The Dissertation Committee for Jeremy Christopher Henderson Certifies that this is the approved version of the following dissertation:

Enzymology of a Three Gene Pathway Required for Cationic Antimicrobial Peptide Resistance by *Vibrio cholerae*

Committee:

M. Stephen Trent, Supervisor

Bryan Davies, Co-Supervisor

Jennifer Brodbelt

Jeffrey Barrick

Marvin Whiteley

Enzymology of a Three Gene Pathway Required for Cationic Antimicrobial Peptide Resistance by *Vibrio cholerae*

by

Jeremy Christopher Henderson, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2016

Dedication

For my grandmother Nina Henderson (1923-2015) and previous graduate mentor Christian R.H. Raetz (1946-2011), in memoriam.

Acknowledgements

It is impossible to adequately express my gratitude to colleagues, faculty, staff, family, and friends whose support over the years made this document possible. Research is a consuming endeavor that requires patience, efficient time-management, and a good sense of humor. First, this work does not exist without Jessica Hankins. Thank you! Without the masterful guidance of my graduate mentor, Jedi Master, and friend, M. Stephen Trent, I wouldn't have the tools to succeed in this demanding profession. Similarly appreciated is the continued support of my undergraduate research advisor, Jane E. Jackman, who catalyzed my pathway to a career in science. The influence on my development by my late graduate mentor Christian R.H. Raetz is immeasurable, sorely missed, and forever appreciated. Hope you enjoy this one from the great beyond, especially Chapter 6 :) The support of my thesis advisory committee at Duke and The University of Texas is equally appreciated. Thank ya'll for taking time from your busy schedules to discuss science. All of these wonderful human beings are responsible for shaping a young punk into a competent scientist, and have my deepest gratitude.

Colleagues at 3 universities made for an enriching and rewarding graduate experience. Hak Suk Chung and Carmen Herrera ensured I always did my best work. 사랑해요. Te amo. Provocative scientific discussions were always had with wonderful lab mates: Alex Crofts, Brittany Needham, Aaron Pride, Emily Nowicki, Davy Giles, Joe Boll, Petko Ivanov, Thomas Cullen, Ryan Emptage, Hayley Young, Jinshi Zhao, Ali Masoudi, Greg Laird, Brandon Tan among many others. Friends from other labs helped in matters of science and were heavily leaned upon for emotional support. Thanks especially to Aaron Conrado, Brad Hover, David Reid, Charlie Cooper, Apollo Stacy, Shane Chen, Ruby Coates and Buddy O'Brien. And of course, thanks Mom! Love ya'll!

Enzymology of a Three Gene Pathway Required for Cationic Antimicrobial Peptide Resistance by *Vibrio cholerae*

Jeremy Christopher Henderson, Ph.D.

The University of Texas at Austin, 2016

Supervisors: M. Stephen Trent & Bryan Davies

Produced by all domains of life, cationic antimicrobial peptides (CAMPs) are the most ubiquitous antibacterial compound in nature. In higher order eukaryotes CAMPs are critical components of innate immunity. Unicellular eukaryotes and prokaryotes use CAMPs for competitive exclusion in ecological niches. Some classes of CAMPs, such as polymyxins, are used as last-resort antibiotics in the treatment of bacterial infections. However, many bacteria have evolved resistance mechanisms to the toxic effects of CAMPs. For example, the current pandemic strain of *Vibrio cholerae* (O1 El Tor) is resistant to polymyxins, whereas the previous pandemic strain (O1 classical) is polymyxin sensitive. For decades how El Tor *V. cholerae* evolved polymyxin resistance was unknown, until the recent identification of a three-gene operon responsible for >100-fold improvement in resistance to polymyxin B. Classical strains lack a functional version of this operon. Renamed *almEFG*, this operon was shown to be necessary for the esterification of one, sometimes two, glycine residues to the major Gram-negative surface molecule lipopolysaccharide (LPS). Covalent attachment of glycine to LPS likely prevents CAMP binding through minimization of the negatively charged bacterial surface, since the amine terminus of glycine or diglycine remains freely exposed. Here, a mechanistic description of the *almEFG* operon is provided. Guided by predictive models, AlmF is shown to be a genuine aminoacyl carrier protein. Also identified is the *V. cholerae* enzyme required for post-translational activation of AlmF to a functional carrier protein. A combination of biochemical approaches reveal that AlmE specifically adds glycine to activated AlmF. Ultimately, AlmG transfers glycine from glycyl-AlmF to the lipid A membrane anchor of surface-displayed LPS *in vivo*, a process that evidence indicates can be done sequentially for diglycine LPS modification. The trio of proteins in the AlmEFG system forms a chemical pathway that resembles the division of labor observed in non-ribosomal peptide synthetases. They also bear striking resemblance to a related Gram-positive cell-wall remodeling strategy that promotes CAMP resistance. This biochemical study has the power to inform therapeutics against *V. cholerae* infection and provides a highly detailed mechanism of Gram-negative CAMP resistance.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction[1]

1.1 The Gram-negative Outer Membrane: Composition, Assembly And Maintenance of An Asymmetric Lipid Bilayer

The defining feature of gram-negative bacteria is the presence of a second membrane bilayer that surrounds the peptidoglycan sacculus (Figure 1.1). The inner bacterial membrane bilayer contains glycerophospholipids (PLs) in both leaflets, whereas the "outer membrane" is an asymmetrical lipid bilayer[1], with PLs in the periplasmic leaflet and a surface-exposed outer leaflet composed of lipopolysaccharide (LPS) molecules. This unique organization affords gram-negative organisms protection from large polar molecules restricted by a typical membrane bilayer, and also from lipophilic compounds. The evidence for the asymmetrical organization of the outer membrane was reported by Nikaido *et al*. in the mid 70s[1]. In their work it was demonstrated that PLs of intact gram-negative bacteria were not susceptible to degradation by phospholipases or to chemical labeling by macromolecular reagents unable to cross the outer membrane. The topological insight provided by Nikaido and colleagues, in combination with the known

[1] Portions of this chapter have been previously published. (i) Henderson, J.C., Zimmerman, S.M., Crofts, A.A., Boll, J.M., Kuhns, L.G., Herrera, C.M., and Trent, M.S. The Power of Asymmetry: Architecture and assembly of the Gram-Negative outer membrane lipid bilayer. *Annu Rev Microbiol* 2016. JCH had primary writing and editing responsibility for this review (ii) Herrera, C.M., Crofts, A.A., Henderson, J.C. , Pingali, S.C., Davies, B.W., and Trent, M.S. Vibrio cholerae VprA-VprB two-component system controls virulence through endotoxin modification *MBio* 2014. JCH contributed data, helped write and edit portions of the manuscript (iii) Henderson, J.C., Fage, C.D., Cannon, J.R., Brodbelt, J.S., Keatinge-Clay, A.T., and Trent, M.S. Antimicrobial peptide resistance of Vibrio cholerae results from an LPS modification pathway related to non-ribosomal peptide synthetases. *ACS Chem Biol*. **86**, 2138-45 2014. JCH along with equal contribution from co-author CDF contributed data, wrote, and edited this report.

lipid composition of isolated outer membrane bilayers, indicated that PLs are strictly confined to the inner leaflet of the outer membrane.

