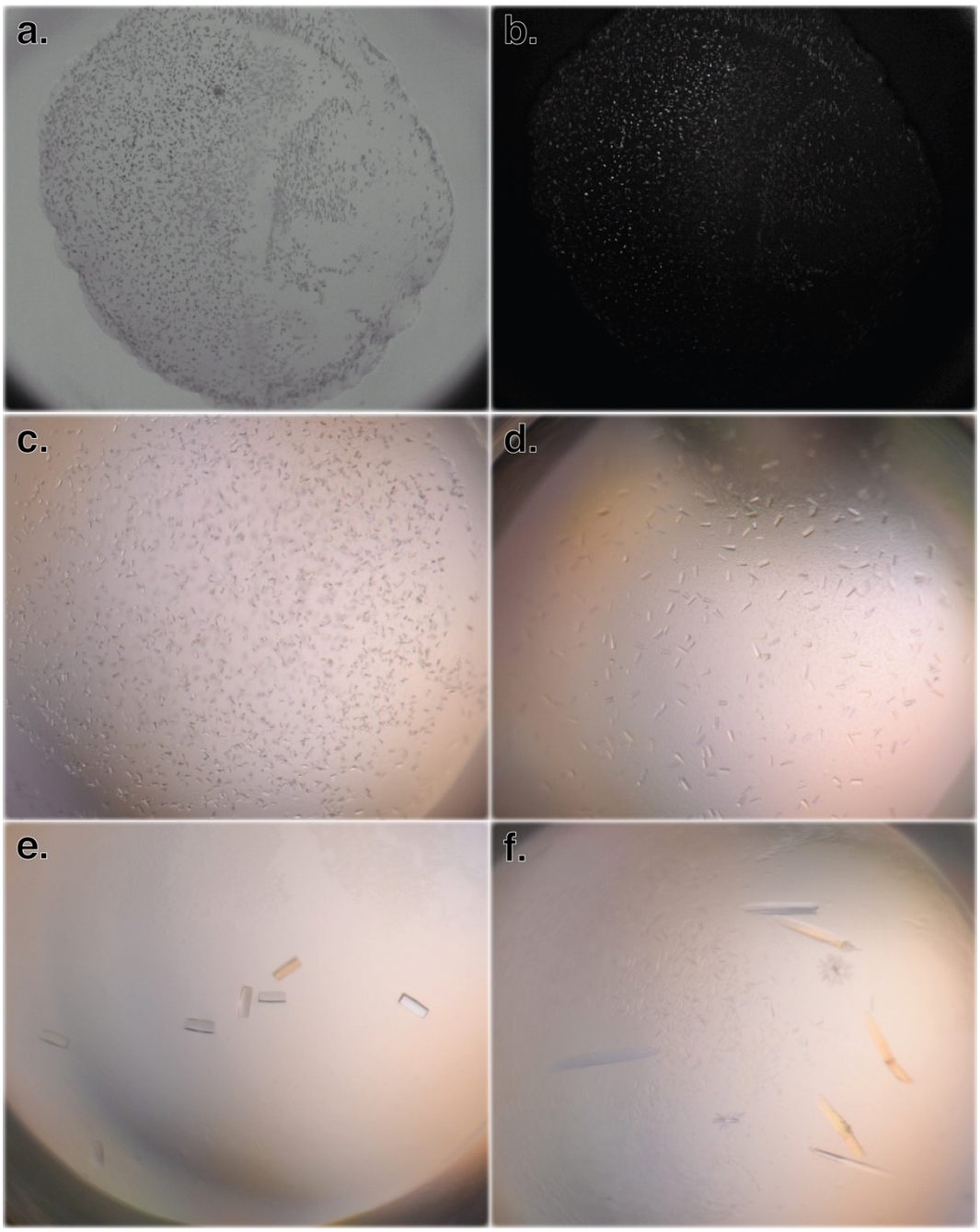


Figure 2.3 Vapour diffusion crystallization of UppP.
a. Crystallization drop containing initial crystals taken 48 hours after set up. **b.** The same image as panel **a** taken with a UV filter. **c.** Initial crystals reproduced in a larger drop **d.** Optimized crystals **e.** Optimized crystals with DM added as a secondary detergent **f.** Optimized crystals set up in complex with C_{15}PP substrate analog.

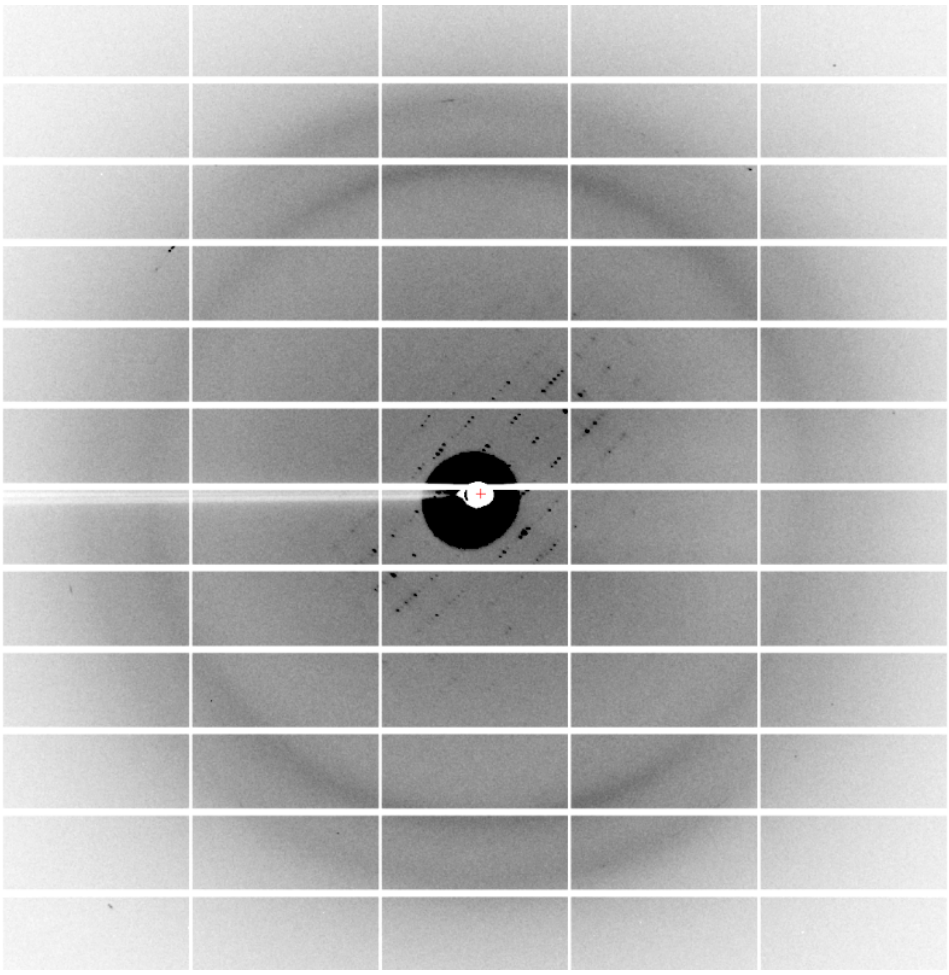


Figure 2.4 Anisotropic diffraction of UppP vapour diffusion crystals.
Collected on Beamline 08ID-1 at the CLS, 0.2° oscillation, 2 s exposure, Pilatus3 S 6M detector.

With the vapour diffusion method failing to yield crystals of appropriate quality for structure determination, we turned to the LCP method[130]. Initial crystallization trials yielded hits in several related conditions, and an initial round of grid optimization around the original hits resulted in crystals that diffracted to 2.9 Å **(**Fig. 2.5a,b**)**. A further round of optimization with fine concentration steps of PEG and salt around the partially optimized condition led to a crystallization condition (100 mM NaCitrate pH 4, 150 mM NaCl, 300 mM LiSO4, 50% PEG 400) that favoured the formation of the sponge phase and enabled the growth of large, rod shaped crystals that

diffracted to 2.0 Å on Beamline 08ID-1 at the CLS **(**Fig. 2.5c, 2.6**)**. The crystals belonged to space group *C*222 with unit cell dimensions 109.88 x 146.08 x 40.19 Å and one molecule in the asymmetric unit (53.2% solvent content as calculated using Xtriage in Phenix[172]).

Experimental phasing with crystals grown *in meso* is particularly difficult because the crystals are generally quite small, radiation-sensitive, and are embedded in excess lipid that results in background scattering that can make it difficult to collect complete data; thus, the vast majority of crystal structures solved using the LCP method have relied on molecular replacement (MR) for phase determination[177]. A search of the PDB for an appropriate model for MR yielded no results and attempts to phase the data using a Rosetta predicted model failed, so experimental phasing was pursued[178]. We first attempted to generate SeMet labelled *Ec*UppP crystals; however, while we were able to obtain crystals, they were much thinner than those obtained with unlabelled protein and their diffraction was of insufficient quality to permit structure solution. We next tried to phase the data using tungstate (WO_4^{-2}), reasoning that it could mimic phosphate and bind in the active site of *Ec*UppP as in the LCP crystal structure of *Bs*PgpB[101]. The crystals obtained with WO_4^{-2} diffracted well, but we detected no anomalous signal, suggesting that the WO_4^{-2} did not bind to the active site as we hoped. Iodide SAD phasing has been suggested as a potential high-throughput phasing method for membrane proteins, facilitated by the large number of uncompensated positive charges often found on the cytoplasmic face of membrane proteins[179]. We explored the use of iodide soaks and co-crystallizations but were again unable to obtain a strong enough anomalous signal for phase determination. Finally, we attempted to obtain mercury derivatized crystals by preincubating concentrated *Ec*UppP with six-fold molar excess ethyl mercury phosphate prior to crystallization trials. We obtained much shorter rod-shaped crystals that diffracted to 3.0 Å and

had sufficient anomalous signal for structure solution **(**Fig. 2.5d**)**. SAD data was collected at the anomalous peak wavelength of Hg (1.0057 Å) on beamline 08ID-1 at the CLS.

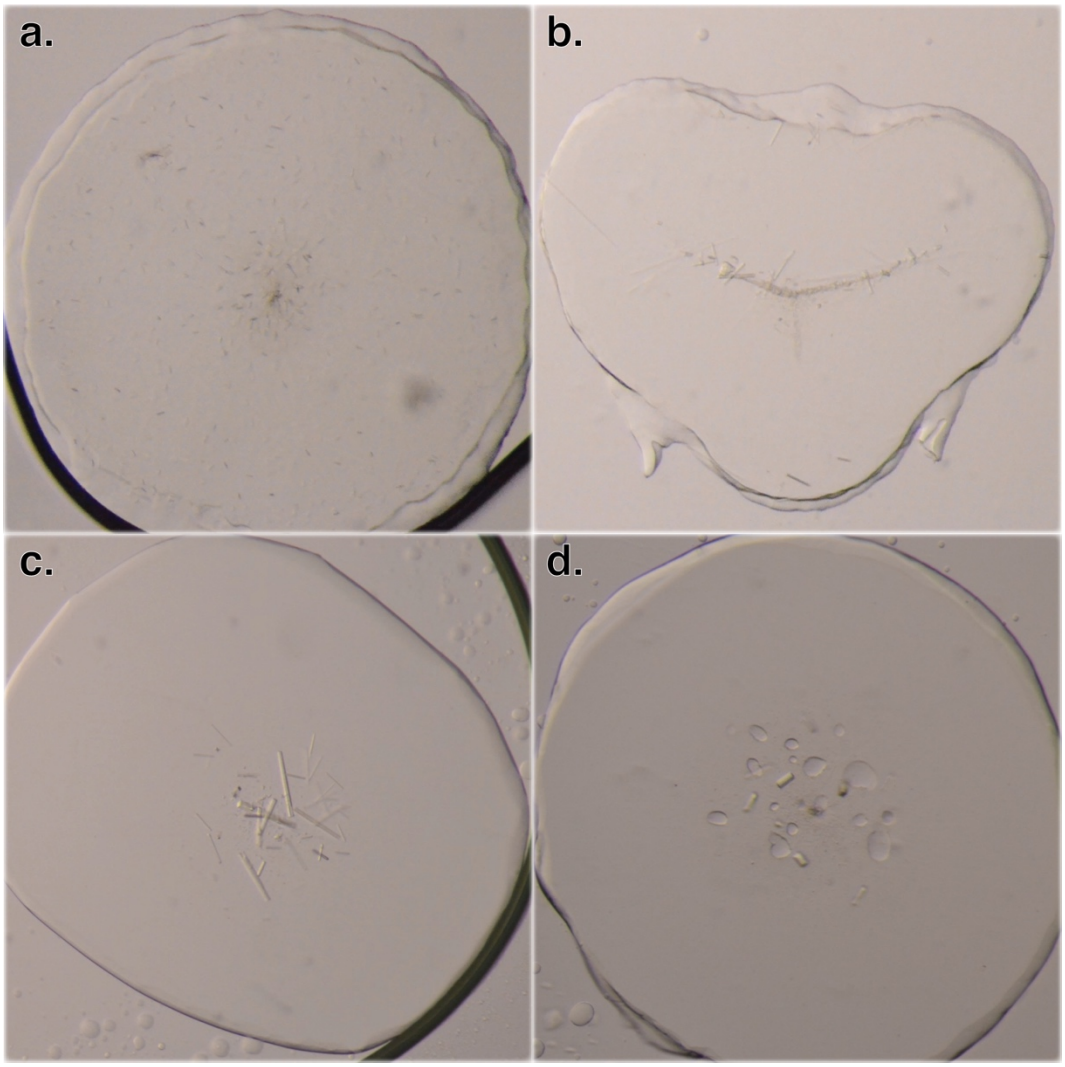


Figure 2.5 LCP crystallization of UppP.
a. The most promising initial crystal hit obtained in MemMeso condition A2 (0.1 M NaCitrate pH 5, 0.1 M NaCl, 0.1 M LiSO4, 40% PEG 200) **b.** Crystals after a single round of optimization that diffracted to 2.9 Å **c.** Fully optimized crystals in the sponge phase that diffracted to 2.0 Å **d.** Heavy atom derivatized crystals grown by pre-incubating UppP with ethyl mercury phosphate.

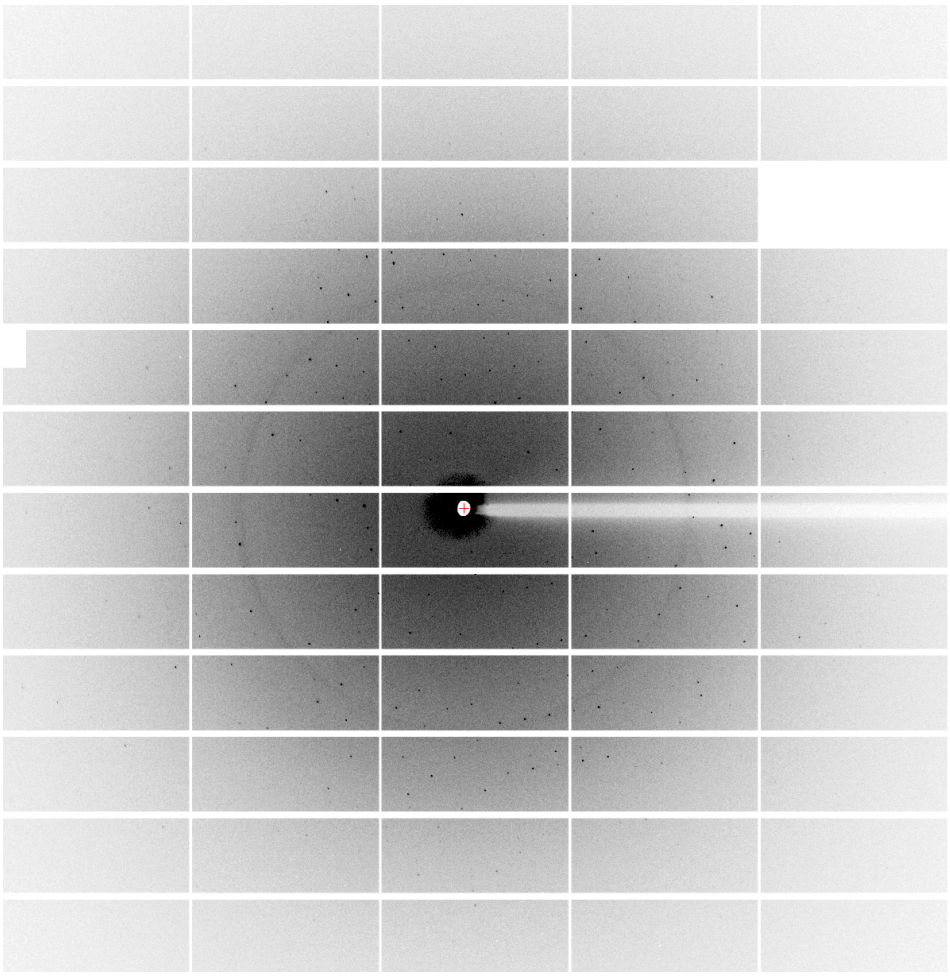


Figure 2.6 Representative diffraction image of optimized UppP LCP crystals.
Collected on Beamline 08ID-1 at the CLS, 0.2° oscillation, 1 s exposure, Pilatus3 S 6M detector.

The initial electron density map that was obtained using experimental SAD phases determined by autoSHARP contained largely unconnected peaks of density into which only simple poly-Ala α-helices could be manually placed in Coot **(**Fig. 2.7a**)**[169,171]. A subsequent round of density modification by SOLOMON in autoSHARP using the α-helices as boundaries for solvent flattening (input solvent of 47.3%) resulted in a more interpretable electron density map that allowed us to build an initial model using Buccaneer **(**Fig. 2.7b**)**[169,180,181]. We used our initial all atom model to phase the high-resolution dataset by molecular replacement using Phaser, and our final model of *Ec*UppP was obtained after iterative cycles of model-building and refinement

using Coot and Phenix **(**Fig. 2.7c,d**)**[171,172,182]. All residues could be readily traced in the resultant electron density maps with the exception of Met1-Asp3. Several monoolein lipid molecules of varying order and occupancy were also modeled and included in the refinement. The final refinement statistics and stereochemical indicators for the resulting model are described in the methods and shown in Table 2.2**.**

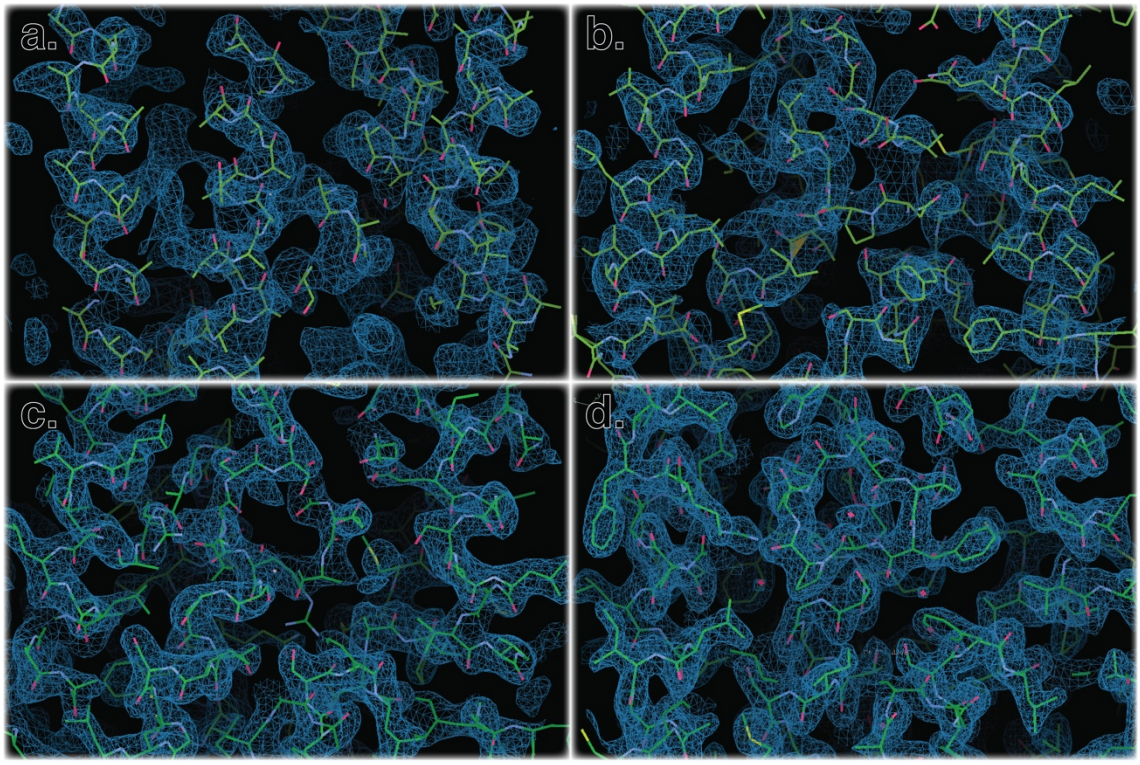


Figure 2.7 Iterative electron density map improvement
a. Manual placement of poly-Ala α-helices following experimental SAD phase determination in autoSHARP. Experimentally phased map from SOLOMON, contoured at 1.5 σ. **b.** Starting model from Buccaneer built into map resulting from subsequent solvent flattening density modification. Density modified experimental phasing map from SOLOMON, contoured at 1.5 σ. **c.** Molecular replacement solution from Phaser using the Buccaneer built model with the high-resolution native data set. $2mF_o$-F_c map contoured at 1.5 σ. **d.** Final electron density map with fully refined UppP model at 2.0 Å resolution. $2mF_o$-F_c map contoured at 1.5 σ.

Table 2.2 UppP data collection and refinement statistics.

	EtHgPO4 UppP	Native UppP
Data collection		
Wavelength (Å)	1.0057	0.9795
Space group	*C*222	*C*222
Cell dimensions		
a, b, c (Å)	111.19, 146.84, 40.20	110.09, 146.20, 40.23
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	40.20-3.00 (3.11 - 3.00)	40.23 - 2.00 (2.07 - 2.00)
R_{pim}	0.0502 (0.2835)	0.0485 (0.7422)
I / σI	14.38 (2.91)	8.12 (1.26)
$CC_{1/2}$	0.998 (0.951)	0.998 (0.463)
Completeness (%)	99.70 (99.28)	94.59 (70.97)
Redundancy	12.1 (11.9)	5.8 (4.0)
Refinement Statistics		
Resolution (Å)		40.23 - 2.00 (2.07 - 2.00)
No. reflections		21,278
R_{work} / R_{free}		0.21 / 0.24
No. atoms		
Protein		2080
Ligand		145
Water		44
B factors (Å^2)		
Protein		50.88
Ligand		75.16
Water		48.51
r.m.s deviations		
Bond lengths (Å)		0.002
Bond angles (°)		0.46

*Numbers in parenthesis refer to the highest resolution shell.
†$R_{pim} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$

2.3.4 X-ray crystallographic structure of *Escherichia coli* UppP

Our refined structure reveals *Ec*UppP is composed largely of ten membrane embedded α-helices (Fig. 2.8), contrary to previous sequence and biochemical based topology predictions that it would contain between seven or eight[106,164,183]. Six helices are full-span transmembrane (TM) helices (α3-α5, α8-α10), while the remaining four make up two unexpected antiparallel reentrant helix-loop-helix regions (α1-α2, α6-α7). The latter are arranged in an inverted manner at what we propose is the enzyme active site and contain highly conserved residues that have been

previously shown by mutagenesis to be crucial for the phosphatase activity of the protein[106,164]. We observe a classic girdle of aromatic residues at the two leaflet interfaces, defining the membrane span (Fig. 2.8a), and analysis of electrostatic surface properties support orientation with respect to the lipid bilayer as depicted in Fig. 2.8c. A significant charge polarity is observed with the hydrophilic surface defined by loops connecting helices α3a-α4 (encompassing helix α3b), α5-α6, α7-α8 and α9-α10 largely electropositive and predicted to be oriented to the cytoplasm with the extended loops α3a-α4 (Gly73-Thr91) and α5-α6 (Lys140-Thr153) harboring a large number of positively charged residues (Fig. 2.8a,c) in keeping with the positive inside rule of membrane protein topology[184]. Conversely, the surface containing the N- and C- termini along with loops connecting helices α2-α3, α4a-α5 (encompassing helix α4b) and α8-α9 project to the periplasm and form a predominantly electronegative surface (Fig. 2.8c).

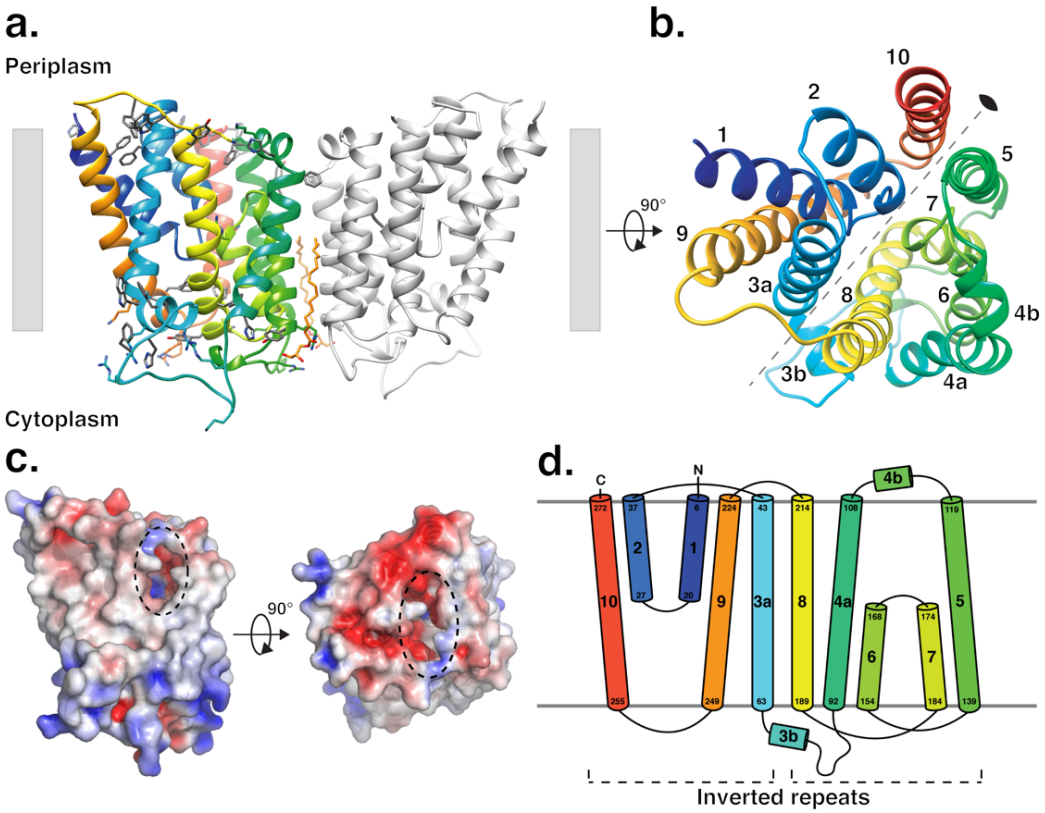


Figure 2.8 The crystal structure of EcUppP at 2.0 Å resolution.
a. Cartoon representation of EcUppP dimer with one monomer-colored rainbow from the N- (blue) to C-terminal (red). Aromatic and positively charged residues at the membrane interface shown as stick and highlight orientation with respect to the inner membrane (shown as gray bars). Two monoolein lipids at the dimer interface shown as stick. **b.** Ninety-degree rotation from **a** viewed from periplasm. Colored as **a** and helices numbered. Twofold pseudosymmetry axis parallel to the plane of the membrane shown as gray dotted line. **c.** Electrostatic surface potential of UppP monomer. Orientations as in **a** and **b**, respectively. Dotted circles indicate location of periplasmic substrate-binding cleft. **d.** Secondary structure topology highlighting interlocked inverted repeat.