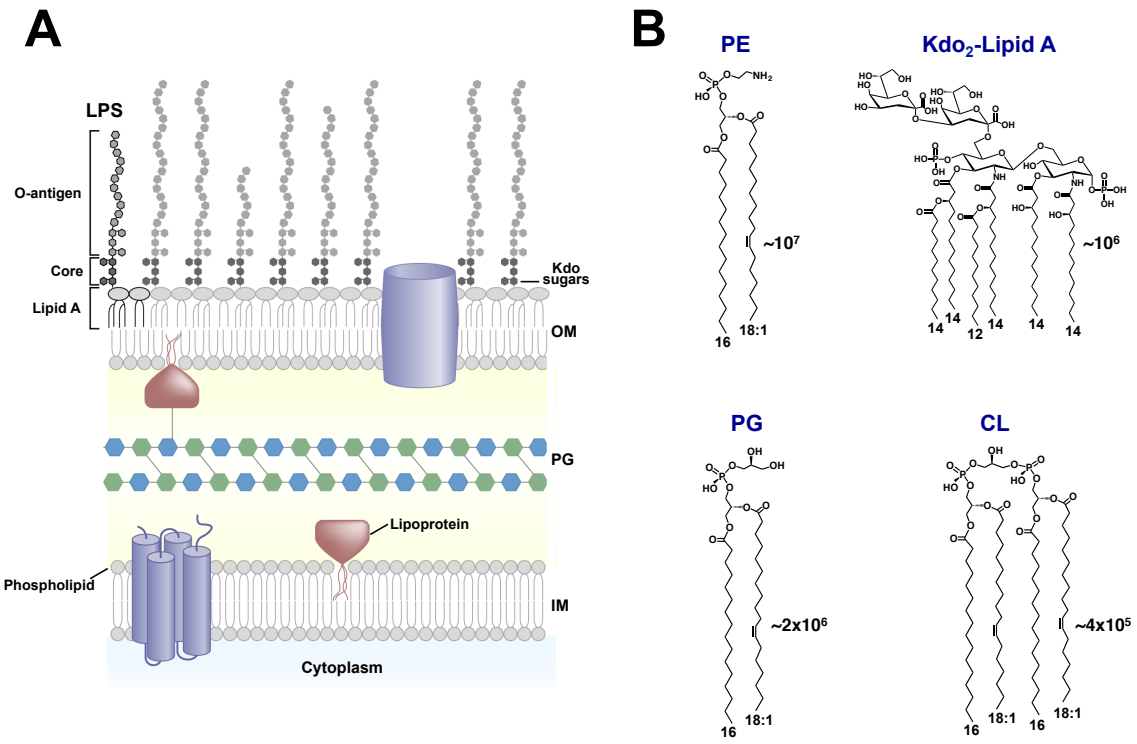


Figure 1.1 The cell envelope and major lipids of prototypical gram-negative bacteria.

(***a***) The typical inner membrane and outer membrane bilayers are separated by the periplasmic compartment, which contains the peptidoglycan layer. The inner membrane is a bilayer of glycerophospholipids, whereas the outer membrane is an asymmetrical bilayer, with glycerophospholipids found in the inner periplasmic leaflet and lipopolysaccharide localized to the outer-surface-exposed leaflet. (***b***) The major lipids of *E. coli* K12, and focus of this dissertation O1 *Vibrio cholerae*, are keto-3-deoxy-D-*manno*-oct-2-ulosonic acid $(Kdo)_2$-lipid A, phosphatidylethanolamine (PE), phosphatidylglycerol (PG), and cardiolipin (CL). The length and composition of fatty acyl chains of each major lipid type, as well as the number of molecules within a bacterial cell during exponential growth phase, are indicated.

Given that the outer membrane is essential for gram-negative organisms and protects bacteria from environmental stresses (e.g. antibiotics) it is of no surprise that it has been the subject of intense research. Over the last two decades, there has been enormous progress towards understanding outer membrane assembly and the maintenance of bilayer lipid asymmetry. One key question has always been how lipophilic components of the cell envelope, assembled at the cytoplasmic leaflet of the inner membrane, are guided through the crowded, aqueous periplasmic compartment. Other areas of intense focus include the biosynthesis and remodeling of key outer membrane components, especially the how chemical modifications to surface lipids, lipoproteins, β-barrel proteins and/or assorted polysaccharides contribute towards bacterial fitness and antibiotic resistance.

1.1.1 Biosynthesis and biology of glycerophospholipids found in model organisms.

Dr. Eugene Kennedy and members of his laboratory combined molecular genetics and biochemical approaches to identify and characterize the genes responsible for glycerophospholipid (PL) metabolism in *Escherichia coli* and *Saccharomyces cerevisiae* over 40 years ago[2–4]. The *E. coli* PL biosynthetic paradigm serves as the backbone for most prokaryotes (Figure 1.2), including *Vibrio cholerae*, while the *S. cerevisiae* pathway generally holds true for related fungal species. Both model organisms contain elements of more complex lipid anabolism observed in higher order eukaryotes. Today, the combination of genomic information with unprecedented sensitivity in detection of lipid species by mass spectrometry, has enabled identification of previously unknown genes

involved in lipid metabolism, even in well-trekked model organisms. Exciting discoveries in *E. coli* include a third phosphatidylglycerophosphate (PGP) phosphatase and two novel cardiolipin (CL) synthases (Fig. 2), while in yeast and humans a new PGP phosphatase and a mitochondrial cytidine diphosphate diacylglycerol synthetase have been identified[5–9]. Mutations in these eukaryotic genes lead to severe mitochondrial dysfunction, as a consequence of CL loss. However, the biological role of CL in *E. coli*, despite the existence of multiple, differentially regulated orthologs, remains poorly understood.