Although *Ec*UppP crystallized with one molecule in the asymmetric unit, analysis of the crystal packing shows a crystallographic dimer with a two-fold symmetry axis perpendicular to the plane of the membrane (Fig. 2.8a**)**. Analysis with PISA shows an interface surface area of 973.5 Å^2, which along with the parallel arrangement with respect to the membrane bilayer suggests it may be of physiological relevance[185]. Gel-filtration analysis demonstrates that detergent solubilized *Ec*UppP elutes with a higher hydrodynamic radius and apparent molecular weight than theoretically expected for a monomer plus detergent micelle (Fig. 2.2c). To further study this we

conducted chemical cross-linking using amine reactive cross-linkers disuccinimidyl suberate and ethylene glycol bis(succinimidyl succinate) and observed the accumulation of a dimer species (Fig 2.2b). Two crystallographically related lipids, modelled here as monoolein and well-ordered in their entirety, are observed flanking the interface from the cytoplasmic surface (Fig. 2.8a) which could play a role in stabilization of the dimeric form, a reoccurring general feature recently shown to be of importance in membrane protein oligomeric interfaces[186]. Beyond this likely stabilization and the observed parallel disposition of the two monomers relative to the membrane leaflets, it is not evident from our structural data if UppP dimerization is also necessary for an additional allosteric (distinct active sites ~ 35 Å distance apart) or other functional role.

Intriguingly, analysis of the helical packing arrangement reveals *Ec*UppP has internal pseudosymmetry within each monomer, with a two-fold rotation axis parallel to and bisecting the membrane midplane which relates a five-helix motif encompassing respective reentrant regions (Fig. 2.8b,d). The contiguous second repeating domain (helices α4-α8) is inserted in sequence between helices α3 and α9 of the first repeat, consequently resulting in an interlocked inverted domain-swap like arrangement where the C-terminal helices of the first repeat are separated in primary structure but tightly associated in tertiary structure. Superposition of the domains reveals a strong degree of structural similarity with a backbone RMSD of ~ 3.2 Å over 218 Cα (Fig 2.9) and a calculated sequence similarity for the structurally aligned regions of 29 % (identity 19 %) suggesting a common evolutionary origin. Inverted two-fold pseudosymmetry is common amongst α-helical transporters, receptors and channels[187]. Intriguingly, this interlocked inverted topology repeat has primarily been described in proteins involved in secondary transport across lipid bilayers and the functional implications of this are discussed below[187–189].

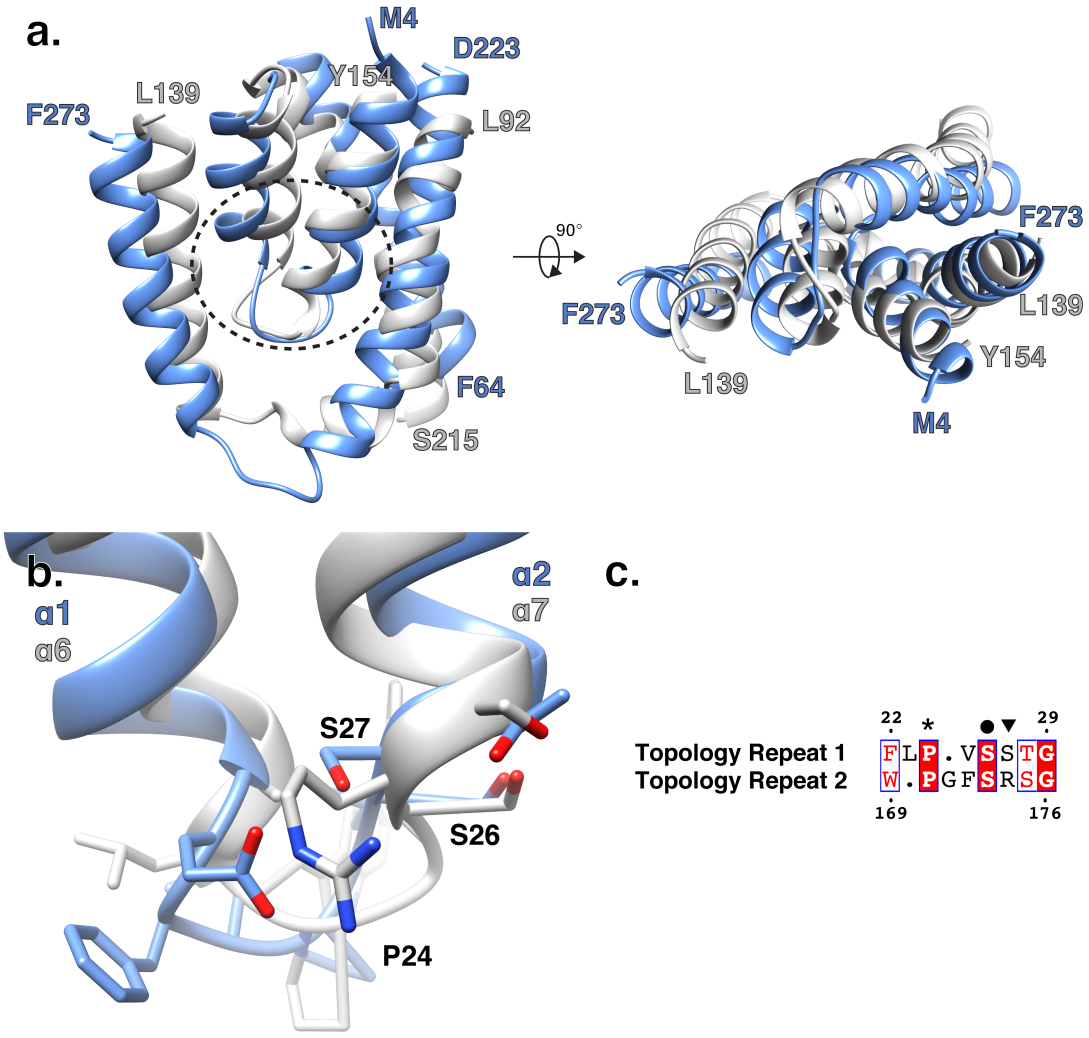


Figure 2.9 Structural superposition of the inverted topology repeats.
a. The first (blue) and second (grey) inverted topology repeats are shown structurally aligned from a side on and top down view. Panel **b** shows region in dashed circle in more detail. **b**. Superposition of just the reentrant helix-loop-helix motifs illustrating conservation of the key structural and catalytic residues. **c.** Sequence alignment of the reentrant loops indicating the crucial helix-breaking prolines (*), helix N-cap serines (●), and the direct alignment of the catalytic serine with the substrate binding arginine (▼).

2.3.5 The substrate binding pocket and active site

A prominent feature of the *Ec*UppP structure is the large cleft open to both the aqueous periplasm and the hydrophobic acyl core of the bilayer. Analysis with the 3V server calculates a cleft volume of 1914 Å^3 with approximate dimensions of 18 Å (l) x 8 Å (w) x 18 Å (h) (Fig. 2.10a,

Fig. 2.11)[174]. Electrostatic surface analysis shows a negatively charged funnel at the periplasmic face feeding into a deep hydrophobic channel which in turn empties into an electropositive basin at the pocket formed by the inverted reentrant helices (Fig. 2.8c, 2.10a).

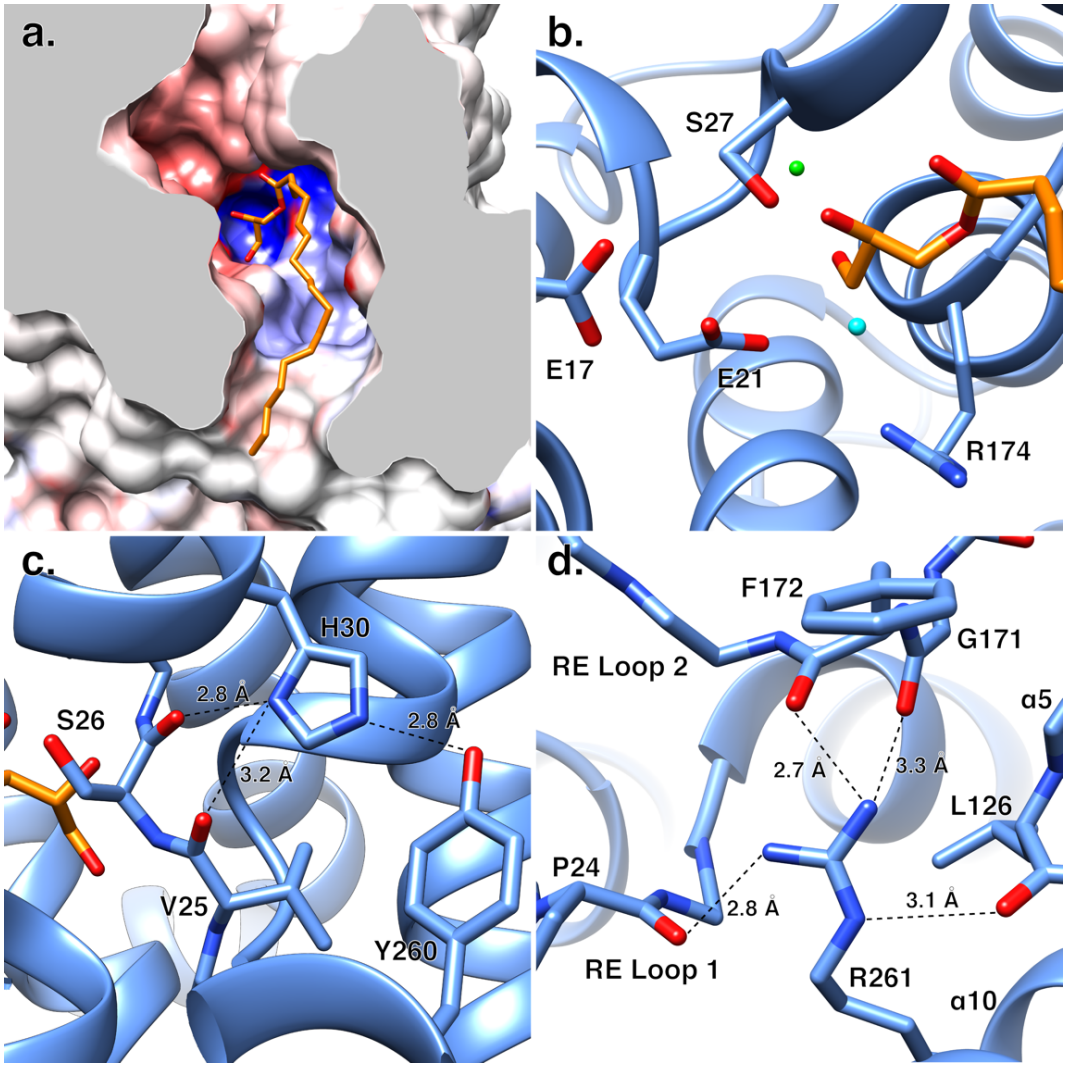


Figure 2.10 The EcUppP substrate-binding pocket.
a. Clipped view of the EcUppP substrate-binding pocket from a periplasmic viewpoint showing the electronegative funnel and deep hydrophobic channel leading into the positively charge basin formed by the antiparallel inverted reentrant helices defining the active site. An observed monoolein lipid is shown in stick with the polar headgroup bound in the active site pocket and lipid tail exiting the cleft along the hydrophobic channel. **b.** Magnified view of the EcUppP active site with key catalytic residues shown in stick. Two modeled active site waters are shown as cyan and green spheres representing the proposed catalytic water and a putative cation-binding site, respectively. **c.** His30 forms structural hydrogen bonds with the backbones amides of Val25 and Ser26, as well as the hydroxyl oxygen of Tyr260. **d.** Arg261 is buried at the membrane midplane

(also see Supplementary Fig. 6) and forms a hydrogen-bonding network that links both reentrant loops (RE Loop 1/2) through the backbone amides of Pro24, Gly171, and F172, in addition to α5 through the backbone amide of Leu126.

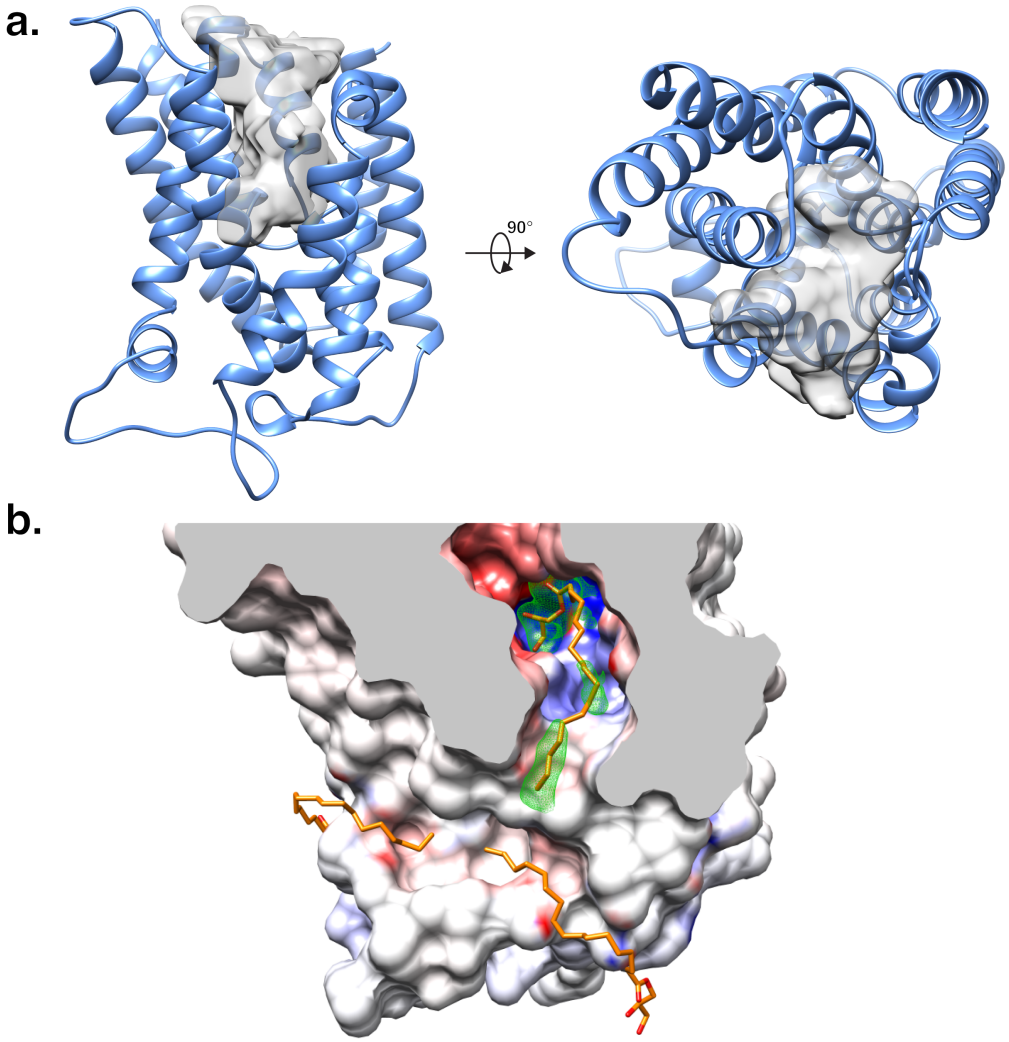


Figure 2.11 The periplasmic oriented substrate binding cleft.
a. The binding cleft volume generated by the 3V server is visualized as a transparent surface modeled on top of a cartoon representation of *Ec*UppP from a side on and top down view showing a deep cleft extending to the membrane midplane (volume 1914 Å^3). **b.** mF$_o$-DF$_c$ simulated annealing omit map for the substrate cleft bound monoolein lipid, contoured at 2σ. The polar and hydrophobic ends are clearly defined while the connecting acyl chain is less well ordered. The acyl chains of two additional monoolein molecules are observed in a hydrophobic pocket below the entrance to the active site cleft, revealing a potential binding site for the hydrophobic tail of C_{55}PP.

The V-shaped opening to the lipid environment of the periplasmic leaflet is framed by helices α4a and α8, which are both kinked at invariant prolines (Pro101 and Pro202) positioned at the midplane of the membrane (Fig 2.11, Fig. 2.12a). Many of the residues lining the cleft are well conserved (Fig. 2.12); previous bioinformatic and biochemical analysis identified two highly conserved sequence motifs important for function which we show map to the loop regions connecting the respective reentrant helices (residues 17-30 and 160-179) situated at the membrane midplane and which structurally define the active site.

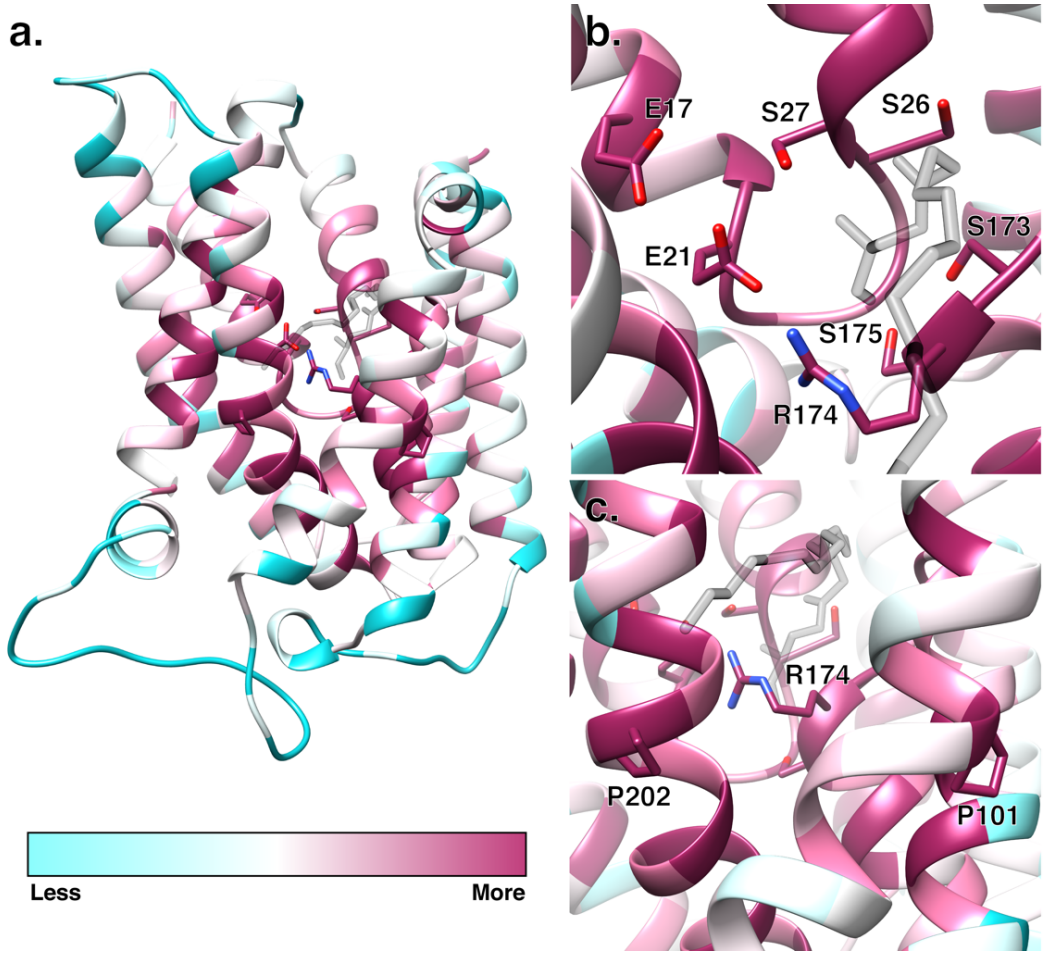


Figure 2.12 Conservation of EcUppP catalytic core.
a. EcUppP structure colored according to sequence conservation from low (cyan) to high (maroon). Highly conserved residues cluster near the catalytic core of the protein. Bound monoolein captured in our structures is displayed in transparent gray. **b.** Top-down view of the active site, showing the highly conserved nature of residues implicated in catalysis (E17, E21, S27,

and R174) and maintenance of the reentrant loop architecture (S26, S173, and S175). **c.** Magnification of strictly conserved proline residues that facilitate the bending of α4a and α8 to create the cleft that gives access to the active site.

The location is in notable contrast to the surface localized active sites previously characterized in the structurally and functionally distinct PAP2 and other integral membrane spanning phosphatases[99,101]. Electron density for a single monoolein lipid is observed in the cleft with the well-ordered glycerol headgroup bound in the deep electropositive basin near a number of highly conserved residues we believe to be central to phosphatase action. Its more dynamic lipid tail is oriented along the hydrophobic channel and exits at the base of the opening to the membrane hydrophobic core (Fig. 2.10a). Like C_{55}PP, monoolein is a single chain lipid with a hydrophilic headgroup and we propose the observed binding orientation is reflective of the native substrate complex.

The reentrant loops are similar in both sequence and structure (Fig. 2.9b,c). Strictly conserved proline residues (Pro24 and Pro170) break the respective N-terminal helices (α1 and α6) and, coupled with the more polar characteristics of the proximal residues contribute to the reentrant conformation, while conserved serine residues (Ser26 and Ser173) form the N-terminal cap of the C-terminal helices (α2 and α7) providing interactions with the peptide backbone that serve to orient the key catalytic residues (Ser27 and Arg174; discussed below) which are conserved in alignment (Fig. 2.9b,c). Preceding biochemical data involving mutation of a number of conserved, titratable residues have been shown to result in substantial decreases in phosphatase activity in *Ec*UppP and the majority of these map to these structural motifs and surrounding residues (Fig. 2.13)[106,164].

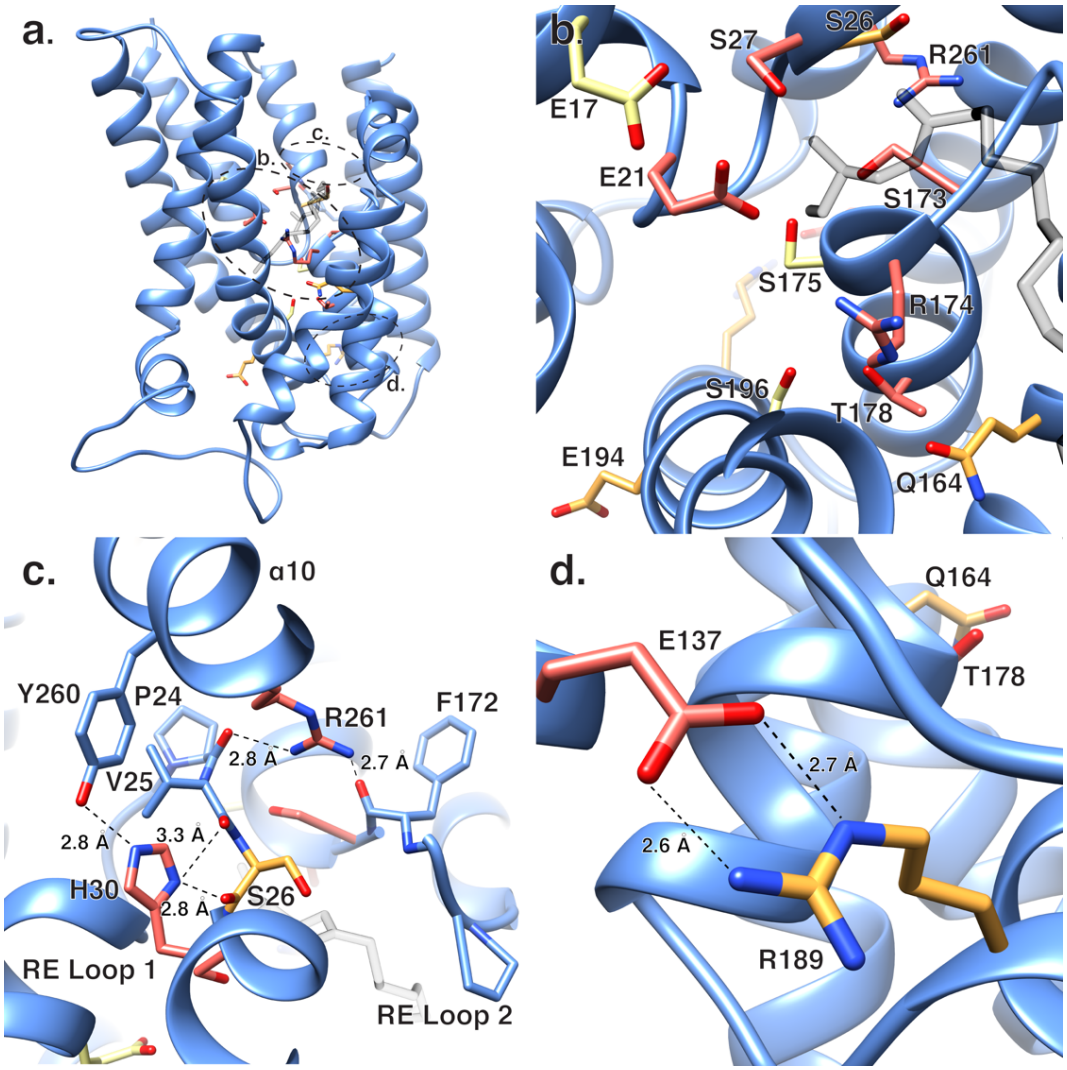
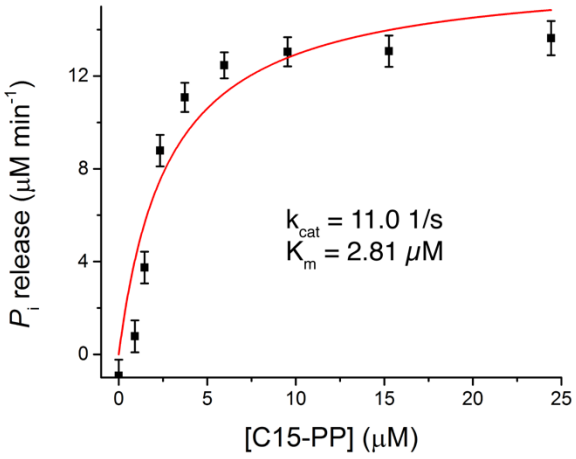


Figure 2.13 Structural mapping of published activity mutants.
a. Published activity mutants shown as stick and colored according to degree of effect from yellow (lowest) to pink (highest). Dotted circles highlight the zoomed regions in **b-d**.

The first reentrant helix-loop-helix contains conserved residues Glu17, Glu21, Ser26, Ser27 and His30, which were previously targeted for mutation[106,164]. Notably, separate studies proposed either Ser27 or His30 as the central nucleophile involved[106,164]. From our structure, Ser27 unambiguously fulfills this role with its side chain projecting from the N-terminal end of helix α2 into the active site where it forms a strong hydrogen bond (2.7 Å) with the secondary

alcohol of the bound monoolein glycerol head group (Fig. 2.10b). Our experimental data supports this with mutation of Ser27 to alanine completely abrogating activity (Fig. 2.14).



	This study	Chang *et al.*	Manat *et al.*
Wild-type	100%	100%	100%
S27A	0.020%	-	0.013%
H30A	27.0%	<1%	6.2%

Figure 2.14 Phosphatase activity of wild-type *Ec*UppP and comparison of wild-type activity to that of proposed catalytic mutants.
The phosphatase assay was performed as described in the Methods. Bars indicate one standard deviation from the mean absorbance measured in three replicates. Initial velocity data were fitted to the Michaelis-Menten equation using Origin data analysis software. Table represents relative activity to the wild-type enzyme.

The carboxylate side chains of both Glu17 and Glu21 are located directly proximal to Ser27, with Glu21 forming a direct hydrogen bond to its side chain hydroxyl that suggests a role in catalysis. His30, on the other hand, lies on the opposite side of helix α2, away from the active site, and appears to play a purely structural role (Fig. 2.10c, Fig. 2.13c), albeit one central to the optimal positioning of Ser27. The side chain imidazole forms hydrogen bonds with the carbonyl oxygens of Val25 and Ser26 (Fig. 2.10c). These hydrogen bonds serve to stabilize the 3_{10} nature in this N-terminal region of an otherwise classic α-helix, that is critical for appropriately orienting Ser27 into the substrate binding pocket. His30 further forms a strong hydrogen bond (2.8 Å) with the side chain hydroxyl oxygen of Tyr260, providing structural stabilization between helices α2 and α10 (Fig. 2.10c). Thus, the decreased phosphatase activity observed for His30Ala (Fig. 2.14)

is presumably due to destabilization from the loss of these key structural interactions, as well as the consequent effect on the sub-optimal positioning of the first reentrant loop in the active site. Both Tyr260 and the neighboring Arg261, mutation of which completely abrogates activity, are also very highly conserved and form multiple noncovalent interactions that further serve to orient both reentrant active site loops (Fig. 2.13c,d)[164]. The side chain guanadinium of Arg261 is remarkably fully buried within a largely hydrophobic interface at the juxtaposition of helices α2, α5, α7 and α10. Here it forms a hydrogen bond network with both reentrant loops at the carbonyls of the invariant Pro24 and Gly171 as well as with the adjacent residue (Phe172 in *Ec*UppP) and the backbone carbonyl of Leu126 on helix α5 (Fig. 2.10d, Fig. 2.15).

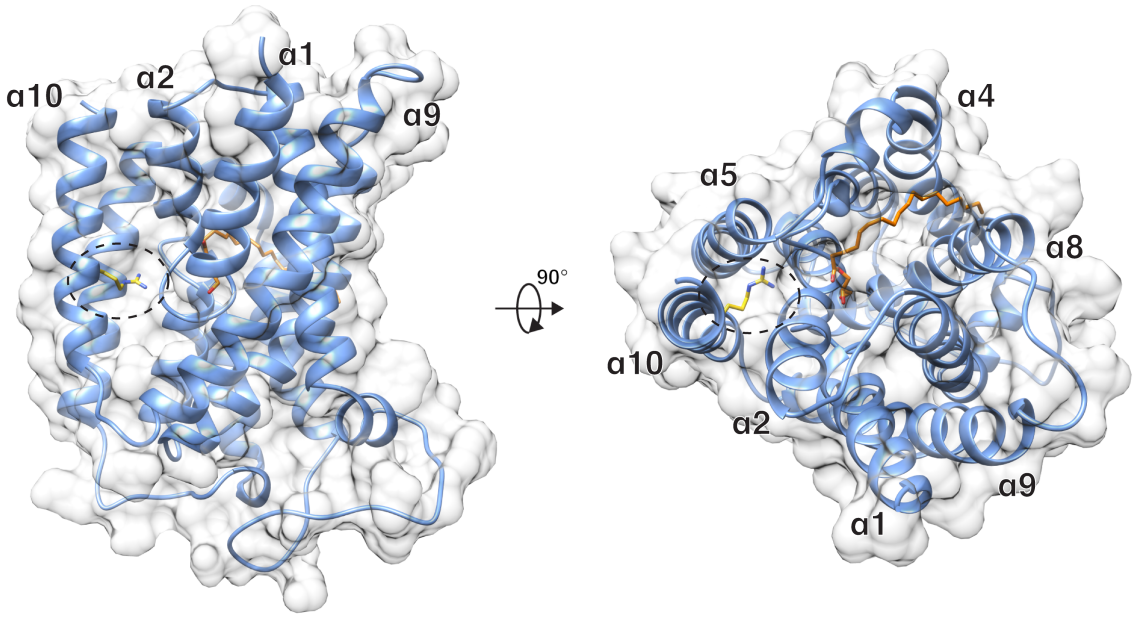


Figure 2.15 Role of Arg261 in structural coordination of the active site architecture.
Highly conserved Arg261 (highlighted by dashed circle) is buried in a hydrophobic interface at the membrane midplane and forms key interactions with the reentrant loops and helix α5 (see also Fig. 2.10d).

The second reentrant loop harbors a highly-conserved Ser-Arg-Ser motif (Fig. 2.9b,c). Remote homology detection with Phyre2[190] identifies similarity to the P-loop motif found in

dual specificity phosphatases (DUSPs)[191–193] where the role of the conserved arginine is coordination and electrostatic polarization of the phosphate moiety of a phosphothreonine or phosphotyrosine allowing a cysteine nucleophile (located at N-6 in the P-loop and not conserved in UppP), to attack and optimally promote subsequent hydrolysis and phosphate release[191–193]. In *Ec*UppP, Arg174 projects into the active site pocket in close proximity to the monoolein head group (Fig. 2.15) and is suitably positioned to fulfil the analogous role in cleavage of the C55 pyrophosphate (see below). Ser173 is the N-terminal capping residue for helix α7 with its side chain hydroxyl forming a hydrogen bond with the backbone carbonyl of Gly176. Mutation of Ser173 and Arg174 to alanine reduced activity ~50-fold and ~1000-fold respectively[106].

In addition to the interactions noted above, it appears that the antiparallel orientation of the four reentrant helices is partially maintained by structural waters, clearly defined in our 2.0 Å resolution maps, that are coordinated by the backbone amides of the residues making up the loop linking the reentrant helices.

2.4 Discussion

We propose the active site bound monoolein lipid we observe is acting as a C_{55}PP mimic. While monoolein lacks the negative charge associated with the native substrate pyrophosphate moiety, like C_{55}PP it has a hydrophilic headgroup and single hydrophobic tail. Overlay of the monoolein headgroup with a pyrophosphate molecule allows us to make inferences about potential interactions, illustrating that UppP utilizes the pseudosymmetry of the reentrant helices to coordinate the two phosphate moieties of the pyrophosphate headgroup at the N-terminus of the adjacent C-terminal reentrant helices α2 and α7 (Fig. 2.16a). The binding orientation is consistent with both a direct interaction with helix backbone amides, as suggested for binding of the Cl^- anion

to the ClC transporter[194], or via an electrostatic interaction with the potentially significant accumulated positive dipole moment at the N-termini of helices[195], here effectively doubled in impact by the disposition of the two inverted C-terminal reentrant helices buried in the middle of a membrane bilayer. Further electrostatic stabilization is realized by the Arg174 side chain guanadinium with the pyrophosphate moiety as well as weaker secondary interactions coming from hydrogen bonds with Glu21, Thr28 (Fig. 2.16b). The UppP phosphatase activity is metal dependent with a strong preference for the divalent cations magnesium and calcium, the latter giving optimal activity *in vitro*[106,164]. Density modelled as a water molecule is observed directly adjacent to the neutral monoolein head group and coordinated by hydrogen bonds with the primary and secondary alcohols of the glycerol moiety, as well as the main chain carbonyl oxygens of Leu23 and Thr20, and the side chain hydroxyls of the essential Ser173 and Ser175 (Fig. 2.16b,c).

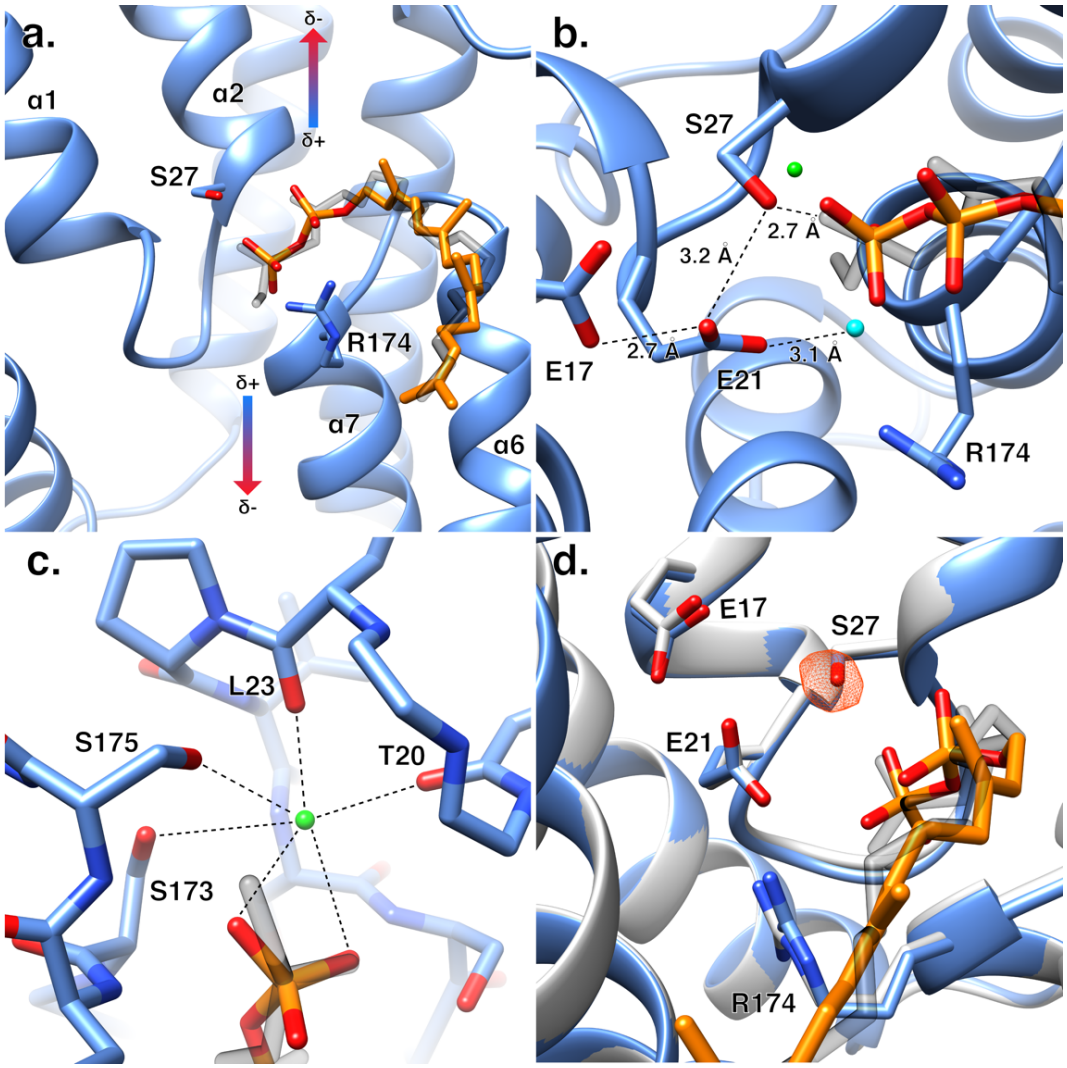


Figure 2.16 Modeling of C_{55}PP in the EcUppP active site.
a. Overlay of the C_{55}PP pyrophosphate on the observed monoolein headgroup illustrates the coordination of the pyrophosphate by the N termini of α2 and α7. **b.** Overlay of the C_{55}PP on the observed monoolein headgroup highlighting the interaction network of the key catalytic residues Ser27 (nucleophile), Glu21 (base), Glu17 (carboxyl–carboxylate pair), Arg174 (coordinates beta-phosphate). **c.** Observed coordination sphere of a bound water we propose may represent a potential binding site for the requisite catalytic cation. The overlaid pyrophosphate would contribute two ligands and a favorable electronegative stabilization to the bound metal. **d.** Overlay of wild-type (blue) and S27A mutant (gray) crystal structures show that the active site architecture is not perturbed by the S27A mutation. To confirm the identity the S27A mutant, the wild-type model was refined against the S27A data and an mFo-DFc map was calculated to show the negative difference peak for the Ser27 side chain hydroxyl (contoured at 3σ)

The hexadentate coordination number, ligand atom type and relative disposition suggest this could represent the putative cation binding site and the docked pyrophosphate positions the

terminal electronegative phosphate for a direct bidentate interaction with the putative cation, providing an apical and equatorial ligand as well as the only complementary electronegative charge in an otherwise electrostatically neutral coordination sphere (the lack of the analogous geometry and charge in the bound monoolein head group is likely the reason we have captured a water rather than a well occupied magnesium or calcium ion in our experimental structures). Indeed, it is known that pyrophosphate containing substrates in other enzyme systems serve to deliver bound catalytic magnesium or calcium ions to the active site[196]. Initial attempts to obtain C_{55}PP substrate or product complexes for both the native and a catalytic Ser27Ala mutant have been unsuccessful so far, and further experiments are required to confirm the substrate binding mechanism and metal binding site.

Based on our observations, we propose a catalytic mechanism for the phosphatase activity of UppP (Fig 2.17) in which Ser27 carries out a nucleophilic attack on the terminal phosphate of C_{55}PP, with the adjacent putative metal ion and Arg174 acting to coordinate, polarize and stabilize the electrophilic phosphocenter and subsequent pentavalent transition state; the nearby Glu21 carboxylate (hydrogen bond of 3.1 Å with the Ser27 side chain hydroxyl) is well positioned for a role as the activating general base (Fig. 2.16b). The close proximity of the Glu17 and Glu21 side chains (2.8Å) further suggests a potential role for a carboxyl-carboxylate interaction in modulating the pKa of Glu21 favorably for its general base role[197]. This facilitated attack of Ser27 results in the formation of a covalent phosphoserine intermediate that is subsequently hydrolyzed by an adjacent water to generate the C_{55}P and Pi reaction products and final regeneration of apo enzyme. A water molecule in our experimental maps is well positioned for this role, with again, general base assistance from the Glu21 carboxylate, its nearest hydrogen bonding partner (Fig. 2.16b). Kinetic analysis of our wild-type, S27A and H30A mutant forms (Fig 2.14) supports our

mechanistic scheme, as does the nearly identical overlap of the native and S27A mutant structures showing the kinetic effects we observe arise primarily due to the catalytic role of this absolutely conserved residue (Fig. 2.16d).

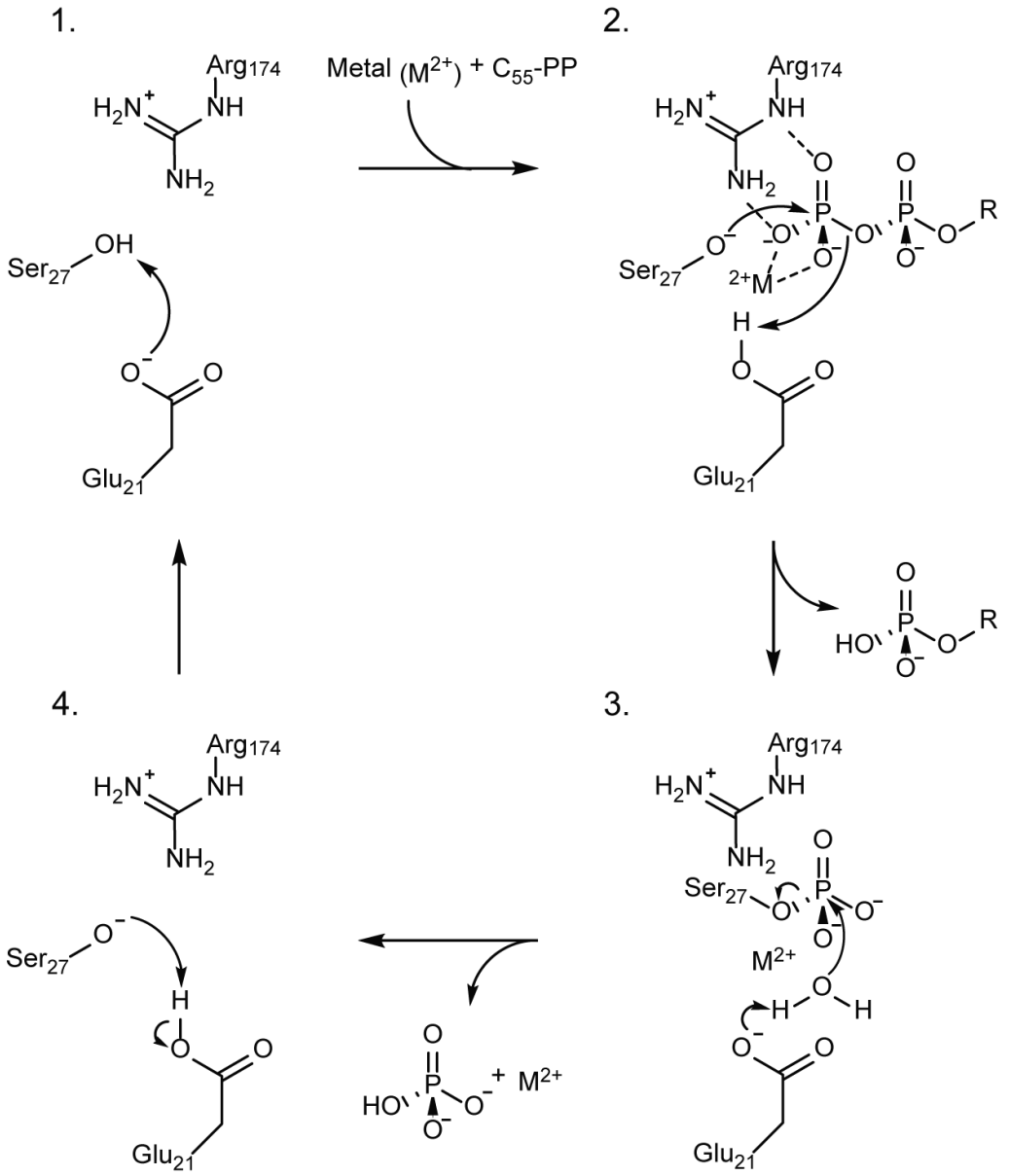


Figure 2.17 Catalytic mechanism of EcUppP phosphatase activity.
1. Activation of Ser27 by Glu21. **2.** Ser27 carries out a nucleophilic attack on the terminal phosphate of C_{55}PP, stabilized by Arg174 and a putative divalent cation (M2+), generating a phosphoenzyme intermediate, and releasing the C_{55}P product **3.** A water is activated for attack by Glu21 to hydrolyze the phosphoenzyme intermediate. **4.** Glu21 acts as a general acid to reprotonate Ser27, returning the enzyme to its starting state.

A search for structural similarity using DALI[198] failed to find any close structural homologues although remarkably the structure is similar to a previous prediction using co-evolutionary analysis coupled to Rosetta structure prediction[178]. Surprisingly, similarities to numerous proteins involved in the cross-membrane transport of small molecules were identified. Notably, similarity was detected to a ZIP zinc transporter[199] (Top hit, DALI Z score 4.1, RMSD 4.1 Å across 139 residues, PDB 5TSA); a major facilitator superfamily (MFS) member multidrug transporter MdfA[200] (DALI Z score 4.1, RMSD 6.0 Å across 140 residues, PDB 4ZP2); the eukaryotic chloride ClC transporter[201] (DALI Z score 3.6, RMSD 5.4 Å across 88 residues, PDB 3ORG); and the sodium-dependent citrate symporter CitS[202] (DALI Z score 3.5, RMSD 4.9 Å across 148 residues, PDB 5A1S). Comparison of these identified recurrent structural features including pseudosymmetry, inverted repeat topologies and, for the ClC transporter and CitS, the presence of reentrant helical/loop regions used to coordinate their anionic substrates similar to that observed here for *Ec*UppP. For CitS[202,203] and the MFS family of transporters[189,204], structures have been solved in discrete functional states highlighting the cycling between conformations open to opposing sides of the membrane to mediate transport. Indeed, this "alternating access" functionality is especially common in membrane proteins with interlocked inverted-repeats which exploit the pseudosymmetric arrangement to switch between conformations open to respective sides of the membrane[187]. Thus, the observed UppP topology - an interlocked inverted repeat with pseudosymmetry relating reentrant helical repeats – raises the intriguing possibility that UppP not only functions as a C_{55}PP phosphatase but concomitantly plays a role in the recycling of C_{55}P back into the bacterial cytoplasm; the identity of such a C-55P flippase remains elusive.

Several glycolipid flippases involved in the synthesis of the bacterial cell wall or outer membrane have been identified including TagGH for teichoic acids[162] and MsbA for lipopolysaccharide[205]. Notably, the identity of the lipid II flippase during peptidoglycan synthesis has been a matter of debate[206] with recent structural and experimental evidence supporting a role of MurJ as a lipid II flippase[47,166]. MurJ and UppP act on similar substrates (C_{55}PP-disaccharide-pentapeptide and C_{55}PP respectively) and comparison of their structures highlights several related features. MurJ consists of 14 transmembrane helices with α1-α6 and α7-α12 related by (distorted) pseudo-symmetry. The two lobes create a large cytoplasmic facing cavity with an exterior hydrophobic groove leading to a portal connecting to a strongly electropositive "proximal" site and weakly anionic "distal" site proposed to be the binding sites for the lipid II C55 tail, pyrophosphate and disaccharide-pentapeptide respectively. The portal and proximal site are similar in characteristics to the hydrophobic cleft and electropositive basin observed in UppP. The UppP structure also reveals a hydrophobic pocket situated below the active site cleft entrance and formed by amphipathic helix α3b, which is oriented parallel to the membrane plane and located at the cytoplasmic interface of the inner leaflet (Fig. 2.11b). The ordered tails of two monoolein lipids are observed in the pocket suggesting it may act as a binding site for the hydrophobic tail of the C_{55}PP substrate (Fig. 2.11b). These shared structural characteristics of the substrate binding sites are in agreement with the related lipid substrates of UppP and MurJ and provide further support for the proposed C_{55}PP binding mechanism to UppP. Whether this comparison can be further extended to a related lipid flippase function for UppP, as its structural characteristics intriguingly suggest, requires further investigation; however, we propose a potential model of carrier lipid flipping as shown in Figure 2.18.

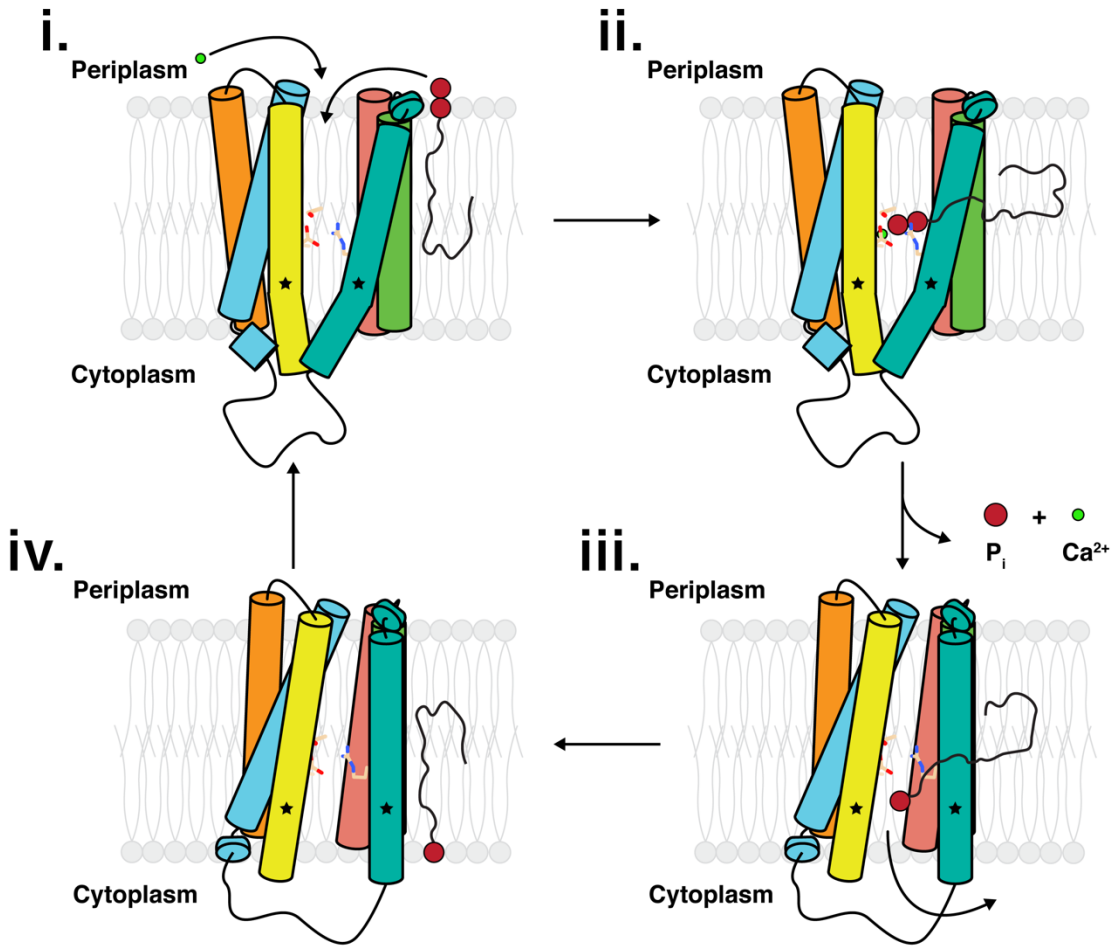


Figure 2.18 Potential flippase mechanism of UppP.
i. The $C_{55}PP$ enters the active site of UppP from the periplasmic leaflet of the inner membrane along with a Ca^{2+} ion. **ii.** UppP carries out its phosphatase activity on $C_{55}PP$, releasing an inorganic phosphate and the Ca^{2+} ion. **iii.** TM helices 5 and 8 straighten, with the strictly conserved Pro101 and Pro202 residues acting as hinges, creating a portal to the cytoplasmic leaflet of the inner membrane. **iv.** The recycled $C_{55}P$ exits the active site of UppP and partitions into the cytoplasmic leaflet of the inner membrane.