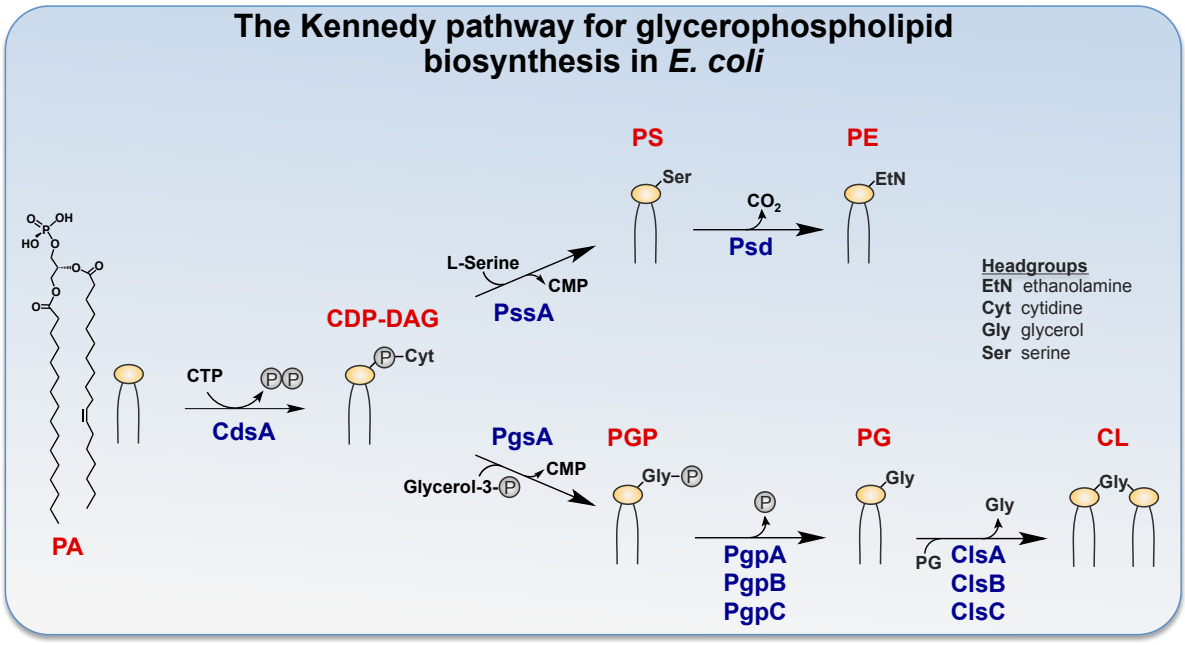


Figure 1.2 The Kennedy pathway for glycerophospholipid biosynthesis in *Escherichia coli,* and prototypical gram-negatives.

Eugene Kennedy and coworkers discovered the critical role of cytidine nucleotides as activating groups for subsequent phosphatidyl-transfer reactions in glycerophospholipid biosynthesis. Enzymes involved in these pathways are either intrinsic membrane proteins or membrane associated. Phosphatidic acid (PA), de novo synthesized on the cytoplasmic surface of the inner membrane, is converted to cytidine diphosphate-diacylglycerol (CDP-DAG) by CDP-DAG synthase (CdsA) in a reaction that utilizes cytidine triphosphate (CTP) with release of pyrophosphate. CDP-DAG functions as a donor of phosphatidyl moieties for the biosynthesis of phosphatidylserine (PS) and phosphatidylglycerol-3-phosphate (PGP) with release of cytidine monophosphate (CMP). The PS formed by phosphatidylserine synthase (PssA) is rapidly decarboxylated by phosphatidylserine decarboxylase (Psd) to generate the predominant glycerophospholipid (~70%) found in *E. coli*, phosphatidylethanolamine (PE). In phosphatidylglycerol (PG) synthesis, PGP is quickly dephosphorylated by one of several inner membrane phosphatases (e.g., PgpA) to form PG. Lastly, cardiolipin is produced through the condensation of two PG molecules by cardiolipin synthase A or B, (ClsA, ClsB) or the condensation of PG and PE by ClsC (ClsA being the major cardiolipin synthase). Enzyme names are in blue, lipid species in brown, and cofactors/ substrates in black.

Despite decades of research on bacterial PL metabolism, CL's enigmatic biological function in bacteria is not exceptional. Overall the physiological importance of PL head-group heterogeneity remains ill defined, yet some insightful reports seem to establish overarching themes. Topology of integral membrane proteins is affected by interaction with specific lipid partners, while perhaps not surprisingly, numerous peripheral membrane proteins associate preferentially with anionic PLs such as phosphatidic acid, phosphatidylglycerol (PG), and CL[10–14], Anionic PL interactions have been implicated as necessary for SecA dependent protein translocation and *oriC* dependent DNA replication[15–19]. It is also well reported that PLs serve as precursors in the chemical modification of sugars, proteins, and even other lipids. In many organisms, the diacylglycerol moiety of PG is used in the post-translational modification of lipoproteins, a chemical modification required for proper insertion of lipoproteins into the membrane[20–24] (Section 1.1.3: Figure 1.3, Panel B). During processing, lipoproteins are also aminoacylated at an *N*-terminal cysteine residue, where PLs serve as substrate acyl chain donors[22,25] (Section 1.1.3: Figure 1.3, Panel B). Membrane derived oligosaccharide, an enigmatic osmoregulatory component found in the periplasm of some gram-negatives, can be modified with PG and phosphatidylethanolamine (PE) headgroups[26–29]. The Kdo-lipid A domain and core oligosaccharide of LPS can also be modified with PL headgroups or acyl chains[30–33]. Bacteria with these modifications show increased resistance to cationic antimicrobial peptides, especially polymyxins that bind Kdo-lipid A specifically, as well an ever-growing list of other phenotypes[31] (Section 1.2).

1.1.2 The mysteries of Gram-negative PL transport: from inner membrane sites of synthesis, through the periplasm, and to the outer membrane.

PL transport to the outer membrane (OM) presents a unique challenge to gram-negative bacteria. Lipoproteins and β-barrel outer membrane proteins contain specific amino acid signal sequences for export to the OM, whereas LPS is strictly compartmentalized to the OM surface by dedicated transport machinery (Section 1.1.5). On the contrary, PLs appear randomly distributed between inner membrane bilayer and the periplasmic leaflet of the outer membrane, with no major observable difference in PL composition between these lipid surfaces. The process by which PLs are distributed between sites of synthesis on the inner membrane (IM) to the periplasmic leaflet of the OM is unknown. Transport machinery exists for other OM components, so an analogous mechanism likely exists to distribute the cellular PL pool between inner and outer membranes. The rate of spontaneous, passive transfer of PLs between separated membranes in vitro is not sufficient to support bacterial growth[34], thus an active mechanism is likely needed to facilitate movement of PLs across the periplasmic compartment.

The most detailed account of PL transport from the inner to outer membrane was done by Donohue-Rolf et al. over 30 years ago[34]. To determine the in vivo translocation rate of newly synthesized PLs, pulse-chase radiolabeling of cellular PL pools was used to track and compare PLs between inner and outer bilayers. The results indicated distinct rates of PL transportation with rapid transport of PG or CL, and slower rates of PE transport to the OM. Depletion of cellular ATP and inhibition of lipid or protein synthesis

did not appear to alter PE translocation rates, but inhibitors that target proton motive force significantly reduced the PE translocation rate. The authors postulate that zones of adhesion, or Bayer's bridges, between OM and IM may provide a route for phospholipid movement between membranes, and are furthermore controlled by the proton motive force[35]. However, the existence of these zones in vivo remains controversial[36], and despite these works no essential anterograde transport mechanisms have been characterized.

Conditionally lethal mutants of MsbA accumulate LPS precursors and PLs in the IM, leading some to conjecture that MsbA serves as a general flippase for both lipid molecules[37,38]. Further characterization has undoubtedly shown MsbA is responsible for flipping nascent LPS intermediates from the cytoplasmic to periplasmic leaflet of the IM. The role of MsbA in general PL transport is less clear, complicated by the fact that accumulation of LPS precursors could impact an unknown, independent PL transport mechanism. Evidence in support of this idea comes from a study in lipooligosaccharide (LOS)-deficient *Neisseria meningitides*, which convincingly demonstrates that MsbA is not required for PL distribution to the OM[39].

1.1.3 Removal of PLs from the outer membrane via retrograde transport.

The Mla (*m*aintenance of OM *l*ipid *a*symmetry) pathway is an intermembrane transport system proposed to function in the removal of accumulated PLs in the OM by transporting them to the IM[40]. The proteins that make up the Mla system contain elements that are similar to those of the Lol lipoprotein transport apparatus[41] (Figure 1.3; compare panel A to B). Mla is made up of an ABC transporter (MlaFEDB), a periplasmic protein

(MlaC), and a lipoprotein tethered to the OM (MlaA). Deletion of any Mla component leads to OM defects quantifiable by combinatorial SDS-EDTA sensitivity. Each single Mla component mutant displays identical sensitivity to this combination of compounds, indicative that they comprise a singular pathway. In *Shigella flexneri,* mutations in Mla homologs fail to form plaques after invasion of human host cells, a proxy for intracellular spread[42]. Like *E. coli*, these mutants are also sensitive to SDS-EDTA.

Suppressor mutations that repair the defective outer membrane phenotype (SDS-EDTA sensitivity) arise in the promoter region of *pldA*, encoding OM phospholipase A [40]. qRT-PCR analysis shows increased *pldA* expression in these stable suppressor strains, where a biological reporter assay, indirectly showed, PldA overexpression can limit PLs mislocalized to the OM surface[40]. PldA has long been thought to contribute to OM lipid asymmetry under certain growth conditions, by degrading PLs that have miscompartmentalized to the outer leaflet of the OM[43–45]. PldA mutants do not accumulate PLs in the OM outer leaflet to the dramatic extent seen in Mla pathway mutants, indicating the necessity of Mla-based maintenance of OM asymmetry. It is important to note, however, that direct biochemical evidence in support of Mla retrograde PL transport from OM to IM is lacking.

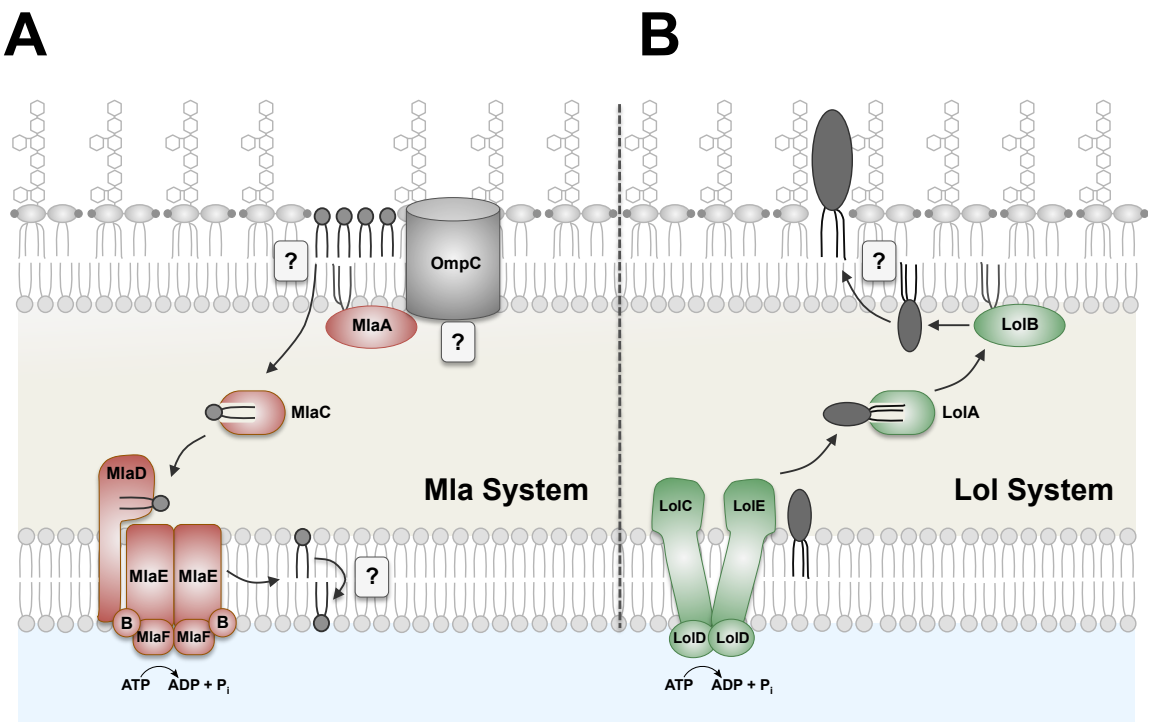


Figure 1.3 Comparison of the Mla and Lol periplasmic transport systems.