In summary, we have solved the structure of *E. coli* UppP at 2.0 Å resolution revealing an unexpected inverted topology repeat similar to many cross-membrane transporters and indicating the basis for phosphatase action deep within the midlayer of a bacterial membrane. Our results provide an important foundation on which to begin to further probe and understand the structural and functional mechanisms of this potential class of enzyme transporter and the design of antimicrobials that targets its essential role in virulence. Additionally, what remains unclear is how

C_{55}PP generated *de novo* in the cytoplasm by the pathway terminating at UppS is dephosphorylated in its final necessary stage for subsequent use as a lipid carrier. Would for example a lipid II flippase such as MurJ, which operates in the opposite direction, promiscuously serve to flip C_{55}PP to the periplasmic space for subsequent phosphatase and recycling action by UppP as previously suggested[106]? Or is there a possibility UppP could access C_{55}PP substrate from both faces at its internalized midlayer active site? These exciting and fundamental questions in bacterial cell wall biogenesis and lipid recycling are now made possible by the structural foundation provided here.

3 Structural characterization of *Bacillus subtilis* UppS inhibition

3.1 Introduction

The bacterial cell wall is a complex polymeric structure that is composed primarily of peptidoglycan (PG), teichoic acids (TAs), and in the case of Gram-negative bacteria, an outer membrane decorated with lipopolysaccharide (LPS)[207]. It provides bacterial cells with a barrier against the often harsh environments in which they live, defines their shape, and allows them to overcome turgor pressure. Due to the broadly conserved nature of its biosynthetic machinery and its absence from mammals, the bacterial cell wall has been a major focus for the development of antibiotics for decades[208]; however, as we continue to discover increasingly complex strategies that bacteria use to overcome our drugs, and resistance continues to spread, we must search for novel targets for the development of therapeutics[21,209].

A feature that is central to the biosynthesis of all bacterial cell wall polymers is the use of extended polyprenyl carrier lipids as scaffolds for their construction on the cytoplasmic leaflet of the inner membrane and subsequent transport to the periplasmic leaflet where the individual subunit can be incorporated into the existing cell wall[10]. The most commonly used carrier lipid in bacteria is undecaprenyl phosphate ($C_{55}P$). The synthesis of $C_{55}P$ is a two-step process, with the carrier lipid initially being synthesized in its pyrophosphate form ($C_{55}PP$) by undecaprenyl pyrophosphate synthase (UppS) before being dephosphorylated through an as yet unidentified mechanism. UppS is a *cis*-prenyltransferase that catalyzes the condensation of one molecule of farnesyl pyrophosphate ($C_{15}PP$) with eight molecules of isopentyl pyrophosphate (C_5PP) to yield di-*trans*-octa-*cis*-undecaprenyl pyrophosphate[10]. The active site of UppS is comprised of two substrate binding sites (S1 and S2 where C_{15} PP and C_5PP bind, respectively) and a long hydrophobic tunnel into which the nascent $C_{55}PP$ product grows during round of C_5PP

incorporation[10]. The entrance to the active site is decorated by conserved Arg residues that have been implicated in the binding of C_{15}PP and C_5PP in addition to an absolutely conserved Asp residue that is critical for catalysis. During catalysis, C_{15}PP binds to the S1 site and C_{15}PP binds to the S2 site in complex with Mg^{2+}. The conserved Asp is proposed to shuttle the Mg^{2+} to the C_{15}PP, allowing for attack of the C_{15}PP by C_5PP[10].

Based on the central role that C_{55}P plays in the biosynthesis of such a broad range of important cell wall polymers, it is of no surprise that UppS is an essential protein, and as such, stands as a promising drug target. In recent years, a series of reports have been published that describe the identification of promising drug leads that target UppS including clomiphene, an FDA approved drug for the treatment of infertility in women, which was identified via an antagonism screen that relied on the dispensability of WTA biosynthesis gene in Gram-positive bacteria[6-18]. WTA biosynthesis is initiated by TarO and TarA, the early-steps, before a series of late-steps, including the translocation of the WTA polymer by TarG from the cytoplasmic leaflet of the plasma membrane to periplasmic leaflet[5]. Inhibition of the late-steps is lethal in a wild-type background but becomes benign in strains with deletions or mutations in *tarO* or *tarA*[210]. This feature of WTA synthesis was exploited by screening for compounds that antagonized the activity of targocil, a lethal compound that targets TarG[35].

Subsequent work built upon the initial antagonism screen in order to develop a streamlined screening platform for the discovery of novel UppS inhibitors using *Bacillus subtilis* due to its amenability to high-throughput screening[37]. A library of 142,000 synthetic compounds were first screened for their ability to inhibit the growth of *B. subtilis* 186 before a secondary screen of antagonism towards the effects of targocil. After potency analysis of 181 targocil antagonists on solid media, 35 priority compounds were assessed for their ability to inhibit purified UppS from

B. subtilis (*Bs*UppS) *in vitro*, which further reduced the pool of potential inhibitors to five compounds[37]. Of these five compounds, MAC-0547630 was shown to be the most potent with an IC50 of ~50 nM against *Bs*UppS; however, it was significantly less effective against UppS from *S. aureus* and showed no inhibition of UppS from *E. coli*, so we sought to use a structural approach to explain the difference in efficacy and provide a basis for the structure guided design of more potent inhibitors against the *S. aureus* orthologue

Here we present the X-ray crystallographic structures of *Bs*UppS in the apo- form, as well as in complex with MAC-0547630 and JPD447, a more potent derivative of MAC-0547630 that we show can potentiate the effect of cefuroxime against MRSA (Fig. 3.1). Our structures reveal that both MAC-0547630 and JPD447 inhibit *Bs*UppS by binding within the hydrophobic cavity into which the C_{55}PP product of UppS grows, and that JPD447's improved inhibition is facilitated by its ability to make more favorable hydrophobic contacts. We also present the X-ray crystallographic structure of *Bs*UppS in complex with clomiphene (Fig. 3.1). This structure has improved density for the compound over our previously published co-structure with *Ec*UppS, allowing for a model of the entire molecule to be built and providing a basis for structure guided drug design.

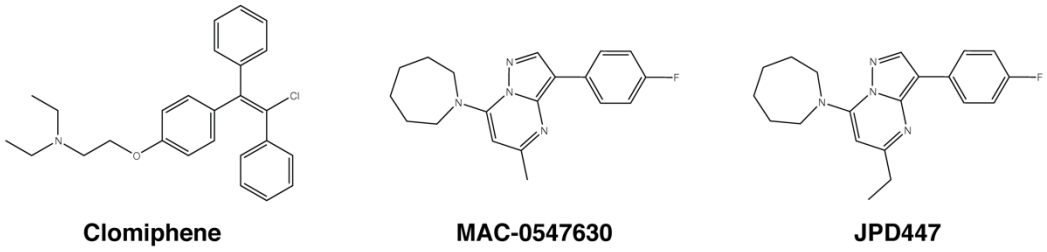


Figure 3.1 Chemical structures of UppS inhibitors.

3.2 Methods

3.2.1 Synthesis of JPD447

All reagents (>95% purity) were purchased from Sigma-Aldrich or Fisher Scientific and were used without purification unless noted otherwise. The chemical shifts for ^{1}H NMR spectra are given in parts per million (ppm) referenced to the residual proton signal of the deuterated solvent; coupling constants are expressed in Hertz (Hz). ^{13}C NMR spectra were referenced to the carbon signals of the deuterated solvent. Melting points were determined using a Fisher-Johns melting point apparatus and are uncorrected. High Resolution Mass Spectrometry (HRMS) analysis was performed at Queens University, Kingston, ON. For detailed synthetic steps, see Appendix A.

3.2.2 Checkerboard analyses and fractional inhibitory concentration index (FICI) determination

FICIs were determined by setting up standard checkerboard broth microdilution assays with two-fold serially diluted 8 (or 10) concentrations of each drug, using conditions based on the Clinical & Laboratory Standards Institute (CLSI) guidelines. Checkerboards were carried out in 96-well plates using Mueller Hinton Broth (MHB) and CA-MRSA USA300 as a test strain. Plates were incubated at 37°C for 18 hours and optical density read at 600nm using a Tecan plate reader. The MIC for each drug was the lowest concentration of drug showing <10% growth. The FIC for each drug was calculated as the concentration of a drug in the presence of co-drug for a well showing <10% growth, divided by the MIC for that drug. The FIC index is the sum of the two FICs. Interactions with FIC Index of less than 0.5 were deemed synergistic.

3.2.3 Cloning, expression, and purification of *Bs*UppS

The gene encoding full-length *Bacillus subtilis* UppS was synthesized by BioBasic and cloned into a modified pET28a vector encoding an N-terminal decahistidine tag and a thrombin cleavage site. *Bs*UppS was overexpressed in *E. coli* BL21 cells grown in ZYP-5052 autoinduction media at 37°C for 3 hours before lowering the temperature to 23°C and allowing growth to continue overnight[167]. Cells were harvested by centrifugation and resuspended in 20 mM Tris pH 7.5, 500 mM NaCl, 10% glycerol. Resuspended cells were lysed 2x using an EmsulsiFlex-C5 homogenizer (Avestin). Unlysed cells and debris were pelleted by centrifugation at 185,000 x *g* for 30 minutes. The resulting supernatant was loaded onto a 5 mL Ni-NTA Superflow (Qiagen) column pre-equilibrated with 20 mM Tris pH 7.5, 500 mM NaCl, 30 mM imidazole, washed with 60 mM imidazole, and *Bs*UppS was eluted with 250 mM imidazole. The purified protein was desalted into 20 mM Tris pH 7.5, 150 mM NaCl, and the decahistidine tag was removed by thrombin cleavage overnight. *Bs*UppS was further purified using a Superdex 10/300 GL column equilibrated in the same buffer. Monodisperse *Bs*UppS was pooled and concentrated to ~8 mg/mL in an Amicon Ultra-15 concentrator (Millipore) with a molecular weight cutoff of 30 kDa.

3.2.4 Crystallization and data collection

Sitting drop vapour diffusion crystallization trials for apo-*Bs*UppS were set up using a Crystal Phoenix (Art Robbins Instruments) robot and were incubated at room temperature. Initial hits were obtained in a number of conditions with the most promising crystals appearing overnight in JCSG+ condition A2 (0.1 M sodium citrate pH 5.5, 20% PEG 3000). Streak seeding was used to obtain optimized apo-*Bs*UppS crystals in 0.2 M sodium citrate pH 5.5, 16% PEG 3000. Crystals were harvested using nylon loops (Hampton Research) and were frozen in liquid nitrogen without

cryoprotectant. Crystals of *Bs*UppS in complex with clomiphene were obtained using the same method as above with the addition of 0.5 mM clomiphene to the protein 10 minutes prior to setting up crystallization experiments and addition of 20% glycerol to the crystallization condition to help maintain the clomiphene in solution. Crystals of *Bs*UppS MAC-0547630 and JPD447 complexes were obtained using the same method as apo- crystals with the addition of 0.4 mM inhibitor to the protein 10 minutes prior to setting up crystallization experiments. All X-ray diffraction data were collected at the Canadian Light Source (CLS) in Saskatoon, Saskatchewan. X-ray diffraction data for apo-*Bs*UppS were collected on Beamline 08B1-1 at a wavelength of 0.97874. X-ray data for inhibitor co-structures were collected on Beamline 08ID-1 at a wavelength of 0.97949 for MAC-0547630 and JPD-447, and at a wavelength of 0.97952 for clomiphene.

3.2.5 Data processing, structure solution, and refinement

Data for apo-*Bs*UppS, clomiphene and JPD447 inhibitor complexes were processed automatically using XDS through Autoprocess at the CLS[168,211]. Data for the MAC-0547630 complex were processed using xia2, DIALS, and Aimless[212–214]. The apo-*Bs*UppS structure was solved by molecular replacement using Phaser with the full length apo- *S. aureus* UppS structure (PDB ID code 3WYI) as a search model, with no truncations or modifications[33]. An initial model of *Bs*UppS was generated using AutoBuild and iterative rounds of manual and automated refinement were carried out using Coot and Phenix[171,172,215]. *Bs*UppS complexes were solved by molecular replacement using the refined apo-*Bs*UppS structure as a search model and were iteratively refined in the same manner. Models of each inhibitor were built in JLigand, and restraints were generated using the eLBOW GUI in Phenix[172,216,217]. Inhibitors were manually placed using Coot after refinement of the protein component of each complex and were

refined assuming full occupancy. mF_o-DF_c omit maps were generated by deletion of the ligands from their respective complex models followed by refinement with simulated annealing.

3.3 Results and discussion

3.3.1 X-ray crystallographic structure of *Bacillus subtilis* UppS

Bipyramidal crystals of apo-*Bs*UppS grew overnight and belonged to the space group $P4_32_12$ with unit cell dimensions 59.58 x 59.58 x 161.05 Å and one molecule in the asymmetric unit. Molecular replacement (MR) using the apo- *Staphylococcus aureus* UppS structure (PDB ID: 3WYI) as a search model yielded a single solution with an LLG of 658.677 and a TFZ of 30.8 [33]. Residues 20-260 could readily be built into the resultant electron density map, though residues 86-88 which correspond to a portion of the flexible loop at the mouth of the protein's active site, exhibited weak density. Data collection and final refinement statistics for all four *Bs*UppS crystals structures are presented in Table 3.1.

Table 3.1 UppS data collection and refinement statistics.

	Apo *Bs*UppS	*Bs*UppS – MAC-0547630	*Bs*UppS – JPD447	*Bs*UppS - Clomiphene
Data collection				
Wavelength (Å)	0.97874	0.97949	0.97949	0.97952
Space group	*P 43 21 2*	*C 2 2 21*	*P 43 21 2*	*P 43 21 2*
Cell dimensions				
a, b, c (Å)	59.576 59.576 161.053	85.723 87.399 159.024	60.115 60.115 161.067	58.501 58.501 161.858
α, β, γ (°)	90 90 90	90 90 90	90 90 90	90 90 90
Resolution (Å)	42.13 - 1.5 (1.554 - 1.5)	61.2 - 2.3 (2.382 - 2.3)	42.51 - 2.2 (2.279 - 2.2)	39.66 - 2.9 (3.004 - 2.9)
R_{pim}	0.030 (0.766)	0.023 (0.205)	0.0214 (0.396)	0.125 (0.903)
I / σI	14.1 (1.0)	16.1 (3.1)	17.3 (2.0)	7.70 (1.32)
$CC_{1/2}$	0.999 (0.489)	1.000 (0.953)	1.000 (0.920)	0.993 (0.7)
Completeness (%)	100.0 (99.7)	100.0 (100.0)	99.46 (99.35)	99.36 (99.37)
Redundancy	6.9 (5.0)	23.6 (23.5)	12.4 (12.8)	9.2 (9.5)
Refinement Statistics				
Resolution (Å)	42.13 - 1.5 (1.554 - 1.5)	61.2 - 2.3 (2.382 - 2.3)	42.51 - 2.2 (2.279 - 2.2)	39.66 - 2.9 (3.004 - 2.9)
No. reflections	47432 (4626)	26914 (2648)	15777 (1532)	61915 (6087)
R_{work} / R_{free}	0.1753 / 0.1982	0.2109 / 0.2492	0.2059 / 0.2383	0.2734 / 0.3315
No. atoms				
Protein	1989	3821	1926	1842
Ligand	7	74	25	29
Water	367	63	37	3
B factors (Å^2)				
Protein	28.5	71.6	67.7	62.1
Ligand	39.2	79.3	93.0	62.2
Water	40.7	57.2	59.0	32.4
r.m.s deviations				
Bond lengths (Å)	0.013	0.002	0.004	0.005
Bond angles (°)	1.25	0.56	0.63	0.81

*Numbers in parenthesis refer to the highest resolution shell.
† $R_{pim} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$

The overall structure of *Bs*UppS is very similar to those of other UppS orthologues, with a highly conserved six-stranded parallel β-sheet surrounded by α1-4 and α7 (Fig. 3.2a)[33,38,218–220]. The β-sheet core dominates structural alignment between *Bs*UppS and other published apo-UppS structures, with the majority of divergence between structures observed in the highly mobile

region linking strand β2 and helix α3, as well as in the N-terminal region of helix α3 itself (Fig. 3.2b). In comparison to both *E. coli* and *S. aureus* apo-UppS structures, helix α3 in the apo-*Bs*UppS structure adopts a conformation that is more akin to the "closed" conformations that have been observed in substrate bound structures of the other orthologues; however, this is potentially due to the presence of a putative polyethylene glycol molecule that occupies the hydrophobic tunnel (Fig. 3.2c,d). This seems to confirm the idea that the induced fit mechanism of substrate binding is mediated by hydrophobic interactions between helix α3 and the polyprenyl tail of C_{15}PP rather than any interaction with its pyrophosphate moiety[77].

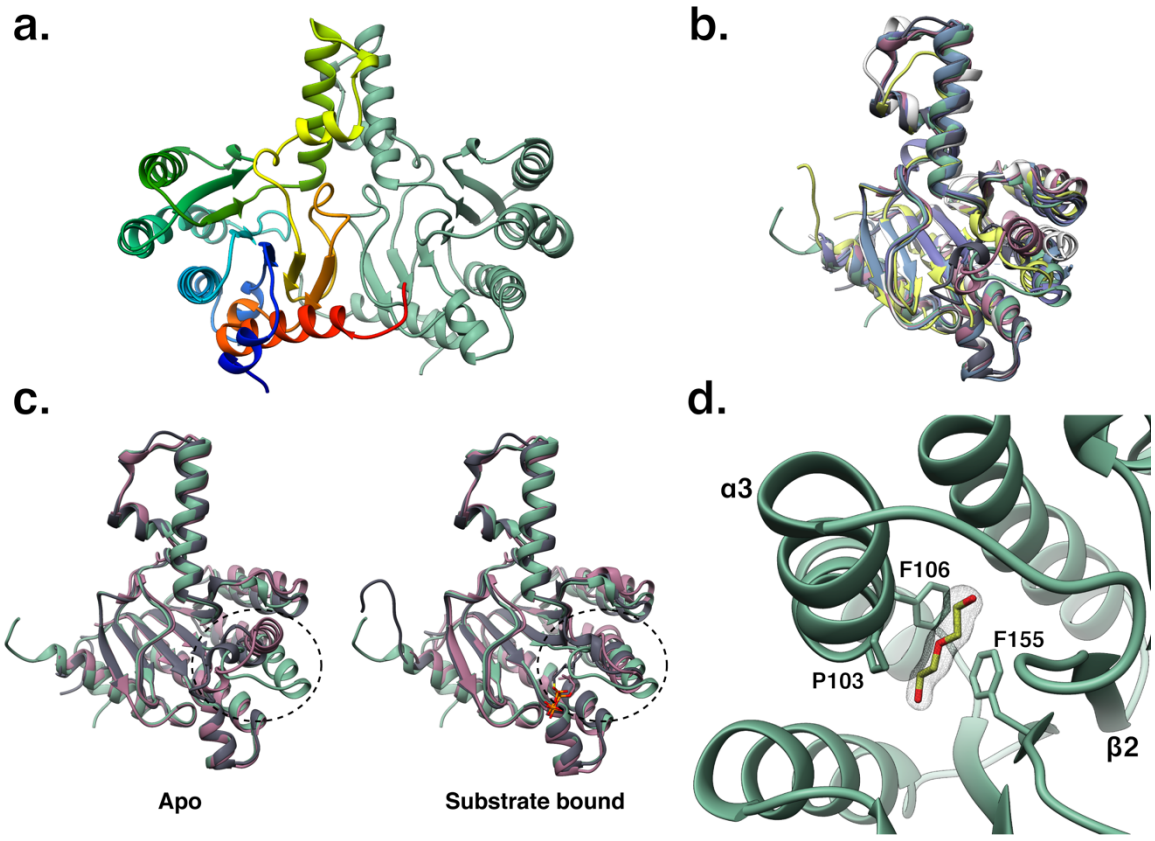


Figure 3.2 The X-ray crystallographic structure of *Bacillus subtilis* UppS.
a. Cartoon representation of *Bs*UppS dimer with one monomer colored rainbow from the N- (blue) to C-terminus (red). **b.** Cartoon overlay of all apo-UppS structures deposited in the PDB with *Bs*UppS showing the conserved β-sheet core and the positional divergence of helix α3. **c.** Comparison of apo-*Bs*UppS (mint) to apo- or substrate bound *Sa*UppS (slate) *Ec*UppS (pink). **d.**

mF_o-DF_c omit map contoured at 2.5σ showing the ordered portion of a PEG molecule in the *Bs*UppS hydrophobic tunnel.

3.3.2 Potentiation of cefuroxime by MAC-0547630 derivatives

As discussed in previous work, there has been a strong interaction demonstrated between inhibition of UppS and the sensitization of pathogens to β-lactam antibiotics[32,221]. The ability of MAC-0547630 to potentiate cefuroxime against CA-MRSA USA-300 was previously demonstrated using a checkerboard dilution assay (FICI of 0.078), and here we used a similar approach to screen a series of compounds that were derived from MAC-0547630 for their ability to do the same (Table 3.2)[37]. The structure activity relationship (SAR) series generated here was primarily focused on the pyrazolopyrimidine core of MAC-0547630. Of the compounds that were tested, three (JPD447, JPD448, and JPD519) demonstrated improved FICIs over MAC-0547630 when used in combination with cefuroxime against CA-MRSA USA 300.

Table 3.2 Combinations of MAC-0547630 and derivatives with cefuroxime against CA-MRSA USA 300.

Compound	MIC analog (µg/mL)*	FIC analog	MIC cefuroxime (µg/mL)*	FIC cefuroxime	FIC index
MAC-0547630	>16	0.0625	>256	0.0156	0.078
JPD447	>128	0.0312	>32	0.0156	0.0468
JPD519	>128	0.0312	>32	0.0312	0.0625
JPD448	>128	0.0312	>32	0.0312	0.0625
JPD451	>128	0.0625	>32	0.0312	0.0937
JPD591	>128	0.0625	>32	0.0625	0.125
JPD593	>128	0.0625	>32	0.0625	0.125
JPD594	>128	0.0625	>32	0.0625	0.125
JPD664	>128	0.0625	>32	0.0625	0.125
JPD518	>128	0.03125	>32	0.125	0.156
JPD457	>128	0.0625	>32	0.125	0.187
JPD613	>128	0.125	>32	0.0625	0.1875
JPD657	>128	0.0625	>32	0.25	0.312
JPD592	>128	0.0625	>32	0.25	0.3125
JPD606	>128	0.25	>32	0.0625	0.3125
JPD658	>128	0.25	>32	0.5	0.75
JPD668	>128	0.25	>32	0.5	0.75
JPD516	>128	0.5	>32	0.5	1
JPD515	>128	1	>32	1	2
JPD452	>128	1	>32	1	2
JPD446	>128	1	>32	1	2
JPD449	>128	1	>32	1	2
JPD524	>128	1	>32	1	2
JPD520	>128	1	>32	1	2
JPD460	>128	1	>32	1	2
JPD517	>128	1	>32	1	2
JPD450	>128	1	>32	1	2
JPD607	>128	1	>32	1	2
JPD608	>128	1	>32	1	2
JPD609	>128	1	>32	1	2
JPD614	>128	1	>32	1	2
JPD665	>128	1	>32	1	2
JPD642	>128	1	>32	1	2
JPD641	>128	1	>32	1	2
JPD666	>128	1	>32	1	2
JPD643	>128	1	>32	1	2
JPD650	>128	1	>32	1	2

* Differences in maximum concentrations tested are due to inhibition data being combined from separate experiments.

3.3.3 MAC-0547630 and JPD447 co-structures

In order to understand the structural basis for UppS inhibition by MAC-0547630 and JPD447, we sought to co-crystallize *Bs*UppS in complex with the two inhibitors. Co-crystals of *Bs*UppS in complex with MAC-0547630 and JPD447 were both obtained in the same condition as apo-*Bs*UppS and displayed the same bipyramidal morphology; however, in contrast to both apo- and JPD447 bound *Bs*UppS, the crystals grown in the presence of MAC-0547630 belonged to the space group $C222_1$ and had two molecules in the asymmetric unit. The MAC-0547630 co-structure also contained a citrate molecule in the pyrophosphate binding site of each monomer in the asymmetric unit. Nucleation and growth of inhibitor bound *Bs*UppS crystals required streak seeding from apo-*Bs*UppS crystals, and the inhibitors displaced the PEG molecule from the hydrophobic tunnel (Fig..2). Overall, the apo-*Bs*UppS and inhibitor bound *Bs*UppS structures show little difference with RMSDs of 0.751 and 0.236 Å over 225 Cαs for the MAC-0547630 and JPD-447 co-structures *versus* apo-*Bs*UppS, respectively.

As with many other candidate inhibitors of UppS, both MAC-057630 and JPD447 bind within the hydrophobic cavity into which the C_{55}PP product of UppS grows **(**Fig. 3.3). Unsurprisingly, based on their similarity, both molecules bind to *Bs*UppS in the same position, with their largely planar nature allowing them to slide deep into the cavity and form an extensive network of hydrophobic interactions. The fluoro-substituted benzyl moieties of both molecules are positioned between the side chain rings of Pro103 and Phe155, allowing for π-π stacking with Phe155, with the fluorine atom directed into a hydrophobic pocket formed by Leu107, Ile123, and Ile138 (Fig. 3.3a,c). The pyrazolopyrimidine ring is positioned to form favourable hydrophobic interactions with the side chains of Ala83, Leu99, Phe106, and Leu157, while the azepane ring is butted up against Gly60 and the side chains of Met39, Met61, Val64, and Trp235 (Fig. 3.3b,d).

Based on substrate bound crystal structures of UppS from other species, the positioning of the inhibitors in the hydrophobic cavity interferes with the ability of the enzyme to effectively bind their C_{15}PP substrate. Overlay of the *Sa*UppS-C_{15}PP co-structure with our inhibitor co-structures shows that rather than interfering with any of the residues involved in binding of the pyrophosphate moiety of C_{15}PP, MAC-0547630 and JPD447 occlude the hydrophobic tunnel into which the polyprenyl tail of C_{15}PP extends upon binding (Fig. 3.4a)[32].