(**A**) The proposed model of the retrograde Mla (**m**aintenance of outer membrane **l**ipid **a**symmetry) system that is responsible for the removal of mislocalized glycerophospholipids from the gram-negative outer membrane. It is unclear whether the Mla system removes excess glycerophospholipids (e.g., phosphatidylethanolamine) from the inner leaflet prior to migration to the cell surface or if the system removes glycerophospholipids directly from the outer leaflet. It has been proposed that OmpC functions in a complex with the lipoprotein MlaA to extract phospholipids from the outer leaflet. Once extracted, the targeted lipid is delivered to the soluble periplasmic transporter MlaC, which delivers the substrate to the MlaFEDB ABC transporter complex. Using energy from ATP hydrolysis, the targeted lipid is reinserted into the inner membrane; however, the fate of the lipid is unknown. The biochemical mechanisms of substrate recognition, membrane extraction, and membrane insertion remain to be elucidated. (**B**) Schematic representation of the Lol (**l**ocalization **o**f **l**ipoproteins) transport system. The LolCDE ABC transporter complex utilizes ATP to direct the movement of mature (i.e., acylated) lipoproteins destined for the outer membrane by LolA. The LolA-lipoprotein complex transports target lipoproteins across the periplasm to LolB, which inserts lipoproteins into the outer membrane. It is unclear how surface-exposed lipoproteins are transported to the outer leaflet of the outer membrane.

1.1.4 Gram-negative biosynthesis of the endotoxic Kdo- lipid A domain of lipopolysaccharide LPS

Over a span of 30 years, the genes and molecular intermediates along the near-completely conserved biosynthetic pathway for Kdo-lipid A assembly in gram-negatives have been characterized (Figure 1.4). Comprehensive biochemical descriptions of pathway enzymes have been revealed through crystallography based structure determination and classical enzymology. Nine enzymes are required for canonical $(Kdo)_2$-lipid A biosynthesis in *E. coli*[33], and are considered essential for growth in nearly all gram-negatives with the exception of the late-step acyltransferases LpxL, LpxM and their various homologs[46–48]. While not all organisms contain high identity homologs to each gene in the *E. coli* pathway, each enzymatic step is universally conserved by organism specific orthologs. For example UDP-2,3-diacyl-GlcN hydrolase activity (Figure 1.4) is catalyzed by LpxH in most γ and β-proteobacteria, whereas LpxI performs this role in α-proteobacteria and other bacterial groups[49]. A very similar taxonomic distribution is seen in late acyltransferases, where LpxJ substitutes for LpxM in many organisms[50] (Figure 1.4). Departures from the canonical biosynthetic structure include variations in acyl chain length and degree of saturation, important for homeoviscous adaptation, as controlled by selective "hydrocarbon ruler" domains of LpxA or LpxD and their functional equivalents[51–54]. Nuanced selectivity in number of Kdo sugars, and modifications thereof (e.g. phosphorylation or hydroxylation), represent other organism specific variations to the canonical biosynthetic route[55,56].

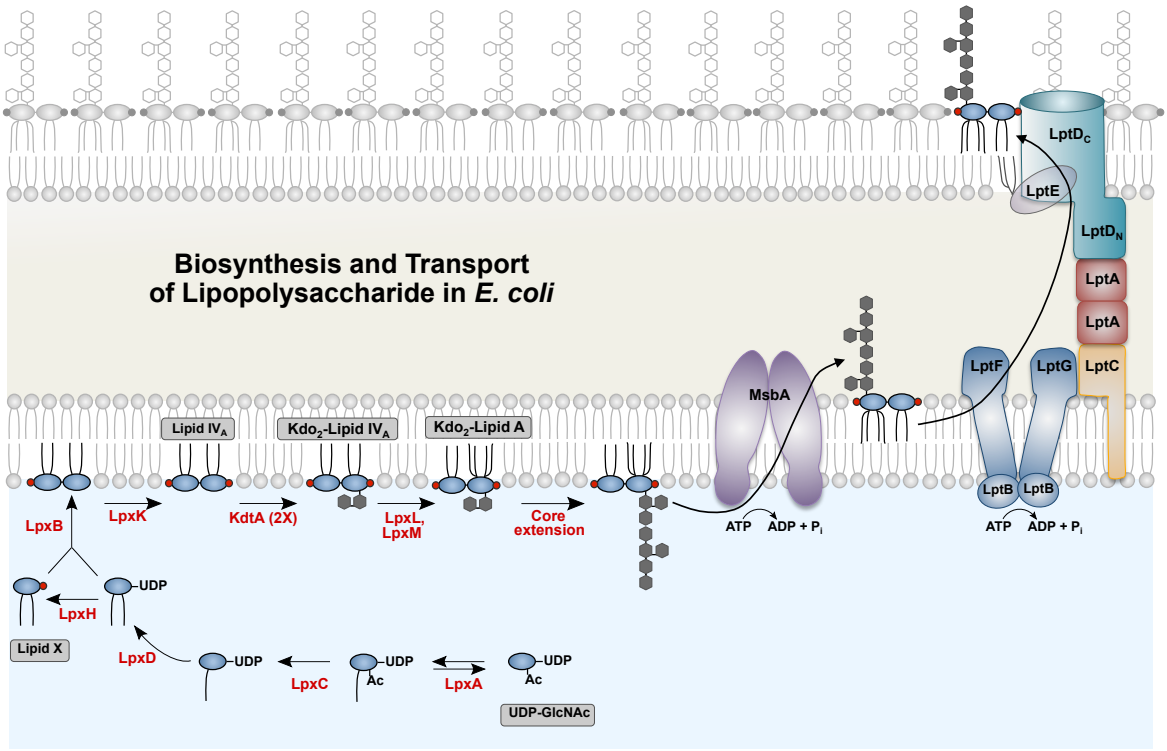


Figure 1.4 Biosynthesis and transport of lipopolysaccharide (LPS) in *E. coli*.

The nine-step Raetz pathway results in the biosynthesis of the canonical Kdo_2-lipid A substructure of LPS, required for the growth of most gram-negative bacteria. All reactions in the pathway are catalyzed by a single enzyme (*red names*) and occur at the interface between the cytosol and the inner membrane. Acyl-ACP (acyl carrier protein) is the preferred donor for each acylation event, and each acyltransferase utilizes an active site hydrocarbon ruler providing acyl chain specificity. Although Kdo is technically part of the core oligosaccharide, Kdo transfer is required for lipid A biosynthesis because the final two steps, catalyzed by LpxL and LpxM, require the presence of covalently attached Kdo. The remaining core oligosaccharide is extended at the cytoplasmic face of the inner membrane, requiring various glycosyl transferases. MsbA, the core Kdo-lipid A domain flippase, functions as an ABC transporter, moving core Kdo-lipid A domains to the periplasmic face of the inner membrane. For simplicity, the addition of O-antigen that typically occurs in the periplasm is not shown. The intermembrane translocation of mature LPS is carried out by the Lpt (LPS transport) system, which forms an envelope-spanning translocation machine. LPS is removed from the inner membrane via the ABC transporter LptBFG and delivered to LptC. Along with the periplasmic domain of LptC, soluble LptA forms a periplasmic bridge with the N-terminal domain of the outer membrane protein LptD. LptE sits within the β barrel formed by the C-terminal domain of LptD and promotes passage to the bacterial surface.

Compounds that inhibit early steps in Kdo-lipid A domain synthesis are candidates for developing broad-spectrum antibacterials. However, impediments to the successful *in vivo* implementation of Raetz pathway inhibitors remain symptoms of antibacterial development at large. Narrow spectrum antibacterial activity, insolubility, and off target protein binding remain formidable barriers to overcome. Compound screens have successfully identified competitive inhibitors towards early-step acyltransferases LpxA and LpxD[57], and a sulfonyl piperazine inhibitor of LpxH was recently discovered[58]. These compounds are effective against specific gram-negative taxa. As the committed step for Kdo-lipid A biosynthesis, and with no homology to any protein domain in humans, LpxC inhibitors remain the most promising broad-spectrum target[59,60]. For many years CHIR-090 has been explored as a powerful inhibitor of many bacterial LpxC homologs[61–64]. A fine-tuned understanding of the CHIR-090-LpxC interaction has lead to the design of dozens of novel compounds with differing specificity and increased potency to various LpxC homologs[63,65–69]. Compounds that inhibit the synthesis or transfer of Kdo to tetra-acylated lipid A have also been described; however, their in vivo efficacy is unclear[70–75].

1.1.5 Final assembly and transport of lipopolysaccharide (LPS) to the outer membrane.

Core-oligosaccharide is assembled stepwise on nascently synthesized Kdo-lipid A domains by specific glycosyltransferases that utilize activated sugar nucleotides. Upon completion MsbA translocates core-Kdo-lipid A from the inner leaflet to the outer leaflet

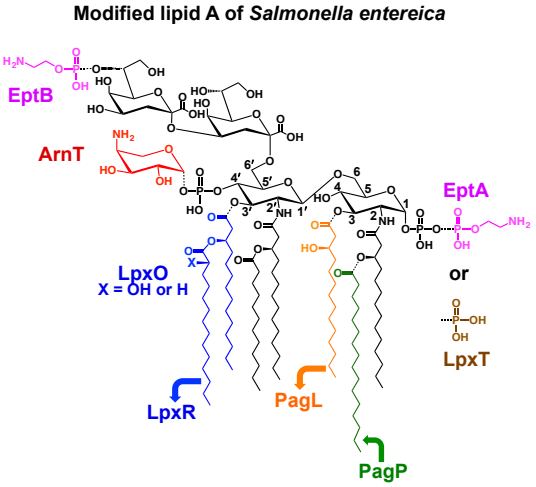
of the IM, in an ATP dependent manner (Figure 1.4). As translocation from the inner to outer leaflet appears essential, MsbA has potential as a target for antimicrobial development. Characterization of MsbA in vitro also suggests it may serve as a capable multidrug efflux pump[76–78]. No compounds that specifically block MsbA in vivo have been reported.

In the periplasmic region, O-antigen is ligated onto core-Kdo-lipid A to form fully mature LPS[49] (not shown in Figure 1.4). In *E. coli,* transport of LPS onto the surface of the OM requires a 7-protein complex (Figure 1.4). An IM ABC transporter, LptBFG is required for LPS extraction from the periplasmic leaflet[79–81]. This protein complex interfaces with an LPS binding protein, LptC, to "transfer" LPS from the periplasmic leaflet of the IM to a periplasmic transporter, LptA[80,82,83]. LptA forms a periplasm-spanning filament bridging the inner to the OM[84,85]. At the OM, LptA transfers LPS to the third and final LPS binding protein, LptE, wherein LptD and LptE form a "plug and barrel" structure for final deposition of LPS into the outer leaflet of the OM[86,87]. How large LPS structures are threaded through the LptDE complex remains an interesting question. This physical constraint is also seen with OM β–barrel assembly machinery (BAM) that helps properly fold OM β–barrel proteins. Often substrate β-barrel proteins are larger than the β–barrel component of BAM complexes. To allow intramembranous transit of bulkier substrate molecules, LptDE, BAM, and related complexs are suggested to contain a gated lateral opening as supported by experimental evidence[88–90]. Since components of the Lpt machinery are on the bacterial surface, and appear essential for growth, LPS transport is a viable target in the development of new antibacterial agents.

Compounds that specifically target LptD in *P. aeruginosa*[91], or LptB of *E. coli* have already been identified[92–94].

1.2 The Biological Significance Of Chemically Altered KDO-Lipid A Domains

Most bacteria modify the basic Kdo-lipid A structure to optimize OM integrity for survival in a given niche. For example, pathogens tailor OM remodeling in response to environmental cues so as to enhance virulence and host colonization. In general, Kdo-lipid A modifications involve changes in the number or composition of acyl chains, phosphate groups or any of various covalently attached functional groups[31,33]. The Kdo-lipid A domain from *Salmonella enterica* serovar Typhimurium serves as an exemplary model of the full range of modifications one organism may contain (Figure 1.5). For more than a decade most efforts have focused on determining structures of Kdo-lipid A domains from pathogenic bacteria[31,33]. In looking towards future research, as monophosphoryl lipid A is the last FDA approved adjuvant, a detailed understanding of Kdo-lipid A domains and their impact on priming the immune system, will certainly contribute to improved vaccine formulations[95]. Similarly, as interest in the human microbiome has reached a fever pitch in recent years, the importance of Kdo-lipid A modifications in maintaining host-bacteria commensal interactions will continue to be an active area of research[32,96–98].