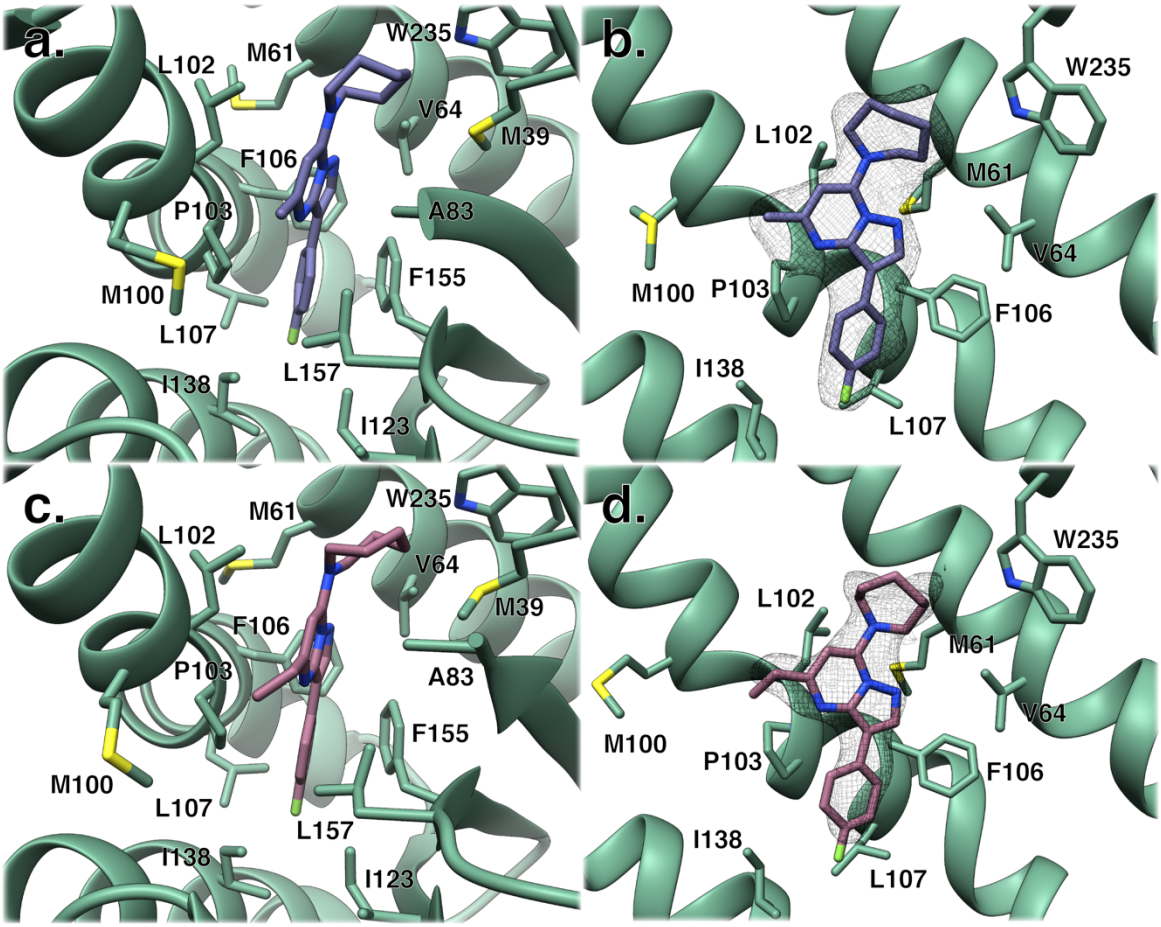


Figure 3.3 Co-crystal structures of MAC-0547630 and JPD447 with *Bacillus subtilis* UppS. **a.** Detailed view of the hydrophobic residues involved in MAC-0547630 binding. **b.** mF_o-DF_c omit map contoured at 2.5σ showing the well resolved density of MAC-0547630 in the *Bs*UppS hydrophobic tunnel. **c.** Detailed view of the hydrophobic residues involved in JPD447 binding. **d.** mF_o-DF_c omit map contoured at 2.5σ showing the density of JPD447 in the *Bs*UppS hydrophobic tunnel. Despite its apparent improved potency, the density for JPD447 was of lower quality than MAC-0547630.

Despite the wide array of derivatives that were generated from MAC-0547630, two of the three most effective potentiators of cefuroxime differed only in the alkyl substitution of the pyrazolopyrimidine ring from a methyl group to an ethyl or butyl group. JPD447 has an ethyl substitution on the pyrazolopyrimidine ring, which could facilitate more favourable hydrophobic interactions with key residues Leu102, Pro103, or Leu157 versus those achieved with MAC-0547630. Interactions with these residues were suggested to be important in previous work in which spontaneous mutants of *B. subtilis* 168 were generated that displayed reduced susceptibility to MAC-0547630[37]. Five unique *uppS* mutants were observed that resulted in a >10-fold increase in MAC-0547630 MIC: Phe84Tyr, Leu102His, Pro103Thr, Ala156Thr, and Leu157His. The electron density for Phe84 is weak in the JPD447 co-structure; however, in the MAC-0547630 co-structure, we are able to see that Phe84 is directed towards Met100 and the methyl substitution of the pyrazolopyrimidine ring, contributing to the network of hydrophobic interactions between protein and inhibitor (Fig 3.4b). Introduction of a hydroxyl group would result in a less favourable environment for the inhibitor to bind, similar in effect to the introduction of an imidazole moiety in the case of the Leu102 and Leu157 mutations. The Pro103Thr mutation would hinder the ability of the inhibitor to slide deep into the hydrophobic pocket to form favourable stacking interactions between the fluoro-benzyl moiety of the inhibitors and Phe155. Presumably, the Ala156Thr mutation would have a similar effect due to its proximity to Phe155 itself, as Ala156 is directed into the hydrophobic core of the protein and its mutation to Thr could result in a bulging of strand β4 to compensate, which would narrow the diameter of the hydrophobic tunnel.

The reduced potency of MAC-0547630 against *S. aureus* and *E. coli* UppS orthologues *versus* UppS from *B. subtilis* can also be explained by our co-structures. Of the five key residues

identified above, Leu157 is the only one not conserved between *B. subtilis* and *S. aureus*, with the latter orthologue containing an isoleucine in its place (Fig. 3.4c). This substitution results in less steric bulk, resulting in a weaker hydrophobic interaction with the inhibitor; thus, the increased potency of the JPD447 inhibitor against MRSA in combination with cefuroxime could largely be mediated by an improved interaction with Ile148 of *S. aureus* UppS. In the *E. coli* orthologue, Leu157 is substituted with an alanine, which likely abrogates a key hydrophobic interaction between the alkyl substituents of the inhibitors and strand β4. The most pronounced differences between the two Gram-positive UppS orthologues tested in previous literature and the *E. coli* orthologue are the substitutions of Pro103 and Phe155 with a phenylalanine and an isoleucine, respectively (Fig. 3.4d). The Pro→Phe substitution completely occludes the site between residues 103 and 155 at which the fluoro-substituted phenyl rings of both inhibitors bind and the Phe→Ile substitution eliminates the potential for π-π stacking interactions that appear to be critical for inhibitor binding; therefore, it's unsurprising that MAC-0547630 was unable to inhibit *Ec*UppS.

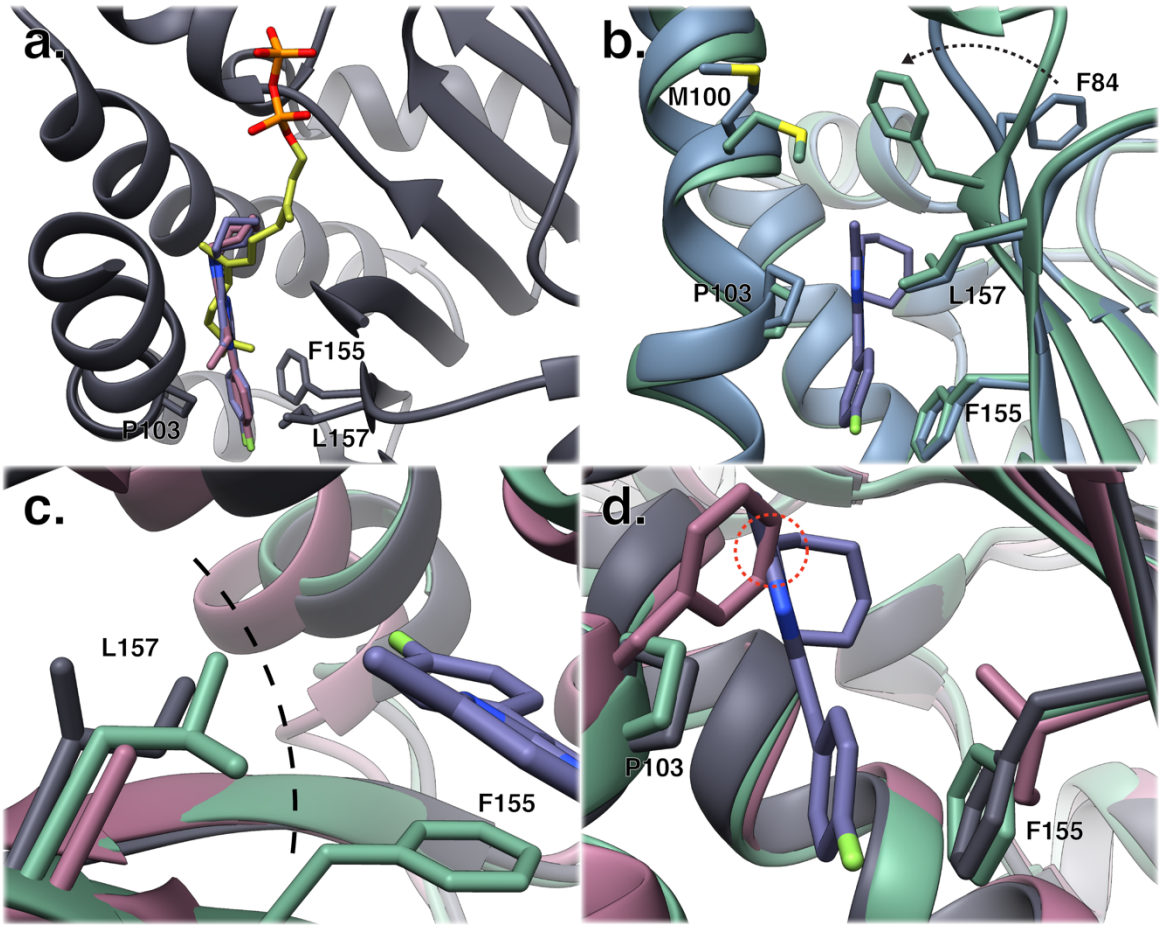


Figure 3.4 Molecular rationale for observed differences in UppS inhibition.
a. Overlay of MAC-0547630 (purple) and JPD447 (pink) with the polyprenyl tail of C_{15}PP (yellow) in the active site of *Sa*UppS. **b.** Overlay of apo-*Bs*UppS (blue) and MAC-0547630 bound *Bs*UppS (mint) structures, showing the rearrangement of the entrance loop to orient Phe84 towards Met100 and the methyl group of MAC-0547630. **c.** Overlay of the MAC-0547630 bound *Bs*UppS structure with *Sa*UppS (slate) and *Ec*UppS (pink), showing the substitutions for the key residue Leu157 which lead to differences in potency. The black dashed line emphasizes the differences in steric bulk. **d.** Comparison of MAC-0547630 bound *Bs*UppS structure with *Sa*UppS and *Ec*UppS structures reveals a structural basis for MAC-0547630's inability to inhibit *Ec*UppS. Red dashed circle highlights the clash between inhibition and the Phe substitution for Pro.

3.3.4 Clomiphene co-structure

Our lab previously crystallized clomiphene in complex with *Ec*UppS and solved the structure to 2.1 Å resolution; however, density for clomiphene was weak and did not allow for a model of entire molecule to be included in refinement[35]. Our co-structure of clomiphene in complex with *Bs*UppS allows for less ambiguous placement of the entire inhibitor, despite having

a more modest resolution of only 2.9 Å (Fig. 3.5a). As with the other inhibitor co-structures, the clomiphene *Bs*UppS co-structure shows very little difference to the apo-*Bs*UppS structure, with an RMSD of 0.541 Å over 225 Cαs, and the clomiphene displaced the PEG molecule from the hydrophobic tunnel.

The chloro-substituted triphenethylene moiety of clomiphene is directed into the hydrophobic cavity of the protein in the same location as it was found in the *Ec*UppS-chlomiphene co-structure, and approximately in the same location that both MAC-0547630 and JPD447 occupy (Fig 3.5b). The hydrophobic interaction network formed between clomiphene and *Bs*UppS is very similar to that of the inhibitors from the previous section, with two of the phenyl rings of the triphenethylene moiety occupying the same region as the fluoro-substituted phenyl ring and pyrazolopyrimidine rings (Fig 3.5b). One phenyl ring stacks between the side chains of Pro103 and Phe155, with the side chains of Phe106, Ile138, and Leu157 contributing additional hydrophobic interactions, while the other projects between strand β2 and helix α2, packing up against Gly60 and forming further hydrophobic interactions with the side chains of Met61, Val64, and Leu102 (Fig. 3.5c). The *N*,*N*-diethyl-2-methoxyethanamine tail, which was directed between helices α2/α3 in the *Ec*UppS co-structure is oriented nearly 180° in the opposite direction towards the loop connecting strand β2 and helix α3 in the *Bs*UppS co-structure, forming hydrophobic interactions with Met100 and Phe84 (Fig. 3.5c,d). In our improved co-structure the loop connecting strand β2 and helix α3 remains almost entirely disordered, with the *N*,*N*-diethyl-2-methoxyethanamine tail pointed towards it. The extended tail could act as a locus for derivatization as there are several residues on the adjacent loop that are strictly conserved, though the flexible nature of the loop could impose too much of an entropic penalty on binding for this to be a viable strategy.

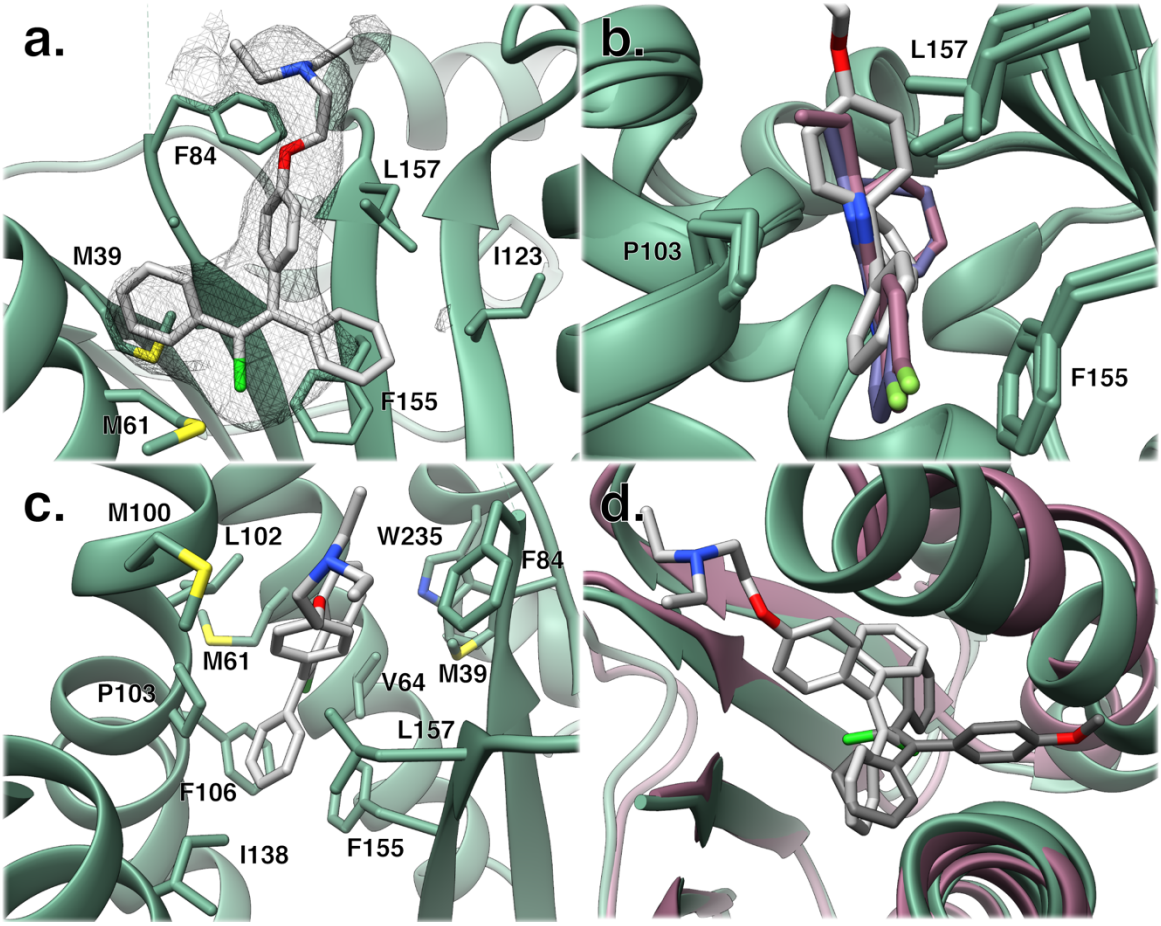


Figure 3.5 Co-crystal structure of clomiphene with *Bacillus subtilis* UppS.
a. mF_o-DF_c omit map contoured at 2.5σ showing the more clearly resolved density for clomiphene in the *Bs*UppS co-structure. **b.** Comparison of clomiphene binding to that of MAC-0547630 and JPD447. **c.** Detailed view of the hydrophobic residues involved in clomiphene binding. **d.** Comparison of the position of clomiphene in the *Bs*UppS co-structure (light grey) vs that in the *Ec*UppS co-structure (dark grey).

3.4 Conclusion

UppS has emerged as a promising target for the development of novel inhibitors due to its central, indispensable role in the biosynthesis of many bacterial cell wall polymers. While a number of inhibitor classes have been identified using *in silico* approaches, a recently developed high-throughput screening platform allowed for the identification of whole cell-active compounds against UppS. One of the compounds that was identified using the screening platform, MAC-

0547630, emerged as a top candidate for use as a starting point in an SAR study. Among the molecules generated for the SAR series, JPD447 was the most promising. Analysis of our *Bs*UppS co-structures indicates that MAC-0547630 and JPD447 bind to UppS primarily through hydrophobic interactions with residues that are well conserved in Gram-positive bacteria.

The identification of clomiphene as an inhibitor of UppS was exciting due to its well characterized toxicology and pharmacological profile, which lowers the barrier for its potential use as a novel antibacterial drug. Our improved crystal structure of clomiphene in complex with *Bs*UppS has allowed us to model the entire molecule, showing that it binds in a very similar fashion to the compounds identified using the high-throughput screening approach. Based on this observation, clomiphene will likely act as a better inhibitor of Gram-positive bacteria than Gram-negative.

Taken together, our crystal structures, combined with our SAR data, provide a foundation for structure guided drug design of more potent UppS inhibitors in the future.

4 Summary and future directions

In the work contained within this thesis, we took a structural approach to characterize two proteins that are central to bacterial cell wall synthesis with the overall aim of using the structural information to facilitate the development of novel antibiotics in the face of the ever-growing threat of antibiotic resistance. We have presented the X-ray crystallographic structure of a polytopic integral membrane protein that plays a central role in the metabolism of all bacterial cell wall polymers. Our structure of UppP reveals a unique fold with similarities to ion channels and transporters, allowing for the possibility that UppP acts as a $C_{55}P(P)$ flippase. Combined with site directed mutagenesis and kinetic assays, the high-resolution structural information gleaned from our crystal structure allowed us to propose a mechanistic basis of UppP's phosphatase activity. We have also presented the structure of UppS from *B. subtilis* both in its apo- form and bound to three different small molecule inhibitors, which has allowed us to rationalize the difference in potency between two closely related inhibitors, as well as provide a foundation for future structure guided drug design.

4.1 UppP

4.1.1 Putative flippase activity

Our X-ray crystallographic structure of UppP represents a landmark in the field of carrier lipid biology as it is the first structure of a unique class of carrier lipid phosphatases and represents a platform from which myriad questions can be asked; however, the most pressing question must be whether or not UppP has the ability to act as a flippase in the recycling of the carrier lipid.

From a purely structural standpoint, capturing UppP in its proposed inward facing conformation would be ideal. In the case of MurJ, the lipid II specific flippase used in

peptidoglycan synthesis, five distinct conformations of the transporter were captured using LCP crystallography[222]. The authors had observed a chloride ion in their previous MurJ structure that they thought could be restricting the protein in a particular conformation, so developed a purification strategy in which chloride was removed from all buffers. Using this optimized purification, and lipid II doped monoolein, they were able to capture five distinct conformations of MurJ; furthermore, they noted that the key determinant of which crystal form was obtained was the concentration of salt in their crystallization condition, lending credence to their hypothesis about chloride restricting the conformation of the protein.

In the case of UppP, there is a monoolein molecule in the active site that could be restricting its movement. Removal of monoolein obviously presents a more difficult challenge than removing chloride, as it makes up the bulk lipid in which the protein is crystallized, though we could attempt incorporating phospholipids such as distearoyl phosphatidylglycerol (DSPG) into the monoolein, as this has been shown to modulate the crystal packing and molecular contacts made in the case of one ligand-gated ion channel[156]. An alternative approach would be to try altering the protein construct for crystallization by mutating a pair of arginine (Arg145/155) residues that are seen to bind two very well-ordered sulfate ions in our crystal structure. Analysis of all the conditions in which crystals of varying quality were grown show that sulfate is present in all of them. The coordination of sulfate ions by Arg145/155 thus appears to be a strong determinant of crystal formation with wild-type UppP, and mutation of these residues, neither of which are conserved, could open up more chemical space for the formation of high-quality crystals.

Beyond capturing a static structure of UppP in an alternative state, we could attempt to use a similar approach as used for several lipid flippases/scramblases to detect flippase activity[223,224]. In such assays 7-nitro-2,1,3-benzoxadiazol (NBD) acyl-labelled phospholipids

are incorporated into the preparation of proteoliposomes containing the lipid flippase of interest and dithionite is added to the reaction mixture, resulting in a decrease of the fluorescence signal as the NBD moieties in the outer membrane leaflet are reduced. A similar assay could be attempted with NBD-C_{15}PP, which is the only commercially available NBD-labelled polyprenyl phosphate lipid; however, a negative result from such an assay could be due a reliance on the longer lipid tail of C_{55}PP rather than indicative of a lack of flippase activity. In order to overcome this, the protocol for synthesis of C_{15}PP could be modified to generate NBD-C_{55}PP, allowing for the detection of putative flippase activity with correct length polyprenyl lipid tail[225].

Another potential avenue for assessing the potential of flippase activity is to measure the changes in global conformational dynamics upon addition of substrate. Several recent publications have taken advantage of hydrogen-deuterium exchange mass spectrometry (HDX-MS) to study the dynamics of transporter proteins[226–229]. Presumably if UppP were utilizing the alternating access model of transport that we propose, and that is common to the MFS transporters with which UppP shares a great deal of similarity, there would be increased HDX upon addition of substrate for residues that are otherwise buried in the core of the protein[204]. We have made initial attempts with collaborators to apply HDX-MS to the study of UppP but have thus far failed at obtaining sufficient peptide coverage to proceed with HDX experiments, though further digestion trials are underway with a panel of aspartic proteases that have been shown to yield better coverage than the porcine pepsin that is generally used for digestion in HDX-MS experiments[230]. Whether we are able to identify flippase activity or not, the combination of HDX-MS data with molecular dynamics (MD) simulations could allow for identification of key protein-lipid interactions, which would give a better understanding of how UppP's environment affects its function and perhaps provide a rationale for its dimerization[231].

4.1.2 Substrate/product co-structure

While the existence of an induced fit mechanism in the transmembrane PAP2 C_{55}PP phosphatases remains to be proven, site directed mutagenesis studies, coupled with crystal structures of soluble PAP2 enzymes in complex with substrate analogs, have allowed for their actual mechanism of phosphate hydrolysis to be inferred with some confidence[95,99–101,232–234]. In the case of UppP, our proposed mechanism of action is based on site directed mutagenesis as well, but the structural data to support the mechanism is limited to overlay of a C_{55}PP pyrophosphate moiety with the glycerol headgroup of a well ordered monoolein. We have been attempting to capture these complexes by using LCP crystallography for some time now; however, we have been hindered by the very thing that allowed us to solve the crystal structure in the first place – the monoolein lipid used in LCP crystallography.

Setting up crystallization experiments with high concentrations of geranylgeranyl pyrophosphate, the longest chain aqueous polyprenyl phosphate that is commercially available, yields high quality crystals that diffract well, but the geranylgeranyl pyrophosphate is unable to displace the monoolein from the active site. We have also attempted an approach similar to that used in the crystallization of ArnT whereby the monoolein is doped with C_{55}PP at a concentration of 1% (w/w)[235]. Again, we were not able to displace the monoolein from the active site of UppP. Finally, we have also attempted to mutate the active site Ser to a Cys residue in the hope that the nucleophilic attack on the terminal phosphate would result in a stable intermediate that we could capture in the crystal, but the mutation seems to disrupt the stability of the protein, yielding a poor quality size exclusion chromatography trace and no crystals whatsoever. Future crystallization experiments that are geared towards obtaining a substrate or product co-structure should likely focus on identifying a novel crystallization condition for UppP

that doesn't require the binding of monoolein in the active site of the protein and preferably one with a more physiological pH, as the moderately acidic environment of our current crystallization conditions (pH 4-5) could potentially be interfering with electrostatic interactions that are required for proper binding of the pyrophosphate moiety. Finally, if crystallization of UppP requires the presence of lipids as in LCP, but monoolein is unable to be displaced from the active site, the HiLiDe method of crystallization could be attempted for capture of substrate or product complexes[236].

4.1.4 Structure guided drug design

As stated above, one of the principal reasons for pursuing the structural characterization of UppP was to facilitate the structure guided design of novel antibiotic therapies. To that end, we have been also been working to capture UppP in complex with a series of phenylthiazole inhibitors that were recently shown to inhibit with the phosphatase activity of UppP with IC50 values in the nanomolar range[28,29,39]. As with the substrate complexes that we have been trying to capture, however, the displacement of monoolein from the active site is a serious barrier to success. In addition, the inhibitors themselves are almost entirely insoluble in aqueous solution, so we have been taking an approach where the inhibitors are dissolved in the molten monoolein used for LCP set up. Some crystals have been obtained, but they have been of poor quality and further optimization will be required; nonetheless, our crystal structure was used for docking studies of the phenylthiazole inhibitors and may already be contributing to the design of improved compounds[39].

4.2 UppS

The initial aim of the work on UppS described in this thesis was to obtain co-crystal structures of MAC-0547630 with both *Bs*UppS and *Sa*UppS, yet *Sa*UppS remained recalcitrant to crystallization despite numerous and varied attempts to identify novel crystallization conditions or reproduce those that have been previously published[32,33]. Thus, while our *Bs*UppS co-structures with both MAC-0547630 and JPD447 allow for rationalization of why the former inhibitor is so much more potent against *Bs*UppS than *Sa*UppS and why the latter inhibitor is better able to potentiate the antibiotic activity of cefuroxime against MRSA, the most logical next step is to obtain co-structure of both inhibitors with *Sa*UppS to confirm our interpretation of the structural data. MAC-0547630 was identified via a high-throughput screening platform, and its more potent JPD447 inhibitor was developed using somewhat of a shotgun approach, altering the different functional groups of the pyrazolopyrimidine core without any structural information for guidance, so it is reasonable to imagine that with our *Bs*UppS co-structures and a future *Sa*UppS co-structure, the compound could be altered to be made even more potent, especially against *S. aureus*. Beyond further optimization of the inhibitor itself, a logical next step would be to test these compounds in an animal model of infection to determine their efficacy and safety.

Because both MAC-0547630 and JPD447 appear to act in the same manner as clomiphene, that is blocking the hydrophobic tunnel into which the C_{55}PP product of UppS grows, it stands to reason that the development of clomiphene as an antibiotic should be prioritized over the other compounds due to its well established safety profile. Our co-structure of *Bs*UppS with clomiphene allowed us to build a model of the entire compound, but the modest 2.9 Å resolution and mobility of the entrance loop make it difficult to predict how exactly the molecule should be functionalized. The presence of two absolutely conserved polar residues (Ser85/Asn88) on the entrance loop do

make the *N*,*N*-diethyl-2-methoxyethanamine tail of clomiphene, which is directed towards the entrance loop, a promising moiety for functionalization that could keep the protein in a "closed" conformation.