Enzyme	Activity	Active site topology
ArnT	Transfers L-4-aminoarabinose to lipid A phosphate groups	Periplasm
EptA	Transfers phosphoethanolamine to lipid A phosphate groups	Periplasm
EptB	Transfers phosphoethanolamine to outer Kdo sugar of inner core	Periplasm
LpxO	Hydroxylates the 3′-fatty acid myristate	Cytoplasm
LpxR	Removes acyl chains from 3′-position	Outer membrane
LpxT	Phosphorylates lipid A forming a pyrophosphate group at the 1-position	Periplasm
PagL	Removes acyl chain from 3-position	Outer membrane
PagP	Palmitoylates (C16:0) the primary linked acyl chain at 2-position	Outer membrane

Figure 1.5 Chemical modifications of the Kdo$_2$-lipid A domain of *Salmonella enterica* serovar Typhimurium.

The Kdo$_2$-lipid A domain of *Salmonella* spp. can be highly modified by highly regulated enzymatic machinery. Arrows indicate either the addition or the removal of acyl chains. Numbers indicate positions on the disaccharide portion of lipid A. Addition of free-amine-containing residues, 4-amino-4-deoxy-l-aminoarabinose (l-4-aminoarabionse) and phosphoethanolamine, or the fatty acid palmitate promotes resistance to antimicrobial peptides. The removal of acyl chains is associated with reduction of endotoxicity and TLR4/MD-2 activation.

1.2.1 Recent observations of chemically modified Kdo-Lipid A domains and their importance in commensal interactions between eukaryotic hosts and bacteria.

Chemical modifications to LPS typically involve smaller chemical alterations of less than 200 daltons. Larger chemical substitutions have been discovered (>500 Da), such as the recently observed covalent attachment of hopanoids to very long chain fatty acids of the Kdo-lipid A domain of Bradyrhizobia LPS (Figure 1.6; far left structure)[98]. Biophysical experiments in liposomes show that hopanoid linked lipid A helps stabilize model asymmetric bilayers, where in nature hopanoid modification is necessary for the

favorable environmental association with the bacteria's leguminous symbiont. Interestingly, strains of *Bradyrhizobium* that were deficient for general hopanoid biosynthesis were less able to maintain "healthy" symbiosis with host *Aeschynomene evenia* legumes[98]. A better understanding of how hopanoid-lipid A impacts symbiosis, could inform agricultural bioengineers interested in reducing our dependence on the environmentally and economically unsustainable use of chemical fertilizers.

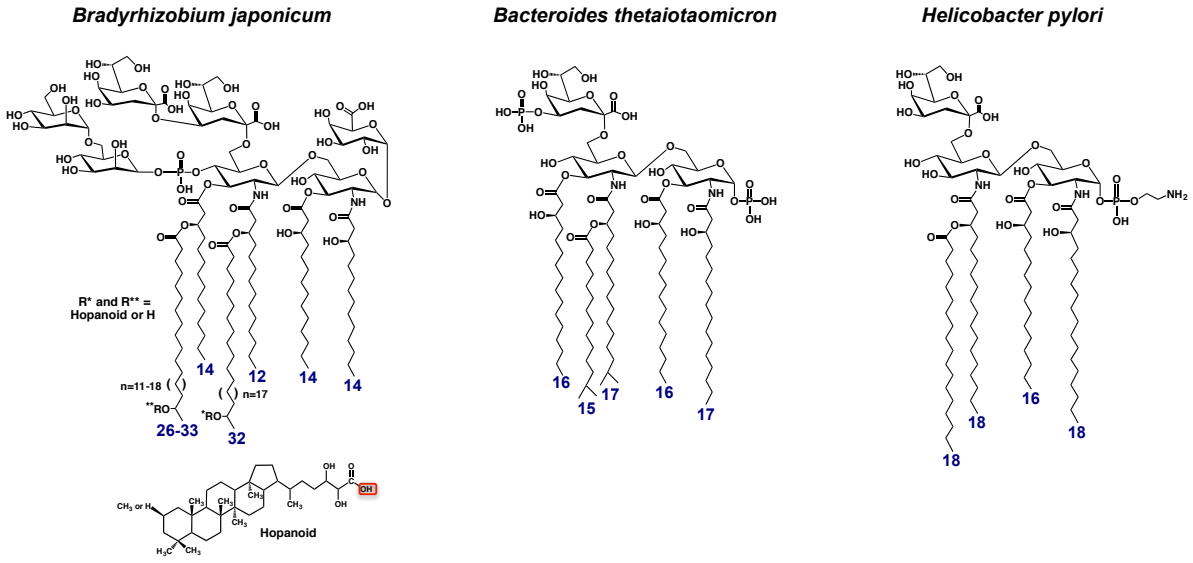


Figure 1.6 Kdo-lipid A domains of *Bradyrhizobium japonicum*, *Bacteroides thetaiotaomicron*, and *Helicobacter pylori*.

The major Kdo-lipid A species presented on the surface of each organism is shown. Numbers indicate the length of acyl chains. For *B. japonicum,* the red box around the hopanoid hydroxyl group indicates the attachment site to indicated acyl chains on the Kdo-lipid A domain of the organism.

Provocative studies reveal the dynamic relationship between surface features of human commensals and host immunity, especially within the gastro-intestinal (GI) tract, which contains over 100 trillion bacteria. Mechanisms that limit over-stimulation of pro-inflammatory receptor mediated pathways by resident GI commensals are of prime interest. One such mechanism involves only sensing bacteria that invade past the luminal epithelial barrier, where immune cells are more abundant. Expression of TLR4/MD-2, the pro-inflammatory receptor that specifically binds to and is activated by Kdo-lipid A, is polarized to the baso-lateral surface of associated enterocytes[99], limiting unnecessary inflammatory responses to bulk commensals in the GI lumen. Similarly, Paneth cells resident to intestinal crypts direct antimicrobial responses in response to TLR4/MD-2 activation[100]. Some commensals appear to disguise their surface, like *Bacteriodetes thetaiotaomicron*, which express a lipid A phosphatase (LpxF). This modification simultaneously provides resistance against host luminal antimicrobial peptides and produces a lipid A molecule that weakly stimulates TLR4/MD-2[96] (Figure 1.6; middle structure). Similar remodeling of the cell surface may be a general commensal strategy as a modified lipid A species, dephosphorylated and under-acylated, is displayed on the surface of the obligate, gastric bacterium *Helicobacter pylori*[97] (Figure 1.6; far right structure). Further understanding of microbial-host homeostasis, in the GI tract, will inform therapeutic treatment options for inflammatory bowel syndromes that arise through diet, pathogen attack, or chemotherapeutic/antibiotic protocols. Other human niches such as the mouth, skin, and respiratory tract, prone to disease states consequent

from perturbations in "healthy" microbiome populations also warrant deeper investigation[101,102].