Finally, efforts should be made to capture co-crystal structure of UppS with the phenylthiazole inhibitors described above, as they have been shown to act on UppS as well[28,29,39]. Dual targeting of both the synthesis and recycling of C_{55}PP is a tantalizing strategy for the development of antibiotics as it would completely halt the ability of the bacteria to synthesize any of their vital cell wall polymers.

Bibliography

[1] W. Vollmer, D. Blanot, M.A. De Pedro, Peptidoglycan structure and architecture, FEMS Microbiol. Rev. 32 (2008) 149–167. https://doi.org/10.1111/j.1574-6976.2007.00094.x.

[2] A.L. Lovering, S.S. Safadi, N.C.J. Strynadka, Structural Perspective of Peptidoglycan Biosynthesis and Assembly, Annu. Rev. Biochem. 81 (2012) 451–478. https://doi.org/10.1146/annurev-biochem-061809-112742.

[3] A.J. Meeske, P. Eammon, P. William, S. Walker, A.C. Kruse, T.G. Bernhardt, D.Z. Rudner, SEDS proteins are a widespread family of bacterial cell wall polymerases, (2016). https://doi.org/10.1038/nature19331.

[4] S. Leclercq, A. Derouaux, S. Olatunji, C. Fraipont, A.J.F. Egan, W. Vollmer, E. Breukink, M. Terrak, Interplay between Penicillin- binding proteins and SEDS proteins promotes bacterial cell wall synthesis, Nat. Publ. Gr. (2017) 1–13. https://doi.org/10.1038/srep43306.

[5] S. Brown, J.P. Santa Maria, S. Walker, Wall Teichoic Acids of Gram-Positive Bacteria, Annu. Rev. Microbiol. 67 (2013) 313–336. https://doi.org/10.1146/annurev-micro-092412-155620.

[6] C.R.H. Raetz, C. Whitfield, Lipopolysaccharide Endotoxins, Annu. Rev. Biochem. 71 (2002) 635–700. https://doi.org/10.1146/annurev.biochem.71.110601.135414.

[7] C. Whitfield, Biosynthesis and Assembly of Capsular Polysaccharides in Escherichia coli , Annu. Rev. Biochem. 75 (2006) 39–68. https://doi.org/10.1146/annurev.biochem.75.103004.142545.

[8] E.P. Kennedy, Osmotic regulation and the biosynthesis of membrane-derived oligosaccharides in Escherichia coli (periplasmic space/fixed anions/membrane phospholipids/outer membrane), Proc. Natl. Acad. Sci. USA. 79 (1982) 1092–1095. https://www.pnas.org/content/pnas/79/4/1092.full.pdf.

[9] H.-M. Kuhn, U. Meier-Dieter, H. Mayer, ECA, the enterobacterial common antigen, FEMS Microbiol. Lett. 54 (1988) 195–222. https://doi.org/10.1111/j.1574-6968.1988.tb02743.x.

[10] G. Manat, S. Roure, R. Auger, A. Bouhss, H. Barreteau, D. Mengin-Lecreulx, T. Touzé, Deciphering the metabolism of undecaprenyl-phosphate: the bacterial cell-wall unit carrier at the membrane frontier., Microb. Drug Resist. 20 (2014) 199–214. https://doi.org/10.1089/mdr.2014.0035.

[11] P. Sarkar, V. Yarlagadda, C. Ghosh, J. Haldar, S. Yuan, R. Altmeyer, G. Zou, J. Haldar, J. Bartlett, K. Kaniga, K.M. Krause, M. Leadbetter, M.S. Linsell, D.G. Marquess, E.J. Moran, M.B. Nodwell, J.L. Pace, S.G. Trapp, S.D. Turner, C. Chen, K. Lewis, A review on cell wall synthesis inhibitors with an emphasis on glycopeptide antibiotics, Med. Chem. Commun. 8 (2017) 516–533. https://doi.org/10.1039/C6MD00585C.

[12] K. Poole, Resistance to β-lactam antibiotics, Cell. Mol. Life Sci. 61 (2004) 2200–2223. https://doi.org/10.1007/s00018-004-4060-9.

[13] D.J. Tipper, J.L. Strominger, Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine., Proc. Natl. Acad. Sci. 54 (1965) 1133–1141. https://doi.org/10.1073/pnas.54.4.1133.

[14] S. Sobhanifar, D.T. King, N.C.J. Strynadka, Fortifying the wall: Synthesis, regulation and degradation of bacterial peptidoglycan, Curr. Opin. Struct. Biol. 23 (2013) 695–703. https://doi.org/10.1016/j.sbi.2013.07.008.

[15] Y. Takahashi, M. Igarashi, T. Miyake, H. Soutome, K. Ishikawa, Y. Komatsuki, Y. Koyama, N. Nakagawa, S. Hattori, K. Inoue, N. Doi, Y. Akamatsu, Novel semisynthetic antibiotics from caprazamycins A-G: Caprazene derivatives and their antibacterial activity, J. Antibiot. (Tokyo). 66 (2013) 171–178. https://doi.org/10.1038/ja.2013.9.
[16] A.L. Lovering, L.H. De Castro, D. Lim, N.C.J. Strynadka, Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis, Science (80-.). 315 (2007) 1402–1405. https://doi.org/10.1126/science.1136611.
[17] L.D. Saravolatz, G.E. Stein, L.B. Johnson, Telavancin: A novel lipoglycopeptide, Chinese J. Infect. Chemother. 11 (2011) 313. https://doi.org/10.1086/648438.
[18] C. Chatterjee, M. Paul, L. Xie, W.A. van der Donk, Biosynthesis and mode of action of lantibiotics, Chem. Rev. 105 (2005) 633–683. https://doi.org/10.1021/cr030105v.
[19] K.J. Stone, J.L. Strominger, Mechanism of action of bacitracin: complexation with metal ion and C 55 -isoprenyl pyrophosphate., Proc. Natl. Acad. Sci. U. S. A. 68 (1971) 3223–3227. https://doi.org/10.1073/pnas.68.12.3223.
[20] J.M. Munita, C.A. Arias, Mechanisms of Antibiotic Resistance, Microbiol. Spectr. 4 (2016) 14207–12. https://doi.org/10.1128/microbiolspec.VMBF-0016-2015.
[21] J.M. Munita, C.A. Arias, Mechanisms of Antibiotic Resistance, Microbiol. Spectr. 4 (2016) 1–24. https://doi.org/10.1128/microbiolspec.VMBF-0016-2015.
[22] L. Röse, S.H.E. Kaufmann, S. Daugelat, Involvement of Mycobacterium smegmatis undecaprenyl phosphokinase in biofilm and smegma formation., Microbes Infect. 6 (2004) 965–71. https://doi.org/10.1016/j.micinf.2004.05.011.
[23] A.F. Chalker, K. a Ingraham, R.D. Lunsford, A.P. Bryant, J. Bryant, N.G. Wallis, J.P. Broskey, S.C. Pearson, D.J. Holmes, The bacA gene, which determines bacitracin susceptibility in Streptococcus pneumoniae and Staphylococcus aureus, is also required for virulence, Microbiology. 146 (2000) 1547–1553. http://mic.sgmjournals.org/content/146/7/1547.abstract.
[24] N. Jalal, X.L. Tian, G. Dong, J. Upham, C. Chen, M. Parcells, Y.H. Li, Identification and characterization of SMU.244 encoding a putative undecaprenyl pyrophosphate phosphatase protein required for cell wall biosynthesis and bacitracin resistance in Streptococcus mutans, Microbiol. (United Kingdom). 161 (2015) 1857–1870. https://doi.org/10.1099/mic.0.000142.
[25] J.D. Ochocki, M.D. Distefano, Prenyltransferase inhibitors: treating human ailments from cancer to parasitic infections, Med. Chem. Commun. 4 (2013) 476–492. https://doi.org/10.1039/C2MD20299A.
[26] R.T. Guo, R. Cao, P.H. Liang, T.P. Ko, T.H. Chang, M.P. Hudock, W.Y. Jeng, C.K.M. Chen, Y. Zhang, Y. Song, C.J. Kuo, F. Yin, E. Oldfield, A.H.J. Wang, Bisphosphonates target multiple sites in both cis- and trans- prenyltransferases, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 10022–10027. https://doi.org/10.1073/pnas.0702254104.
[27] J. Desai, Y. Wang, K. Wang, S.R. Malwal, E. Oldfield, Isoprenoid Biosynthesis Inhibitors Targeting Bacterial Cell Growth, ChemMedChem. (2016) 2205–2215. https://doi.org/10.1002/cmdc.201600343.
[28] H. Mohammad, W. Younis, L. Chen, C.E. Peters, J. Pogliano, K. Pogliano, B. Cooper, J. Zhang, A. Mayhoub, E. Oldfield, M. Cushman, M.N. Seleem, Phenylthiazole Antibacterial Agents Targeting Cell Wall Synthesis Exhibit Potent Activity in Vitro and in Vivo against Vancomycin-Resistant Enterococci, J. Med. Chem. 60 (2017) 2425–2438. https://doi.org/10.1021/acs.jmedchem.6b01780.

[29] M.M. Elsebaei, H. Mohammad, M. Abouf, N.S. Abutaleb, Y.A. Hegazy, A. Ghiaty, L. Chen, J. Zhang, S.R. Malwal, E. Oldfield, M.N. Seleem, A.S. Mayhoub, Alkynyl-containing phenylthiazoles: Systemically active antibacterial agents effective against methicillin-resistant Staphylococcus aureus (MRSA), Eur. J. Med. Chem. 148 (2018) 195–209. https://doi.org/10.1016/j.ejmech.2018.02.031.
[30] M. Jukic, K. Rožman, M. Sova, H. Barreteau, S. Gobec, Anthranilic acid inhibitors of Undecaprenyl Pyrophosphate Synthase (UppS), an essential enzyme for bacterial cell wall biosynthesis, Front. Microbiol. 10 (2019) 1–9. https://doi.org/10.3389/fmicb.2018.03322.
[31] Y. Zhang, F.Y. Lin, K. Li, W. Zhu, Y.L. Liu, R. Cao, R. Pang, E. Lee, J. Axelson, M. Hensler, K. Wang, K.J. Molohon, Y. Wang, D.A. Mitchell, V. Nizet, E. Oldfield, HIV-1 integrase inhibitor-inspired antibacterials targeting isoprenoid biosynthesis, ACS Med. Chem. Lett. 3 (2012) 402–406. https://doi.org/10.1021/ml300038t.
[32] W. Zhu, Y. Zhang, W. Sinko, M.E. Hensler, J. Olson, K.J. Molohon, S. Lindert, R. Cao, K. Li, K. Wang, Y. Wang, Y.-L. Liu, A. Sankovsky, C.A.F. de Oliveira, D.A. Mitchell, V. Nizet, J.A. McCammon, E. Oldfield, Antibacterial drug leads targeting isoprenoid biosynthesis, Proc. Natl. Acad. Sci. 110 (2013) 123–128. https://doi.org/10.1073/pnas.1219899110.
[33] W. Zhu, Y. Wang, K. Li, J. Gao, C.H. Huang, C.C. Chen, T.P. Ko, Y. Zhang, R.T. Guo, E. Oldfield, Antibacterial drug leads: DNA and enzyme multitargeting, J. Med. Chem. 58 (2015) 1215–1227. https://doi.org/10.1021/jm501449u.
[34] D.E. Danley, E.T. Baima, M. Mansour, K.F. Fennell, B.A. Chrunyk, J.P. Mueller, S. Liu, X. Qiu, Discovery and structural characterization of an allosteric inhibitor of bacterial cis-prenyltransferase, Protein Sci. 24 (2015) 20–26. https://doi.org/10.1002/pro.2579.
[35] M.A. Farha, T.L. Czarny, C.L. Myers, L.J. Worrall, S. French, D.G. Conrady, Y. Wang, E. Oldfield, N.C.J. Strynadka, E.D. Brown, Antagonism screen for inhibitors of bacterial cell wall biogenesis uncovers an inhibitor of undecaprenyl diphosphate synthase, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 11048–11053. https://doi.org/10.1073/pnas.1511751112.
[36] Y. Wang, J. Desai, Y. Zhang, S.R. Malwal, C.J. Shin, X. Feng, H. Sun, G. Liu, R.T. Guo, E. Oldfield, Bacterial Cell Growth Inhibitors Targeting Undecaprenyl Diphosphate Synthase and Undecaprenyl Diphosphate Phosphatase, ChemMedChem. 11 (2016) 2311–2319. https://doi.org/10.1002/cmdc.201600342.
[37] T.L. Czarny, E.D. Brown, A Small-Molecule Screening Platform for the Discovery of Inhibitors of Undecaprenyl Diphosphate Synthase, ACS Infect. Dis. 2 (2016) 489–499. https://doi.org/10.1021/acsinfecdis.6b00044.
[38] N. Concha, J. Huang, X. Bai, A. Benowitz, P. Brady, L.C. Grady, L.H. Kryn, D. Holmes, K. Ingraham, Q. Jin, L. Pothier Kaushansky, L. McCloskey, J.A. Messer, H. O'Keefe, A. Patel, A.L. Satz, R.H. Sinnamon, J. Schneck, S.R. Skinner, J. Summerfield, A. Taylor, J.D. Taylor, G. Evindar, R.A. Stavenger, Discovery and Characterization of a Class of Pyrazole Inhibitors of Bacterial Undecaprenyl Pyrophosphate Synthase, J. Med. Chem. 59 (2016) 7299–7304. https://doi.org/10.1021/acs.jmedchem.6b00746.
[39] M.M. Elsebaei, H. Mohammad, A. Samir, N.S. Abutaleb, A.B. Norvil, A.R. Michie, M.M. Moustafa, H. Samy, H. Gowher, M.N. Seleem, A.S. Mayhoub, Lipophilic efficient phenylthiazoles with potent undecaprenyl pyrophosphatase inhibitory activity, Eur. J. Med. Chem. 175 (2019) 49–62. https://doi.org/10.1016/j.ejmech.2019.04.063.
[40] L. Surmacz, E. Swiezewska, Polyisoprenoids - Secondary metabolites or physiologically

important superlipids?, Biochem. Biophys. Res. Commun. 407 (2011) 627–632. https://doi.org/10.1016/j.bbrc.2011.03.059.
[41] M.D. Hartley, B. Imperiali, At the membrane frontier: A prospectus on the remarkable evolutionary conservation of polyprenols and polyprenyl-phosphates, Arch. Biochem. Biophys. 517 (2012) 83–97. https://doi.org/10.1016/j.abb.2011.10.018.
[42] B.G. Ng, H.H. Freeze, Perspectives on Glycosylation and Its Congenital Disorders, Trends Genet. 34 (2018) 466–476. https://doi.org/10.1016/j.tig.2018.03.002.
[43] M. Welti, Regulation of dolichol-linked glycosylation, Glycoconj. J. 30 (2013) 51–56. https://doi.org/10.1007/s10719-012-9417-y.
[44] A. Bouhss, A.E. Trunkfield, T.D.H. Bugg, D. Mengin-Lecreulx, The biosynthesis of peptidoglycan lipid-linked intermediates., FEMS Microbiol. Rev. 32 (2008) 208–33. https://doi.org/10.1111/j.1574-6976.2007.00089.x.
[45] S. Kalynych, R. Morona, M. Cygler, Progress in understanding the assembly process of bacterial O-antigen, FEMS Microbiol. Rev. 38 (2014) 1048–1065. https://doi.org/10.1111/1574-6976.12070.
[46] D. Kaur, P.J. Brennan, D.C. Crick, Decaprenyl diphosphate synthesis in Mycobacterium tuberculosis, J. Bacteriol. 186 (2004) 7564–7570. https://doi.org/10.1128/JB.186.22.7564-7570.2004.
[47] L.T. Sham, E.K. Butler, M.D. Lebar, D. Kahne, T.G. Bernhardt, N. Ruiz, MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis, Science (80-.). 345 (2014) 220–222. https://doi.org/10.1126/science.1254522.
[48] S. Mulholland, E.R. Turpin, B.B. Bonev, J.D. Hirst, Docking and molecular dynamics simulations of the ternary complex nisin2:lipid II, Sci. Rep. 6 (2016) 1–11. https://doi.org/10.1038/srep21185.
[49] C. Valtersson, G. van Duyn, A.J. Verkleij, T. Chojnacki, B. de Kruijff, G. Dallner, The influence of dolichol, dolichol esters, and dolichyl phosphate on phospholipid polymorphism and fluidity in model membranes, J. Biol. Chem. 260 (1985) 2742–2751.
[50] G. ban Duijn, C. Valtersson, T. Chojnacki, A.J. Verkleij, G. Dallner, B. de Kruijff, Dolichyl phosphate induces non-bilayer structures, vesicle fusion and transbilayer movement of lipids: a model membrane study, BBA - Biomembr. 861 (1986) 211–223. https://doi.org/10.1016/0005-2736(86)90423-2.
[51] X. Wang, A.R. Mansourian, P.J. Quinn, The effect of dolichol on the structure and phase behaviour of phospholipid model membranes., Mol. Membr. Biol. 25 (2008) 547–556. https://doi.org/10.1080/09687680802520684.
[52] M. El Ghachi, A. Bouhss, D. Blanot, D. Mengin-Lecreulx, The bacA gene of Escherichia coli encodes an undecaprenyl pyrophosphate phosphatase activity., J. Biol. Chem. 279 (2004) 30106–13. https://doi.org/10.1074/jbc.M401701200.
[53] M. El Ghachi, A. Derbise, A. Bouhss, D. Mengin-Lecreulx, Identification of multiple genes encoding membrane proteins with undecaprenyl pyrophosphate phosphatase (UppP) activity in Escherichia coli., J. Biol. Chem. 280 (2005) 18689–95. https://doi.org/10.1074/jbc.M412277200.
[54] L.D. Tatar, C.L. Marolda, A.N. Polischuk, D. van Leeuwen, M. a. Valvano, An Escherichia coli undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling, Microbiology. 153 (2007) 2518–2529. https://doi.org/10.1099/mic.0.2007/006312-0.
[55] M.A. McCloskey, F.A. Troy, Paramagnetic Isoprenoid Carrier Lipids. 1. Chemical

Synthesis and Incorporation into Model Membranes, Biochemistry. 19 (1980) 2056–2060. https://doi.org/10.1021/bi00551a008.
[56] M.A. McCloskey, F.A. Troy, Paramagnetic Isoprenoid Carrier Lipids. 2. Dispersion and Dynamics in Lipid Membranes, Biochemistry. 19 (1980) 2061–2066. https://doi.org/10.1021/bi00551a009.
[57] D. Mengin-Lecreulx, J. Van Heijenoort, Effect of growth conditions on peptidoglycan content and cytoplasmic steps of its biosynthesis in Escherichia coli, J. Bacteriol. 163 (1985) 208–212.
[58] S. Heuston, M. Begley, C.G.M. Gahan, C. Hill, Isoprenoid biosynthesis in bacterial pathogens, Microbiol. (United Kingdom). 158 (2012) 1389–1401. https://doi.org/10.1099/mic.0.051599-0.
[59] N. Ladygina, E.G. Dedyukhina, M.B. Vainshtein, A review on microbial synthesis of hydrocarbons, Process Biochem. 41 (2006) 1001–1014. https://doi.org/10.1016/j.procbio.2005.12.007.
[60] T. TouzÉ, D. Mengin-Lecreulx, Undecaprenyl Phosphate Synthesis, EcoSal Plus. 1 (2013) 1–20. https://doi.org/10.1128/ecosalplus.4.7.1.7.
[61] Z.A. Demissie, L.A.E. Erland, M.R. Rheault, S.S. Mahmoud, The biosynthetic origin of irregular monoterpenes in lavandula: Isolation and biochemical characterization of a novel cis-prenyl diphosphate synthase gene, lavandulyl diphosphate synthase, J. Biol. Chem. 288 (2013) 6333–6341. https://doi.org/10.1074/jbc.M112.431171.
[62] M. Liu, C.C. Chen, L. Chen, X. Xiao, Y. Zheng, J.W. Huang, W. Liu, T.P. Ko, Y.S. Cheng, X. Feng, E. Oldfield, R.T. Guo, Y. Ma, Structure and Function of a “head-to-Middle” Prenyltransferase: Lavandulyl Diphosphate Synthase, Angew. Chemie - Int. Ed. 55 (2016) 4721–4724. https://doi.org/10.1002/anie.201600656.
[63] J. Gao, T.P. Ko, L. Chen, S.R. Malwal, J. Zhang, X. Hu, F. Qu, W. Liu, J.W. Huang, Y.S. Cheng, C.C. Chen, Y. Yang, Y. Zhang, E. Oldfield, R.T. Guo, “Head-to-Middle” and “Head-to-Tail” cis-Prenyl Transferases: Structure of Isosesquilavandulyl Diphosphate Synthase, Angew. Chemie - Int. Ed. 57 (2018) 683–687. https://doi.org/10.1002/anie.201710185.
[64] J.S. Anderson, M. Matsuhashi, M.A. Haskin, J.L. Strominger, Lipid-Phosphoacetylmuramyl-Pentapeptide and Lipid-Phosphodisaccharide-Pentapeptide: Presumed Membrane Transport Intermediates in Cell Wall Synthesis, Proc. Natl. Acad. Sci. 53 (1965) 881–889. https://doi.org/10.1073/pnas.53.4.881.
[65] A. Wright, M. Dankert, P.W. Robbins, Evidence for an intermediate stage in the biosynthesis of the Salmonella O-antigen., Proc. Natl. Acad. Sci. U. S. A. 54 (1965) 235–241. https://doi.org/10.1073/pnas.54.1.235.
[66] Y. Higashi, J.L. Strominger, C.C. Sweeley, Structure of a lipid intermediate in cell wall peptidoglycan synthesis: a derivative of a C55 isoprenoid alcohol., Proc. Natl. Acad. Sci. U. S. A. 57 (1967) 1878–84. https://doi.org/10.1073/pnas.57.6.1878.
[67] A. Wright, M. Dankert, P. Fennessey, P.W. Robbins, Characterization of a polyisoprenoid compound functional in O-antigen biosynthesis., Proc. Natl. Acad. Sci. U. S. A. 57 (1967) 1798–1803. https://doi.org/10.1073/pnas.57.6.1798.
[68] M. Ito, M. Kobayashi, T. Koyama, K. Ogura, Stereochemical Analysis of Prenyltransferase Reactions Leading to (Z)- and (E)-Polyprenyl Chains, Biochemistry. 26 (1987) 4745–4750. https://doi.org/10.1021/bi00389a022.
[69] M. V. Keenan, C.M. Allen, Characterization of undecaprenyl pyrophosphate synthetase

from Lactobacillus plantarum, Arch. Biochem. Biophys. 161 (1974) 375–383. https://doi.org/10.1016/0003-9861(74)90318-X.
[70] C.M. Allen, M. V. Keenan, J. Sack, Lactobacillus plantarum undecaprenyl pyrophosphate synthetase: Purification and reaction requirements, Arch. Biochem. Biophys. 175 (1976) 236–248. https://doi.org/10.1016/0003-9861(76)90504-x.
[71] C.M. Apfel, B. Takács, M. Fountoulakis, M. Stieger, W. Keck, Use of genomics to identify bacterial undecaprenyl pyrophosphate synthetase: Cloning, expression, and characterization of the essential uppS gene, J. Bacteriol. 181 (1999) 483–492.
[72] T. Koyama, I. Yoshida, K. Ogura, Undecaprenyl diphosphate synthase from Micrococcus luteus B-P 26: Essential factors for the enzymatic activity, J. Biochem. 103 (1988) 867–871. https://doi.org/10.1093/oxfordjournals.jbchem.a122363.
[73] Y.P. Lu, H.G. Liu, K.H. Teng, P.H. Liang, Mechanism of cis-prenyltransferase reaction probed by substrate analogues, Biochem. Biophys. Res. Commun. 400 (2010) 758–762. https://doi.org/10.1016/j.bbrc.2010.09.001.
[74] T.P. Ko, Y.K. Chen, H. Robinson, P.C. Tsai, Y.G. Gao, A.P.C. Chen, A.H.J. Wang, P.H. Liang, Mechanism of Product Chain Length Determination and the Role of a Flexible Loop in Escherichia coli Undecaprenyl-pyrophosphate Synthase Catalysis, J. Biol. Chem. 276 (2001) 47474–47482. https://doi.org/10.1074/jbc.M106747200.
[75] J.J. Pan, L.W. Yang, P.H. Liang, Effect of site-directed mutagenesis of the conserved aspartate and glutamate on E. coli undecaprenyl pyrophosphate synthase catalysis, Biochemistry. 39 (2000) 13856–13861. https://doi.org/10.1021/bi001226h.
[76] S.-Y. Chang, Substrate binding mode and reaction mechanism of undecaprenyl pyrophosphate synthase deduced from crystallographic studies, Protein Sci. 13 (2004) 971–978. https://doi.org/10.1110/ps.03519904.
[77] R.T. Guo, T.P. Ko, A.P.C. Chen, C.J. Kuo, A.H.J. Wang, P.H. Liang, Crystal structures of undecaprenyl pyrophosphate synthase in complex with magnesium, isopentenyl pyrophosphate, and farnesyl thiopyrophosphate: Roles of the metal ion and conserved residues in catalysis, J. Biol. Chem. 280 (2005) 20762–20774. https://doi.org/10.1074/jbc.M502121200.
[78] P.H. Liang, T.P. Ko, A.H.J. Wang, Structure, mechanism and function of prenyltransferases, Eur. J. Biochem. 269 (2002) 3339–3354. https://doi.org/10.1046/j.1432-1033.2002.03014.x.
[79] Y.P. Lu, H.G. Liu, P.H. Liang, Different reaction mechanisms for cis- and trans-prenyltransferases, Biochem. Biophys. Res. Commun. 379 (2009) 351–355. https://doi.org/10.1016/j.bbrc.2008.12.061.
[80] R.T. Guo, C.J. Kuo, C.C. Chou, T.P. Ko, H.L. Shr, P.H. Liang, A.H.J. Wang, Crystal Structure of Octaprenyl Pyrophosphate Synthase from Hyperthermophilic Thermotoga maritima and Mechanism of Product Chain Length Determination, J. Biol. Chem. 279 (2004) 4903–4912. https://doi.org/10.1074/jbc.M310161200.
[81] Y.H. Chen, A.P.C. Chen, C.T. Chen, A.H.J. Wang, P.H. Liang, Probing the conformational change of Escherichia coli undecaprenyl pyrophosphate synthase during catalysis using an inhibitor and tryptophan mutants, J. Biol. Chem. 277 (2002) 7369–7376. https://doi.org/10.1074/jbc.M110014200.
[82] J. Inokoshi, Y. Nakamura, S. Komada, K. Komatsu, H. Umeyama, H. Tomoda, Inhibition of bacterial undecaprenyl pyrophosphate synthase by small fungal molecules, J. Antibiot. (Tokyo). 69 (2016) 798–805. https://doi.org/10.1038/ja.2016.35.

[83] H.C. Chan, X. Feng, T.P. Ko, C.H. Huang, Y. Hu, Y. Zheng, S. Bogue, C. Nakano, T. Hoshino, L. Zhang, P. Lv, W. Liu, D.C. Crick, P.H. Liang, A.H.J. Wang, E. Oldfield, R.T. Guo, Structure and inhibition of tuberculosinol synthase and decaprenyl diphosphate synthase from mycobacterium tuberculosis, J. Am. Chem. Soc. 136 (2014) 2892–2896. https://doi.org/10.1021/ja413127v.
[84] S.R. Malwal, L. Chen, H. Hicks, F. Qu, W. Liu, A. Shillo, W.X. Law, J. Zhang, N. Chandnani, X. Han, Y. Zheng, C.C. Chen, R.T. Guo, A. Abdelkhalek, M.N. Seleem, E. Oldfield, Discovery of Lipophilic Bisphosphonates That Target Bacterial Cell Wall and Quinone Biosynthesis, J. Med. Chem. 62 (2019) 2564–2581. https://doi.org/10.1021/acs.jmedchem.8b01878.
[85] L. Rothfield, D. Romeo, Role of lipids in the biosynthesis of the bacterial cell envelope., Bacteriol. Rev. 35 (1971) 14–38.
[86] R. Goldman, J.L. Strominger, Purification and properties of C 55 -isoprenylpyrophosphate phosphatase from Micrococcus lysodeikticus., J. Biol. Chem. 247 (1972) 5116–5122.
[87] G. Siewert, J.L. Strominger, Bacitracin: an inhibitor of the dephosphorylation of lipid pyrophosphate, an intermediate in the biosynthesis of the peptidoglycan of bacterial cell walls., Proc. Natl. Acad. Sci. 57 (1967) 767–773. https://doi.org/10.1073/pnas.57.3.767.
[88] B.D. Cain, P.J. Norton, W. Eubanks, H.S. Nick, C.M. Allen, Amplification of the bacA gene confers bacitracin resistance to Escherichia coli, J. Bacteriol. 175 (1993) 3784–3789. http://jb.asm.org/content/175/12/3784.abstract.
[89] J.K. Kim, H.J. Lee, Y. Kikuchi, W. Kitagawa, N. Nikoh, T. Fukatsu, B.L. Lee, Bacterial cell wall synthesis gene uppP is required for burkholderia colonization of the stinkbug gut, Appl. Environ. Microbiol. 79 (2013) 4879–4886. https://doi.org/10.1128/AEM.01269-13.
[90] W. MEYER, G. SCHÄFER, Characterization and purification of a membrane-bound archaebacterial pyrophosphatase from Sulfolobus acidocaldarius, Eur. J. Biochem. 207 (1992) 741–746. https://doi.org/10.1111/j.1432-1033.1992.tb17104.x.
[91] 28852–7 (2005). Bernard, Remi1. Bernard, R., El Ghachi, M., Mengin-Lecreulx, D., Chippaux, M. & Denizot, F. BcrC from Bacillus subtilis acts as an undecaprenyl pyrophosphate phosphatase in bacitracin resistance. J. Biol. Chem. 280, M. El Ghachi, D. Mengin-Lecreulx, M. Chippaux, F. Denizot, BcrC from Bacillus subtilis acts as an undecaprenyl pyrophosphate phosphatase in bacitracin resistance., J. Biol. Chem. 280 (2005) 28852–7. https://doi.org/10.1074/jbc.M413750200.
[92] J. Stukey, G.M. Carman, Identification of a novel phosphatase sequence motif, Protein Sci. 6 (2008) 469–472. https://doi.org/10.1002/pro.5560060226.
[93] A. Messerschmidt, R. Wever, X-ray structure of a vanadium-containing enzyme: Chloroperoxidase from the fungus Curvularia inaequalis, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 392–396. https://doi.org/10.1073/pnas.93.1.392.
[94] A.F. Neuwald, An unexpected structural relationship between integral membrane phosphatases and soluble haloperoxidases, Protein Sci. 6 (1997) 1764–1767. https://doi.org/10.1002/pro.5560060817.
[95] K. Ishikawa, X-ray structures of a novel acid phosphatase from Escherichia blattae and its complex with the transition-state analog molybdate, EMBO J. 19 (2000) 2412–2423. https://doi.org/10.1093/emboj/19.11.2412.
[96] T. Touzé, A.X. Tran, J. V Hankins, D. Mengin-Lecreulx, M.S. Trent, Periplasmic phosphorylation of lipid A is linked to the synthesis of undecaprenyl phosphate., Mol. Microbiol. 67 (2008) 264–77. https://doi.org/10.1111/j.1365-2958.2007.06044.x.

[97] T. Touzé, D. Blanot, D. Mengin-Lecreulx, Substrate specificity and membrane topology of Escherichia coli PgpB, an undecaprenyl pyrophosphate phosphatase., J. Biol. Chem. 283 (2008) 16573–83. https://doi.org/10.1074/jbc.M800394200.
[98] D.A. Dillon, W.I. Wut, B. Riedel, J.B. Wissing, W. Dowhan, G.M. Carman, The Escherichia coli pgpB gene encodes for a diacylglycerol pyrophosphate phosphatase activity, J. Biol. Chem. 271 (1996) 30548–30553. https://doi.org/10.1074/jbc.271.48.30548.
[99] J. Fan, D. Jiang, Y. Zhao, J. Liu, X.C. Zhang, Crystal structure of lipid phosphatase Escherichia coli phosphatidylglycerophosphate phosphatase B, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 7636–7640. https://doi.org/10.1073/pnas.1403097111.
[100] S. Tong, Y. Lin, S. Lu, M. Wang, M. Bogdanov, L. Zheng, Structural insight into substrate selection and catalysis of lipid phosphate phosphatase PgpB in the cell membrane, J. Biol. Chem. 291 (2016) 18342–18352. https://doi.org/10.1074/jbc.M116.737874.
[101] M. El Ghachi, N. Howe, R. Auger, A. Lambion, A. Guiseppi, F. Delbrassine, G. Manat, S. Roure, S. Peslier, E. Sauvage, L. Vogeley, J.C. Rengifo-Gonzalez, P. Charlier, D. Mengin-Lecreulx, M. Foglino, T. Touzé, M. Caffrey, F. Kerff, Crystal structure and biochemical characterization of the transmembrane PAP2 type phosphatidylglycerol phosphate phosphatase from Bacillus subtilis, Cell. Mol. Life Sci. 74 (2017) 2319–2332. https://doi.org/10.1007/s00018-017-2464-6.
[102] J.B. Vincent, M.W. Crowder, B.A. Averill, Hydrolysis of phosphate monoesters: a biological problem with multiple chemical solutions, Trends Biochem. Sci. 17 (1992) 105–110. https://doi.org/10.1016/0968-0004(92)90246-6.
[103] W.D. Van Horn, C.R. Sanders, Prokaryotic Diacylglycerol Kinase and Undecaprenol Kinase, Annu. Rev. Biophys. 41 (2012) 81–101. https://doi.org/10.1146/annurev-biophys-050511-102330.
[104] M. Lis, H.K. Kuramitsu, The stress-responsive dgk gene from Streptococcus mutans encodes a putative undecaprenol kinase activity, Infect. Immun. 71 (2003) 1938–1943. https://doi.org/10.1128/IAI.71.4.1938-1943.2003.
[105] L.Y. Huang, S.C. Wang, T.J.R. Cheng, C.H. Wong, Undecaprenyl Phosphate Phosphatase Activity of Undecaprenol Kinase Regulates the Lipid Pool in Gram-Positive Bacteria, Biochemistry. 56 (2017) 5417–5427. https://doi.org/10.1021/acs.biochem.7b00603.
[106] G. Manat, M. El Ghachi, R. Auger, K. Baouche, S. Olatunji, F. Kerff, T. Touzé, D. Mengin-Lecreulx, A. Bouhss, Membrane Topology and Biochemical Characterization of the Escherichia coli BacA Undecaprenyl-Pyrophosphate Phosphatase, PLoS One. 10 (2015) e0142870. https://doi.org/10.1371/journal.pone.0142870.
[107] E. Gasiorowski, R. Auger, X. Tian, S. Hicham, C. Ecobichon, S. Roure, M. V Douglass, M.S. Trent, I. Gomperts, B. Id, HupA , the main undecaprenyl pyrophosphate and phosphatidylglycerol phosphate phosphatase in Helicobacter pylori is essential for colonization of the stomach, (2019) 1–24.
[108] J. Radeck, N. Lautenschläger, T. Mascher, The essential UPP phosphatase pair BcrC and UppP connects cell wall homeostasis during growth and sporulation with cell envelope stress response in Bacillus subtilis, Front. Microbiol. 8 (2017) 1–14. https://doi.org/10.3389/fmicb.2017.02403.
[109] H.Y. Chang, C.C. Chou, M.F. Hsu, A.H.J. Wang, Proposed carrier lipid-binding site of undecaprenyl pyrophosphate phosphatase from escherichia coli, J. Biol. Chem. 289 (2014)

18719–18735. https://doi.org/10.1074/jbc.M114.575076.
[110] E. Wallin, G. Von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, Protein Sci. 7 (2008) 1029–1038. https://doi.org/10.1002/pro.5560070420.
[111] Y. Arinaminpathy, E. Khurana, D.M. Engelman, M.B. Gerstein, Computational analysis of membrane proteins: the largest class of drug targets, Drug Discov. Today. 14 (2009) 1130–1135. https://doi.org/10.1016/j.drudis.2009.08.006.
[112] J. Deisenhofer, O. Epp, K. Miki, R. Huber, H. Michel, Structure of the protein subunits in the photosynthetic reaction center of, Nature. (1985). https://www.nature.com/articles/318618a0.pdf%0Ahttp://www4.utsouthwestern.edu/MichaelyLab/Nature318-618.pdf.
[113] S.J. Opella, F.M. Marassi, Applications of NMR to membrane proteins, Arch. Biochem. Biophys. 628 (2017) 92–101. https://doi.org/10.1016/j.abb.2017.05.011.
[114] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, M. Miyano, Crystal structure of rhodopsin: A G protein-coupled receptor, Science (80-.). 289 (2000) 739–745. https://doi.org/10.1126/science.289.5480.739.
[115] S. Wagner, L. Baarst, A.J. Ytterberg, A. Klussmerer, C.S. Wagner, O. Nord, P.Å. Nygren, K.J. Van Wijks, J.W. De Gier, Consequences of membrane protein overexpression in Escherichia coli, Mol. Cell. Proteomics. 6 (2007) 1527–1550. https://doi.org/10.1074/mcp.M600431-MCP200.
[116] H.E. Autzen, D. Julius, Y. Cheng, Membrane mimetic systems in CryoEM: keeping membrane proteins in their native environment, Curr. Opin. Struct. Biol. 58 (2019) 259–268. https://doi.org/10.1016/j.sbi.2019.05.022.
[117] M.L. Carlson, J.W. Young, Z. Zhao, L. Fabre, D. Jun, J. Li, J. Li, H.S. Dhupar, I. Wason, A.T. Mills, J.T. Beatty, J.S. Klassen, I. Rouiller, F. Duong, The peptidisc, a simple method for stabilizing membrane proteins in detergent-free solution, Elife. 7 (2018) 1–23. https://doi.org/10.7554/eLife.34085.
[118] A. Ishchenko, E.E. Abola, V. Cherezov, Crystallization of Membrane Proteins: An Overview, in: Protein Crystallogr. Methods Protoc. Methods Mol. Biol., 2017: pp. 117–141. https://doi.org/10.1007/978-1-4939-7000-1_5.
[119] J.M. Dörr, S. Scheidelaar, M.C. Koorengevel, J.J. Dominguez, M. Schäfer, C.A. van Walree, J.A. Killian, The styrene–maleic acid copolymer: a versatile tool in membrane research, Eur. Biophys. J. 45 (2016) 3–21. https://doi.org/10.1007/s00249-015-1093-y.
[120] Z. Stroud, S.C.L. Hall, T.R. Dafforn, Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights, Methods. 147 (2018) 106–117. https://doi.org/10.1016/j.ymeth.2018.03.011.
[121] T. Kawate, E. Gouaux, Fluorescence-Detection Size-Exclusion Chromatography for Precrystallization Screening of Integral Membrane Proteins, Structure. 14 (2006) 673–681. https://doi.org/10.1016/j.str.2006.01.013.
[122] Y. Cheng, Membrane protein structural biology in the era of single particle cryo-EM, Curr. Opin. Struct. Biol. 52 (2018) 58–63. https://doi.org/10.1016/j.sbi.2018.08.008.
[123] S.S. Pang, C. Bayly-Jones, M. Radjainia, B.A. Spicer, R.H.P. Law, A.W. Hodel, E.S. Parsons, S.M. Ekkel, P.J. Conroy, G. Ramm, H. Venugopal, P.I. Bird, B.W. Hoogenboom, I. Voskoboinik, Y. Gambin, E. Sierecki, M.A. Dunstone, J.C. Whisstock, The cryo-EM structure of the acid activatable pore-forming immune effector Macrophage-expressed

gene 1, Nat. Commun. 10 (2019) 1–9. https://doi.org/10.1038/s41467-019-12279-2.
[124] B. Hu, M. Lara-Tejero, Q. Kong, J.E. Galán, J. Liu, In Situ Molecular Architecture of the Salmonella Type III Secretion Machine, Cell. 168 (2017) 1065-1074.e10. https://doi.org/10.1016/j.cell.2017.02.022.
[125] C. Butan, M. Lara-Tejero, W. Li, J. Liu, J.E. Galán, High-resolution view of the type III secretion export apparatus in situ reveals membrane remodeling and a secretion pathway., Proc. Natl. Acad. Sci. U. S. A. 116 (2019). https://doi.org/10.1073/pnas.1916331116.
[126] Y. Liu, D.T. Huynh, T.O. Yeates, A 3.8 Å resolution cryo-EM structure of a small protein bound to an imaging scaffold, Nat. Commun. 10 (2019) 1–7. https://doi.org/10.1038/s41467-019-09836-0.
[127] X. Fan, J. Wang, X. Zhang, Z. Yang, J.C. Zhang, L. Zhao, H.L. Peng, J. Lei, H.W. Wang, Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution, Nat. Commun. 10 (2019) 1–11. https://doi.org/10.1038/s41467-019-10368-w.
[128] J.R. Luft, J. Newman, E.H. Snell, Crystallization screening: The influence of history on current practice, Acta Crystallogr. Sect. FStructural Biol. Commun. 70 (2014) 835–853. https://doi.org/10.1107/S2053230X1401262X.
[129] P.J. Loll, Membrane proteins, detergents and crystals: What is the state of the art?, Acta Crystallogr. Sect. FStructural Biol. Commun. 70 (2014) 1576–1583. https://doi.org/10.1107/S2053230X14025035.
[130] E.M. Landau, J.P. Rosenbusch, Lipidic cubic phases: a novel concept for the crystallization of membrane proteins., Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 14532–14535. https://doi.org/10.1073/pnas.93.25.14532.
[131] S. Faham, J.U. Bowie, Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure., J. Mol. Biol. 316 (2002) 1–6. https://doi.org/10.1006/jmbi.2001.5295.
[132] B. Miroux, J.E. Walker, Over-production of Proteins in Escherichia coli: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels, J. Mol. Biol. 260 (1996) 289–298. https://doi.org/10.1006/jmbi.1996.0399.
[133] S. Schlegel, P. Genevaux, J.W. de Gier, De-convoluting the Genetic Adaptations of E.coli C41(DE3) in Real Time Reveals How Alleviating Protein Production Stress Improves Yields, Cell Rep. 10 (2015) 1758–1766. https://doi.org/10.1016/j.celrep.2015.02.029.
[134] K. Kanonenberg, J. Royes, A. Kedrov, G. Poschmann, F. Angius, A. Solgadi, O. Spitz, D. Kleinschrodt, K. Stühler, B. Miroux, L. Schmitt, Shaping the lipid composition of bacterial membranes for membrane protein production, Microb. Cell Fact. 18 (2019). https://doi.org/10.1186/s12934-019-1182-1.
[135] M.S. King, C. Boes, E.R.S. Kunji, Membrane protein expression in Lactococcus lactis, 1st ed., Elsevier Inc., 2015. https://doi.org/10.1016/bs.mie.2014.12.009.
[136] T.P. Roosild, J. Greenwald, M. Vega, S. Castronovo, R. Riek, S. Choe, NMR structure of mistic, a membrane-integrating protein for membrane protein expression, Science (80-.). 307 (2005) 1317–1321. https://doi.org/10.1126/science.1106392.
[137] G.G. Privé, Detergents for the stabilization and crystallization of membrane proteins, Methods. 41 (2007) 388–397. https://doi.org/10.1016/j.ymeth.2007.01.007.
[138] P.S. Chae, S.G.F. Rasmussen, R.R. Rana, K. Gotfryd, R. Chandra, M.A. Goren, A.C. Kruse, S. Nurva, C.J. Loland, Y. Pierre, D. Drew, J.-L. Popot, D. Picot, B.G. Fox, L. Guan, U. Gether, B. Byrne, B. Kobilka, S.H. Gellman, Maltose–neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins, Nat.

Methods. 7 (2010) 1003–1008. https://doi.org/10.1038/nmeth.1526.
[139] P.S. Chae, S.G.F. Rasmussen, R.R. Rana, K. Gotfryd, A.C. Kruse, A. Manglik, K.H. Cho, S. Nurva, U. Gether, L. Guan, C.J. Loland, B. Byrne, B.K. Kobilka, S.H. Gellman, A New Class of Amphiphiles Bearing Rigid Hydrophobic Groups for Solubilization and Stabilization of Membrane Proteins, Chem. - A Eur. J. 18 (2012) 9485–9490. https://doi.org/10.1002/chem.201200069.
[140] M. Hattori, R.E. Hibbs, E. Gouaux, A fluorescence-detection size-exclusion chromatography-based thermostability assay to identify membrane protein expression and crystallization conditions, Structure. 20 (2012) 1293–1299. https://doi.org/10.1016/j.str.2012.06.009.A.
[141] P.S. Miller, A.R. Aricescu, Crystal structure of a human GABAA receptor, Nature. 512 (2014) 270–275. https://doi.org/10.1038/nature13293.
[142] J.A. Coleman, E.M. Green, E. Gouaux, X-ray structures and mechanism of the human serotonin transporter, Nature. 532 (2016) 334–339. https://doi.org/10.1038/nature17629.
[143] P.S. Miller, S. Scott, S. Masiulis, L. De Colibus, E. Pardon, J. Steyaert, A.R. Aricescu, Structural basis for GABA A receptor potentiation by neurosteroids, Nat. Struct. Mol. Biol. 24 (2017) 986–992. https://doi.org/10.1038/nsmb.3484.
[144] R.M. Walsh, S.H. Roh, A. Gharpure, C.L. Morales-Perez, J. Teng, R.E. Hibbs, Structural principles of distinct assemblies of the human α4β2 nicotinic receptor, Nature. 557 (2018) 261–265. https://doi.org/10.1038/s41586-018-0081-7.
[145] S. Phulera, H. Zhu, J. Yu, D.P. Claxton, N. Yoder, C. Yoshioka, E. Gouaux, Cryo-EM structure of the benzodiazepine-sensitive α1β1γ2S tri-heteromeric GABA A receptor in complex with GABA, Elife. 7 (2018) 1–21. https://doi.org/10.7554/eLife.39383.
[146] K. Morimoto, R. Suno, Y. Hotta, K. Yamashita, K. Hirata, M. Yamamoto, S. Narumiya, S. Iwata, T. Kobayashi, Crystal structure of the endogenous agonist-bound prostanoid receptor EP3, Nat. Chem. Biol. 15 (2019) 8–10. https://doi.org/10.1038/s41589-018-0171-8.
[147] H.E. Kato, Y. Zhang, H. Hu, C.M. Suomivuori, F.M.N. Kadji, J. Aoki, K. Krishna Kumar, R. Fonseca, D. Hilger, W. Huang, N.R. Latorraca, A. Inoue, R.O. Dror, B.K. Kobilka, G. Skiniotis, Conformational transitions of a neurotensin receptor 1–Gi1 complex, Nature. 572 (2019) 80–85. https://doi.org/10.1038/s41586-019-1337-6.
[148] Y. Zhao, S. Chen, E. Gouaux, Architecture and subunit arrangement of native AMPA receptors elucidated by cryo-EM., Science (80-.). 364 (2019) 355–362. https://doi.org/10.2210/pdb6njm/pdb.
[149] V. Cherezov, E. Yamashita, W. Liu, M. Zhalnina, W.A. Cramer, M. Caffrey, In Meso Structure of the Cobalamin Transporter, BtuB, at 1.95 Å Resolution, J. Mol. Biol. 364 (2006) 716–734. https://doi.org/10.1016/j.jmb.2006.09.022.
[150] D. Li, M. Caffrey, Lipid cubic phase as a membrane mimetic for integral membrane protein enzymes, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 8639–8644. https://doi.org/10.1073/pnas.1101815108.
[151] M. Caffrey, A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes, Acta Crystallogr. Sect. FStructural Biol. Commun. 71 (2015) 3–18. https://doi.org/10.1107/S2053230X14026843.
[152] M. Caffrey, Crystallizing Membrane Proteins for Structure Determination: Use of Lipidic Mesophases, Annu. Rev. Biophys. 38 (2009) 29–51.

https://doi.org/10.1146/annurev.biophys.050708.133655.
[153] N. Johner, S. Mondal, G. Morra, M. Caffrey, H. Weinstein, G. Khelashvili, Protein and lipid interactions driving molecular mechanisms of in meso crystallization, J. Am. Chem. Soc. 136 (2014) 3271–3284. https://doi.org/10.1021/ja4129839.
[154] A. Zabara, T.G. Meikle, R. Trenker, S. Yao, J. Newman, T.S. Peat, F. Separovic, C.E. Conn, M.J. Call, M.E. Call, E.M. Landau, C.J. Drummond, Lipidic Cubic Phase-Induced Membrane Protein Crystallization: Interplay between Lipid Molecular Structure, Mesophase Structure and Properties, and Crystallogenesis, Cryst. Growth Des. 17 (2017) 5667–5674. https://doi.org/10.1021/acs.cgd.7b00519.
[155] L. Van 't Hag, L. De Campo, N. Tran, A. Sokolova, R. Trenker, M.E. Call, M.J. Call, C.J. Garvey, A.E. Leung, T.A. Darwish, A. Krause-Heuer, R. Knott, T.G. Meikle, C.J. Drummond, R. Mezzenga, C.E. Conn, Protein-eye view of the in meso crystallization mechanism, Langmuir. 35 (2019) 8344–8356. https://doi.org/10.1021/acs.langmuir.9b00647.
[156] A. Zabara, J.T.Y. Chong, I. Martiel, L. Stark, B.A. Cromer, C. Speziale, C.J. Drummond, R. Mezzenga, Design of ultra-swollen lipidic mesophases for the crystallization of membrane proteins with large extracellular domains, Nat. Commun. 9 (2018). https://doi.org/10.1038/s41467-018-02996-5.
[157] L. V. Misquitta, Y. Misquitta, V. Cherezov, O. Slattery, J.M. Mohan, D. Hart, M. Zhalnina, W.A. Cramer, M. Caffrey, Membrane protein crystallization in lipidic mesophases with tailored bilayers, Structure. 12 (2004) 2113–2124. https://doi.org/10.1016/j.str.2004.09.020.
[158] C.Y. Huang, V. Olieric, P. Ma, E. Panepucci, K. Diederichs, M. Wang, M. Caffrey, In meso in situ serial X-ray crystallography of soluble and membrane proteins, Acta Crystallogr. Sect. D Biol. Crystallogr. 71 (2015) 1238–1256. https://doi.org/10.1107/S1399004715005210.
[159] C.Y. Huang, V. Olieric, P. Ma, N. Howe, L. Vogeley, X. Liu, R. Warshamanage, T. Weinert, E. Panepucci, B. Kobilka, K. Diederichs, M. Wang, M. Caffrey, In meso in situ serial X-ray crystallography of soluble and membrane proteins at cryogenic temperatures, Acta Crystallogr. Sect. D Struct. Biol. 72 (2016) 93–112. https://doi.org/10.1107/S2059798315021683.
[160] P. Nogly, V. Panneels, G. Nelson, C. Gati, T. Kimura, C. Milne, D. Milathianaki, M. Kubo, W. Wu, C. Conrad, J. Coe, R. Bean, Y. Zhao, P. Båth, R. Dods, R. Harimoorthy, K.R. Beyerlein, J. Rheinberger, D. James, D. DePonte, C. Li, L. Sala, G.J. Williams, M.S. Hunter, J.E. Koglin, P. Berntsen, E. Nango, S. Iwata, H.N. Chapman, P. Fromme, M. Frank, R. Abela, S. Boutet, A. Barty, T.A. White, U. Weierstall, J. Spence, R. Neutze, G. Schertler, J. Standfuss, Lipidic cubic phase injector is a viable crystal delivery system for time-resolved serial crystallography, Nat. Commun. 7 (2016) 1–9. https://doi.org/10.1038/ncomms12314.
[161] J. Standfuss, J. Spence, Serial crystallography at synchrotrons and X-ray lasers, IUCrJ. 4 (2017) 100–101. https://doi.org/10.1107/S2052252517001877.
[162] V. Lazarevic, D. Karamata, The tagGH operon of Bacillus subtilis 168 encodes a two-component ABC transporter involved in the metabolism of two wall teichoic acids, Mol. Microbiol. 16 (1995) 345–355. https://doi.org/10.1111/j.1365-2958.1995.tb02306.x.
[163] P. Sarkar, V. Yarlagadda, C. Ghosh, J. Haldar, A review on cell wall synthesis inhibitors with an emphasis on glycopeptide antibiotics, Medchemcomm. 8 (2017) 516–533.

https://doi.org/10.1039/c6md00585c.
[164] H.-Y. Chang, C.-C. Chou, M.-F. Hsu, A.H.J. Wang, Proposed carrier lipid-binding site of undecaprenyl pyrophosphate phosphatase from Escherichia coli., J. Biol. Chem. 289 (2014) 18719–35. https://doi.org/10.1074/jbc.M114.575076.
[165] C. Perez, S. Gerber, J. Boilevin, M. Bucher, T. Darbre, M. Aebi, J. Reymond, K.P. Locher, Structure and mechanism of an active lipid-linked oligosaccharide flippase, Nature. 524 (2015) 433–438. https://doi.org/10.1038/nature14953.
[166] A.C.Y. Kuk, E.H. Mashalidis, S.-Y. Lee, Crystal structure of the MOP flippase MurJ in an inward-facing conformation, Nat. Struct. Mol. Biol. 24 (2016) 171–176. https://doi.org/10.1038/nsmb.3346.
[167] F.W. Studier, Protein production by auto-induction in high-density shaking cultures, Protein Expr. Purif. 41 (2005) 207–234. https://doi.org/10.1016/j.pep.2005.01.016.
[168] W. Kabsch, Integration, scaling, space-group assignment and post-refinement, Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2010) 133–144. https://doi.org/10.1107/S0907444909047374.
[169] C. Vonrhein, E. Blanc, P. Roversi, G. Bricogne, Automated Structure Solution With autoSHARP, Macromol. Crystallogr. Protoc. Vol. 2. 364 (n.d.) 215–230. https://doi.org/10.1385/1-59745-266-1:215.
[170] G.M. Sheldrick, A short history of SHELX, Acta Crystallogr. Sect. A Found. Crystallogr. 64 (2007) 112–122. https://doi.org/10.1107/S0108767307043930.
[171] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2010) 486–501. https://doi.org/10.1107/S0907444910007493.
[172] P.D. Adams, P. V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2010) 213–221. https://doi.org/10.1107/S0907444909052925.
[173] G. Celniker, G. Nimrod, H. Ashkenazy, F. Glaser, E. Martz, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf: Using evolutionary data to raise testable hypotheses about protein function, Isr. J. Chem. 53 (2013) 199–206. https://doi.org/10.1002/ijch.201200096.
[174] N.R. Voss, M. Gerstein, 3V: cavity, channel and cleft volume calculator and extractor., Nucleic Acids Res. 38 (2010) W555-62. https://doi.org/10.1093/nar/gkq395.
[175] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612. https://doi.org/10.1002/jcc.20084.
[176] B.C. Chung, J. Zhao, R. Gillespie, D.Y. Kwon, Z. Guan, J. Hong, P. Zhou, S.-Y. Lee, Crystal Structure of MraY, an Essential Membrane Enzyme for Bacterial Cell Wall Synthesis, Biophys. J. 106 (2014) 14a. https://doi.org/10.1016/j.bpj.2013.11.131.
[177] D. Li, V.E. Pye, M. Caffrey, Experimental phasing for structure determination using membrane-protein crystals grown by the lipid cubic phase method, Acta Crystallogr. Sect. D Biol. Crystallogr. 71 (2015) 104–122. https://doi.org/10.1107/S1399004714010360.
[178] S. Ovchinnikov, L. Kinch, H. Park, Y. Liao, J. Pei, D.E. Kim, H. Kamisetty, N. V. Grishin, D. Baker, Large-scale determination of previously unsolved protein structures using evolutionary information, Elife. 4 (2015) 1–25. https://doi.org/10.7554/eLife.09248.

[179] I. Melnikov, V. Polovinkin, K. Kovalev, I. Gushchin, M. Shevtsov, V. Shevchenko, A. Mishin, A. Alekseev, F. Rodriguez-Valera, V. Borshchevskiy, V. Cherezov, G.A. Leonard, V. Gordeliy, A. Popov, Fast iodide-SAD phasing for high-throughput membrane protein structure determination, Sci. Adv. 3 (2017). https://doi.org/10.1126/sciadv.1602952.
[180] J.P. Abrahams, A.G.W. Leslie, Methods used in the structure determination of bovine mitochondrial F1 ATPase, Acta Crystallogr. Sect. D Biol. Crystallogr. 52 (1996) 30–42. https://doi.org/10.1107/S0907444995008754.
[181] K. Cowtan, The Buccaneer software for automated model building., Acta Crystallogr. Sect. D Biol. Crystallogr. 62 (2006) 1002–1011. https://doi.org/10.1107/S0907444906022116.
[182] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674. https://doi.org/10.1107/S0021889807021206.
[183] J.S. Bickford, H.S. Nick, Conservation of the PTEN catalytic motif in the bacterial undecaprenyl pyrophosphate phosphatase, BacA/UppP., Microbiology. 159 (2013) 2444–55. https://doi.org/10.1099/mic.0.070474-0.
[184] G. von Heijne, The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology., EMBO J. 5 (1986) 3021–3027. https://doi.org/10.1002/j.1460-2075.1986.tb04601.x.
[185] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state., J. Mol. Biol. 372 (2007) 774–97. https://doi.org/10.1016/j.jmb.2007.05.022.
[186] K. Gupta, J.A.C. Donlan, J.T.S. Hopper, P. Uzdavinys, M. Landreh, W.B. Struwe, D. Drew, A.J. Baldwin, P.J. Stansfeld, C. V. Robinson, The role of interfacial lipids in stabilizing membrane protein oligomers., Nature. 541 (2017) 421–424. https://doi.org/10.1038/nature20820.
[187] L.R. Forrest, Structural Symmetry in Membrane Proteins, Annu. Rev. Biophys. 44 (2015) 311–337. https://doi.org/10.1146/annurev-biophys-051013-023008.
[188] V.S. Reddy, M.A. Shlykov, R. Castillo, E.I. Sun, M.H. Saier, The major facilitator superfamily (MFS) revisited, FEBS J. 279 (2012) 2022–2035. https://doi.org/10.1111/j.1742-4658.2012.08588.x.
[189] N. Yan, Structural advances for the major facilitator superfamily (MFS) transporters, Trends Biochem. Sci. 38 (2013) 151–159. https://doi.org/10.1016/j.tibs.2013.01.003.
[190] L.A. Kelly, S. Mezulis, C. Yates, M. Wass, M. Sternberg, The Phyre2 web portal for protein modelling, prediction, and analysis, Nat. Protoc. 10 (2015) 845–858. https://doi.org/10.1038/nprot.2015-053.
[191] J. Yuvaniyama, J.M. Denu, J.E. Dixon, M.A. Saper, Crystal structure of the dual specificity protein phosphatase VHR., Science (80-.). 272 (1996) 1328–31. http://www.ncbi.nlm.nih.gov/pubmed/8650541.
[192] E.Y. Won, Y. Xie, C. Takemoto, L. Chen, Z.J. Liu, B.C. Wang, D. Lee, E.J. Woo, S.G. Park, M. Shirouzu, S. Yokoyama, S.J. Kim, S.W. Chi, High-resolution crystal structure of the catalytic domain of human dual-specificity phosphatase 26, Acta Crystallogr. Sect. D Biol. Crystallogr. 69 (2013) 1160–1170. https://doi.org/10.1107/S0907444913004770.
[193] R.S. Sankhala, R.K. Lokareddy, G. Cingolani, Structure of human PIR1, an atypical dual-specificity phosphatase., Biochemistry. 53 (2014) 862–71. https://doi.org/10.1021/bi401240x.

[194] J.D. Faraldo-Gómez, B. Roux, Electrostatics of ion stabilization in a ClC chloride channel homologue from Escherichia coli, J. Mol. Biol. 339 (2004) 981–1000. https://doi.org/10.1016/j.jmb.2004.04.023.
[195] W.G. Hol, P.T. van Duijnen, H.J. Berendsen, The alpha-helix dipole and the properties of proteins., Nature. 273 (1978) 443–446. https://doi.org/10.1038/273443a0.
[196] J. Cowan, Structural and catalytic chemistry of magnesium dependent enzymes, Biometals. 15 (2002) 225–235. https://doi.org/1016022730880.
[197] L. Sawyer, M.N. James, Carboxyl-carboxylate interactions in proteins., Nature. 295 (1982) 79–80. https://doi.org/10.1038/295079a0.
[198] L. Holm, L.M. Laakso, Dali server update., Nucleic Acids Res. 44 (2016) W351-5. https://doi.org/10.1093/nar/gkw357.
[199] T. Zhang, J. Liu, M. Fellner, C. Zhang, D. Sui, J. Hu, Crystal structures of a ZIP zinc transporter reveal a binuclear metal center in the transport pathway, Sci. Adv. 3 (2017) e1700344. https://doi.org/10.1126/sciadv.1700344.
[200] J. Heng, Y. Zhao, M. Liu, Y. Liu, J. Fan, X. Wang, Y. Zhao, X.C. Zhang, Substrate-bound structure of the E. coli multidrug resistance transporter MdfA, Cell Res. 25 (2015) 1060–1073. https://doi.org/10.1038/cr.2015.94.
[201] R. Dutzler, E.B. Campbell, M. Cadene, B.T. Chait, R. MacKinnon, X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity, Nature. 415 (2002) 287–294. https://doi.org/10.1038/415287a.
[202] D. Wöhlert, M.J. Grötzinger, W. Kühlbrandt, Ö. Yildiz, Mechanism of Na(+)-dependent citrate transport from the structure of an asymmetrical CitS dimer., Elife. 4 (2015) e09375. https://doi.org/10.7554/eLife.09375.
[203] J.W. Kim, S. Kim, S. Kim, H. Lee, J.-O. Lee, M.S. Jin, Structural insights into the elevator-like mechanism of the sodium/citrate symporter CitS, Sci. Rep. 7 (2017) 2548. https://doi.org/10.1038/s41598-017-02794-x.
[204] S. Radestock, L.R. Forrest, The alternating-access mechanism of MFS transporters arises from inverted-topology repeats, J. Mol. Biol. 407 (2011) 698–715. https://doi.org/10.1016/j.jmb.2011.02.008.
[205] W. Mi, Y. Li, S.H. Yoon, R.K. Ernst, T. Walz, M. Liao, Structural basis of MsbA-mediated lipopolysaccharide transport, Nature. 549 (2017) 233–237. https://doi.org/10.1038/nature23649.
[206] N. Ruiz, Lipid flippases for bacterial peptidoglycan biosynthesis, Lipid Insights. 2015 (2015) 21–31. https://doi.org/10.4137/Lpi.s31783.
[207] T.J. Silhavy, D. Kahne, S. Walker, The Bacterial Cell Envelope1 T. J. Silhavy, D. Kahne and S. Walker, ., Cold Spring Harb Perspect Biol. 2 (2010) 1–16. https://doi.org/10.1101/cshperspect.a000414.
[208] T.D.H. Bugg, D. Braddick, C.G. Dowson, D.I. Roper, Bacterial cell wall assembly: Still an attractive antibacterial target, Trends Biotechnol. 29 (2011) 167–173. https://doi.org/10.1016/j.tibtech.2010.12.006.
[209] N.A. Caveney, G. Caballero, H. Voedts, A. Niciforovic, L.J. Worrall, M. Vuckovic, M. Fonvielle, J.E. Hugonnet, M. Arthur, N.C.J. Strynadka, Structural insight into YcbB-mediated beta-lactam resistance in Escherichia coli, Nat. Commun. 10 (2019) 1–11. https://doi.org/10.1038/s41467-019-09507-0.
[210] M.A. D'Elia, M.P. Pereira, Y.S. Chung, W. Zhao, A. Chau, T.J. Kenney, M.C. Sulavik, T.A. Black, E.D. Brown, Lesions in teichoic acid biosynthesis in Staphylococcus aureus

lead to a lethal gain of function in the otherwise dispensable pathway, J. Bacteriol. 188 (2006) 4183–4189. https://doi.org/10.1128/JB.00197-06.
[211] M. Fodje, K. Janzen, S. Labiuk, J. Gorin, P. Grochulski, AutoProcess:Automated strategy calculation, data processing & structure solution, Acta Crystallogr. Sect. A Found. Adv. 70 (2014) C791–C791. https://doi.org/10.1107/s2053273314092080.
[212] G. Winter, Xia2: An expert system for macromolecular crystallography data reduction, J. Appl. Crystallogr. 43 (2010) 186–190. https://doi.org/10.1107/S0021889809045701.
[213] G. Winter, D.G. Waterman, J.M. Parkhurst, A.S. Brewster, R.J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I.D. Young, N.K. Sauter, G. Evans, DIALS: Implementation and evaluation of a new integration package, Acta Crystallogr. Sect. D Struct. Biol. 74 (2018) 85–97. https://doi.org/10.1107/S2059798317017235.
[214] P.R. Evans, G.N. Murshudov, How good are my data and what is the resolution?, Acta Crystallogr. Sect. D Biol. Crystallogr. 69 (2013) 1204–1214. https://doi.org/10.1107/S0907444913000061.
[215] T.C. Terwilliger, R.W. Grosse-Kunstleve, P. V. Afonine, N.W. Moriarty, P.H. Zwart, L.W. Hung, R.J. Read, P.D. Adams, Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard, Acta Crystallogr. Sect. D Biol. Crystallogr. 64 (2007) 61–69. https://doi.org/10.1107/S090744490705024X.
[216] A.A. Lebedev, P. Young, M.N. Isupov, O. V. Moroz, A.A. Vagin, G.N. Murshudov, JLigand: A graphical tool for the CCP4 template-restraint library, Acta Crystallogr. Sect. D Biol. Crystallogr. 68 (2012) 431–440. https://doi.org/10.1107/S090744491200251X.
[217] N.W. Moriarty, R.W. Grosse-Kunstleve, P.D. Adams, Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation, Acta Crystallogr. Sect. D Biol. Crystallogr. 65 (2009) 1074–1080. https://doi.org/10.1107/S0907444909029436.
[218] M. Fujihashi, Y.W. Zhang, Y. Higuchi, X.Y. Li, T. Koyama, K. Miki, Crystal structure of cis-prenyl chain elongating enzyme, undecaprenyl diphosphate synthase, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 4337–4342. https://doi.org/10.1073/pnas.071514398.
[219] W. Sinko, C. de Oliveira, S. Williams, A. Van Wynsberghe, J.D. Durrant, R. Cao, E. Oldfield, J.A. Mccammon, Applying molecular dynamics simulations to identify rarely sampled ligand-bound conformational states of undecaprenyl pyrophosphate synthase, an antibacterial target, Chem. Biol. Drug Des. 77 (2011) 412–420. https://doi.org/10.1111/j.1747-0285.2011.01101.x.
[220] C.J. Kuo, R.T. Guo, I.L. Lu, H.G. Liu, S.Y. Wu, T.P. Ko, A.H.J. Wang, P.H. Liang, Structure-based inhibitors exhibit differential activities against Helicobacter pylori and Escherichia coli undecaprenyl pyrophosphate synthases, J. Biomed. Biotechnol. 2008 (2008) 1–7. https://doi.org/10.1155/2008/841312.
[221] M.A. Farha, E.D. Brown, Unconventional screening approaches for antibiotic discovery, Ann. N. Y. Acad. Sci. 1354 (2015) 54–66. https://doi.org/10.1111/nyas.12803.
[222] A.C.Y. Kuk, A. Hao, Z. Guan, S.Y. Lee, Visualizing conformation transitions of the Lipid II flippase MurJ, Nat. Commun. 10 (2019). https://doi.org/10.1038/s41467-019-09658-0.
[223] J.D. Brunner, S. Schenck, R. Dutzler, Structural basis for phospholipid scrambling in the TMEM16 family, Curr. Opin. Struct. Biol. 39 (2016) 61–70. https://doi.org/10.1016/j.sbi.2016.05.020.
[224] M. Marek, T. Günther-Pomorski, Assay of Flippase Activity in Proteoliposomes Using Fluorescent Lipid Derivatives, in: P-Type ATPases Methods Protoc., 2016: pp. 181–191.

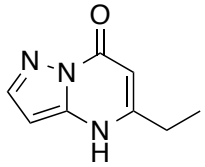
https://doi.org/10.1007/978-1-4939-3179-8_18.
[225] Y.W. Wu, K. Alexandrov, L. Brunsveld, Synthesis of a fluorescent analogue of geranylgeranyl pyrophosphate and its use in a high-throughput fluorometric assay for rab geranylgeranyltransferase, Nat. Protoc. 2 (2007) 2704–2711. https://doi.org/10.1038/nprot.2007.401.
[226] M.L. Eisinger, A.R. Dörrbaum, H. Michel, E. Padan, J.D. Langer, Ligand-induced conformational dynamics of the Escherichia coli Na+/H+ antiporter NhaA revealed by hydrogen/deuterium exchange mass spectrometry, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) 11691–11696. https://doi.org/10.1073/pnas.1703422114.
[227] A.K. Nielsen, I.R. Möller, Y. Wang, S.G.F. Rasmussen, K. Lindorff-Larsen, K.D. Rand, C.J. Loland, Substrate-induced conformational dynamics of the dopamine transporter, Nat. Commun. 10 (2019) 1–14. https://doi.org/10.1038/s41467-019-10449-w.
[228] I.R. Möller, M. Slivacka, A.K. Nielsen, S.G.F. Rasmussen, U. Gether, C.J. Loland, K.D. Rand, Conformational dynamics of the human serotonin transporter during substrate and drug binding, Nat. Commun. 10 (2019) 1–13. https://doi.org/10.1038/s41467-019-09675-z.
[229] P.S. Merkle, K. Gotfryd, M.A. Cuendet, K.Z. Leth-Espensen, U. Gether, C.J. Loland, K.D. Rand, Substrate-modulated unwinding of transmembrane helices in the NSS transporter LeuT, Sci. Adv. 4 (2018) 1–12. https://doi.org/10.1126/sciadv.aar6179.
[230] I.R. Möller, M. Slivacka, J. Hausner, A.K. Nielsen, E. Pospíšilová, P.S. Merkle, R. Lišková, M. Polák, C.J. Loland, A. Kádek, P. Man, K.D. Rand, Improving the Sequence Coverage of Integral Membrane Proteins during Hydrogen/Deuterium Exchange Mass Spectrometry Experiments, Anal. Chem. 91 (2019) 10970–10978. https://doi.org/10.1021/acs.analchem.9b00973.
[231] C. Martens, M. Shekhar, A.M. Lau, E. Tajkhorshid, A. Politis, Integrating hydrogen–deuterium exchange mass spectrometry with molecular dynamics simulations to probe lipid-modulated conformational changes in membrane proteins, Nat. Protoc. 14 (2019) 3183–3204. https://doi.org/10.1038/s41596-019-0219-6.
[232] M. Kesherwani, D. Velmurugan, Molecular insights into substrate binding mechanism of undecaprenyl pyrophosphate with membrane integrated phosphatidyl glycerophosphate phosphatase B (PgpB) using molecular dynamics simulation approach, J. Biomol. Struct. Dyn. 37 (2019) 1062–1089. https://doi.org/10.1080/07391102.2018.1449666.
[233] J.B. Fournier, E. Rebuffet, L. Delage, R. Grijol, L. Meslet-Cladière, J. Rzonca, P. Potin, G. Michel, M. Czjzek, C. Leblanc, The vanadium iodoperoxidase from the marine Flavobacteriaceae species Zobellia galactanivorans reveals novel molecular and evolutionary features of halide specificity in the vanadium haloperoxidase enzyme family, Appl. Environ. Microbiol. 80 (2014) 7561–7573. https://doi.org/10.1128/AEM.02430-14.
[234] R.D. Makde, S.K. Mahajan, V. Kumar, Structure and mutational analysis of the PhoN protein of Salmonella typhimurium provide insight into mechanistic details, Biochemistry. 46 (2007) 2079–2090. https://doi.org/10.1021/bi062180g.
[235] V.I. Petrou, C.M. Herrera, K.M. Schultz, O.B. Clarke, J. Vendome, D. Tomasek, S. Banerjee, K.R. Rajashankar, M.B. Dufrisne, B. Kloss, E. Kloppmann, B. Rost, C.S. Klug, M.S. Trent, L. Shapiro, F. Mancia, Structures of aminoarabinose transferase ArnT suggest a molecular basis for lipid A glycosylation, Science (80-.). 351 (2016) 608–612. https://doi.org/10.1126/science.aad1172.
[236] P. Gourdon, J.L. Andersen, K.L. Hein, M. Bublitz, B.P. Pedersen, X.-Y. Liu, L. Yatime,

M. Nyblom, T.T. Nielsen, C. Olesen, J.V. Møller, P. Nissen, J.P. Morth, HiLiDe—Systematic Approach to Membrane Protein Crystallization in Lipid and Detergent, Cryst. Growth Des. 11 (2011) 2098–2106. https://doi.org/10.1021/cg101360d.

Appendices

Appendix A Experimental details for the synthesis of JPD-447

5-Ethylpyrazolo[1,5-a]pyrimidin-7(4H)-one

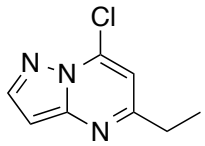


Beige solid (920 mg, 94%). m.p. 246-248°C; H-NMR (400 MHz, d^6-DMSO) δ 12.23 (s, 1H), 7.83 (s, 1H), 6.10 (s, 1H), 5.58 (s, 1H), 2.57 (q, *J*= 7.3 Hz, 2H), 1.21 (t, *J*= 7.3 Hz, 3H); ^{13}C-NMR (100 MHz, d^6-DMSO) δ 156.9, 155.6, 143.0, 142.0, 93.7, 88.7, 26.0, 12.8; HRMS (EI) ([M] calcd. for $C_8H_9N_3O$ 163.0746) found 163.0742.

General procedure:

3-Aminopyrazole (1 equiv.) and corresponding keto-ester (1.2 equiv.) were stirred in acetic acid (1M) at 90°C for 16 h. After cooling to room temperature, Et_2O was added and the heterogeneous mixture stirred for 30 mins at 0°C. The product was then isolated by filtration, washed with further Et_2O and dried under vacuum.

7-Chloro-5-ethylpyrazolo[1,5-*a*]pyrimidine



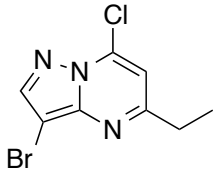
Isolated by flash chromatography (0→30% Et_2O/pentane, R_f = 0.4) as a green solid (285 mg, 85%). m.p. 68-70 °C; ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.02 (d, *J*= 2.1 Hz, 1H), 6.71 (s, 1H), 6.54 (d, *J*= 2.1 Hz, 1H), 2.71 (q, *J*= 7.4 Hz, 2H), 1.21 (t, *J*= 7.4 Hz, 3H); ^{13}C-NMR (100 MHz,

$CDCl_3$) δ 162.9, 149.4, 145.0, 138.0, 107.7, 97.2, 31.1, 12.4; HRMS (EI) ([M] calcd. for $C_8H_8N_3Cl$ 181.0407) found 181.0402.

General procedure:

To a flame-dried flask under argon was added pyrazolopyrimidinone (1 equiv.). Anhydrous toluene was then added, followed by sequential addition of dimethylaniline (1.25 equiv.) and phosphorous oxychloride (3 equiv.). The mixture was then placed in a pre-heated oil-bath and stirred at 90°C for 16 h. The mixture was cooled to 0°C, quenched with 2M $NaOH_{(aq.)}$ and the organics extracted with CH_2Cl_2 (3x). Combined organics were washed with brine, dried over $MgSO_4$ filtered and evaporated. The crude product was then purified by flash column chromatography.

3-Bromo-7-chloro-5-ethylpyrazolo[1,5-*a*]pyrimidine

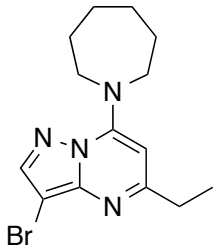


Isolated by flash chromatography (0→30% Et_2O/pentane, R_f = 0.3) as a yellow oil (96%). ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.18 (s, 1H), 6.86 (s, 1H), 2.85 (q, *J*= 7.4 Hz, 2H), 1.31 (t, *J*= 7.4 Hz, 3H); ^{13}C-NMR (100 MHz, $CDCl_3$) δ 164.5, 145.9, 145.0, 138.6, 108.5, 84.6, 31.4, 12.7; HRMS (EI) ([M] calcd. for $C_8H_7BrClN_3$ 260.9491, 258.9512) found 260.9486, 258.9501.

General procedure:

The corresponding 7-chloropyrimidine was stirred in dry CH_2Cl_2 (0.25M) under argon, followed by addition of N-bromosuccinimide (1.1 equiv.) in one portion. The reaction mixture was stirred at room temperature for 16 h and the solvent evaporated. The crude product was purified by flash column chromatography.

7-(Azepan-1-yl)-3-bromo-5-ethylpyrazolo[1,5-*a*]pyrimidine

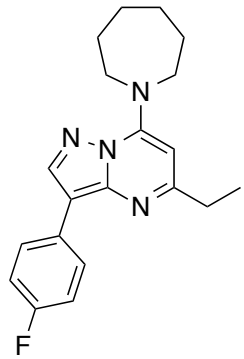


Isolated by flash chromatography (0→30% Et_2O/pentane, R_f = 0.2) as a beige solid (42.1 mg, 97%). m.p. 96-98°C; ^{1}H-NMR (400 MHz, $CDCl_3$) δ 7.89 (s, 1H), 5.85 (s, 1H), 4.00-3.92 (m, 4H), 2.77 (q, *J*= 7.4 Hz, 2H), 1.95-1.87 (m, 4H), 1.68-1.60 (m, 4H), 1.32 (t, *J*= 7.4 Hz, 3H); ^{13}C-NMR (100 MHz, $CDCl_3$) δ 165.0, 149.6, 147.9, 142.6, 89.4, 80.7, 52.0, 31.7, 28.5, 26.6, 13.6; HRMS (EI) ([M] calcd. for $C_{14}H_{19}BrN_4$ 324.0773, 322.0793) found 324.0774, 322.0793.

General procedure:

The corresponding 3-bromo-7-chloropyrimidine (1 equiv.) was stirred in isopropanol (0.4M), followed by addition of amine (1.2 equiv.) and DIPEA (1.5 equiv.). The mixture was then placed in a pre-heated oil-bath at 70°C and stirred for 16 h. After cooling to room temperature, CH_2Cl_2 was added and the organics washed with water (3x) and brine. The organics were then dried over $MgSO_4$, filtered and evaporated. The crude product was purified by flash column chromatography or by passing through a silica gel plug.

7-(Azepan-1-yl)-5-ethyl-3-(4-fluorophenyl)pyrazolo[1,5-*a*]pyrimidine



Isolated by flash chromatography (0→10% Et_2O/pentane, R_f = 0.2) as an off-white solid (120 mg, 96%). m.p. 111-113°C; ^{1}H-NMR (400 MHz, $CDCl_3$) δ 8.23 (s, 1H), 8.08 (dd, 7.8, 5.8 Hz, 2H), 7.13 (t, *J*= 8.8 Hz, 2H), 5.91 (s, 1H), 4.01 (t, *J*= 5.6 Hz, 4H), 2.84 (q, *J*= 7.5 Hz, 2H), 2.00-1.92 (m, 4H), 1.72-1.66 (m, 4H), 1.40 (t, *J*= 7.5 Hz, 3H); ^{13}C-NMR (100 MHz, $CDCl_3$) δ 163.9, 160.8 (d, J^{C-F}= 241 Hz), 149.6, 147.6, 140.7, 129.3, 127.2 (d, J^{C-F}= 8 Hz), 115.2 (d, J^{C-F}= 21 Hz), 106.1, 89.1, 51.9, 31.5, 28.5, 26.7, 13.1; ^{19}F NMR (376.5 MHz, $CDCl_3$) δ -117.8; HRMS (EI) ([M] calcd. for $C_{20}H_{23}FN_4$ 338.1907.2063) found 338.1918.

General procedure:

Heteroaryl bromide (0.1 mmol), boronic acid (0.15 mmol), Pd-PEPPSI IPent^{Cl}-2-picoline (2.5 mg, 0.003 mmol) and Cs_2CO_3 (81 mg, 0.25 mmol) were added to an oven-dried vial equipped with a magnetic stirrer bar. Degassed dioxane/H_2O (5:1, 300 μL) was then added and the vial purged with argon. The mixture was then placed in a pre-heated oil-bath at 80°C and stirred for 20 h. After cooling to room temperature, solvent was evaporated and the crude product purified by flash column chromatography.