Perhaps one of the best-appreciated models of host-bacterial symbiosis is provided by the exclusive recruitment of bioluminescent *Vibrio fischeri* by the light organ of the Hawaiian bobtail squid (*Euprymna scolopes)*. In juvenile squid, development of light organ ciliated epithelial cells requires bacterial components, peptidoglycan tracheal cytotoxin and LPS[103]. These *V. fischeri* surface components are sensed through pattern recognition receptors of the innate immune system, and are essential for morphogenesis of "gut" epithelial cells into tissues of the mature light organ. After recruitment of bioluminescent *V. fischeri* into a mature light organ, over-activation of the squid immune system (TLR4/MD-2 pathway) is thought to be avoided through enzymatic detoxification of *V. fischeri* LPS by a secreted host alkaline phosphatase[104]. Whereas human gut commensals *Helicobacter pylori* and *Bacteriodetes thetaiotaomicron* encode their own phosphatases, the squid produces the phosphatase that modifies *V. fischeri* lipid A. Unique to Kdo-lipid A domains of *V. fischeri* is the addition of specific phospholipid chemical moieties: phosphatidic acid, lyso-phosphatidic acid, and phosphoglycerol that have been observed in mass-spectral analysis of lipid A isolated from laboratory grown *V. fischeri*[32] (Figure 1.7). This first-of-its-kind Kdo-lipid A modification is of unknown biological significance and the synthetic pathway awaits elucidation. Of important note, these modifications may be a general feature of the marine bacterial genus *Vibrio,* as *Vibrio cholerae* produces a minor population of this chemotype under specific growth conditions (Chapter 5).

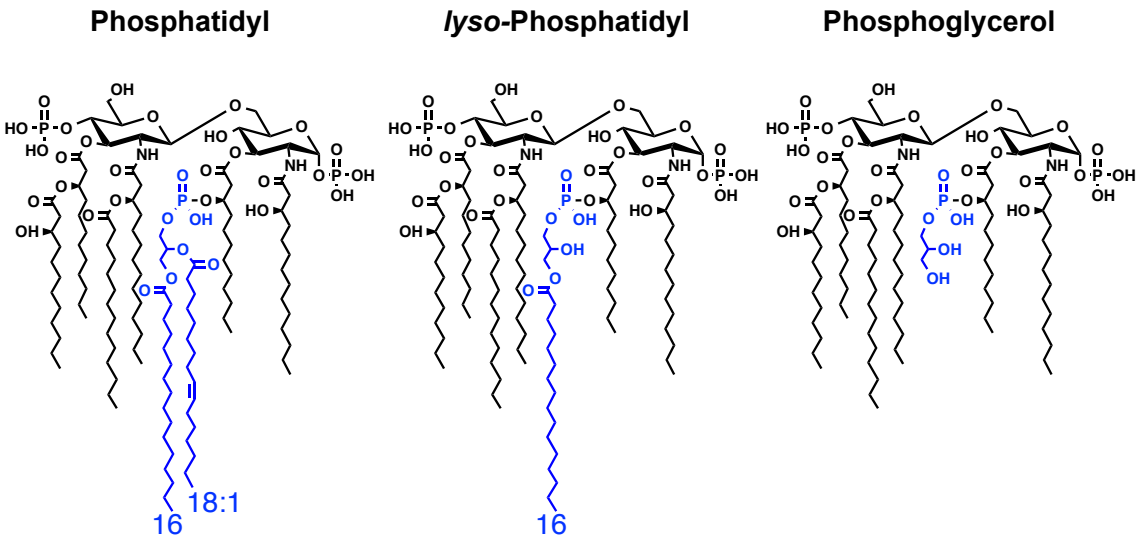


Figure 1.7 Lipid A species isolated from the *Euprymenes scolopes* symbiont, *Vibrio fischeri*.

Lipid A species isolated from *V. fischeri* grown in defined aquatic seawater medium. Type of chemical modification appears above each lipid A spcies. These same structures have been observed as minor molecular species in mass-spectrometric analysis of lipid A material isolated from *V. cholerae* under specific growth conditions. Acyl chain length shown for phosphatidyl and lyso-phosphatidyl are representative of the major chain length observed, other lengths and degrees of unsaturation are also observed, albeit in more minor amounts.

1.2.2 Surface remodeling of Kdo-Lipid A domains and other outer membrane components promotes resistance to cationic antimicrobial peptides (CAMPs)

Clinical use of polymyxins is an antibacterial strategy of last resort in treating infections with multi-drug resistant Gram-negative bacteria such as *Acinetobacter baumanii*, *Pseudomonas aeruginosa*, *Klebsiella spp*. and *E. coli spp*.[105–107]. Polymyxins are cationic antimicrobial peptides (CAMPs) comprised of a cationic/hydrophobic cyclic

decapeptide linked to a fatty acyl chain (Figure 1.8; top structure panel A). Like host defensins, CAMPs disrupt the Gram-negative cell envelope by associating with the anionic lipid A membrane anchor of the major cell-surface molecule lipopolysaccharide (LPS) as well as acidic glycerophospholipids[107]. Detergent-like aggregates of membrane-bound polymyxins deplete lipids of associated membranes (Figure 1.8; panel B), disrupting vital bacterial processes, and at high concentrations leads to fatal cell lysis[107]. Accumulating evidence suggests that many bacteria have evolved elegant resistance strategies toward polymyxins and related families of CAMPs[107,108]. Of recent interest is the *mcr-1* allele observed in strains of *E. coli*, identified first from a pig farm in China and later from uropathogenic isolates obtained from a human patient in the United States[109,110]. It is the first identified colistin (polymyxin E) resistance gene transmitted via plasmid-mediated horizontal gene transfer. A comprehensive understanding of the molecular basis for resistance to CAMPs is paramount in designing new, or modifying current, antibiotics and in developing improved clinical regimens with existing drugs.

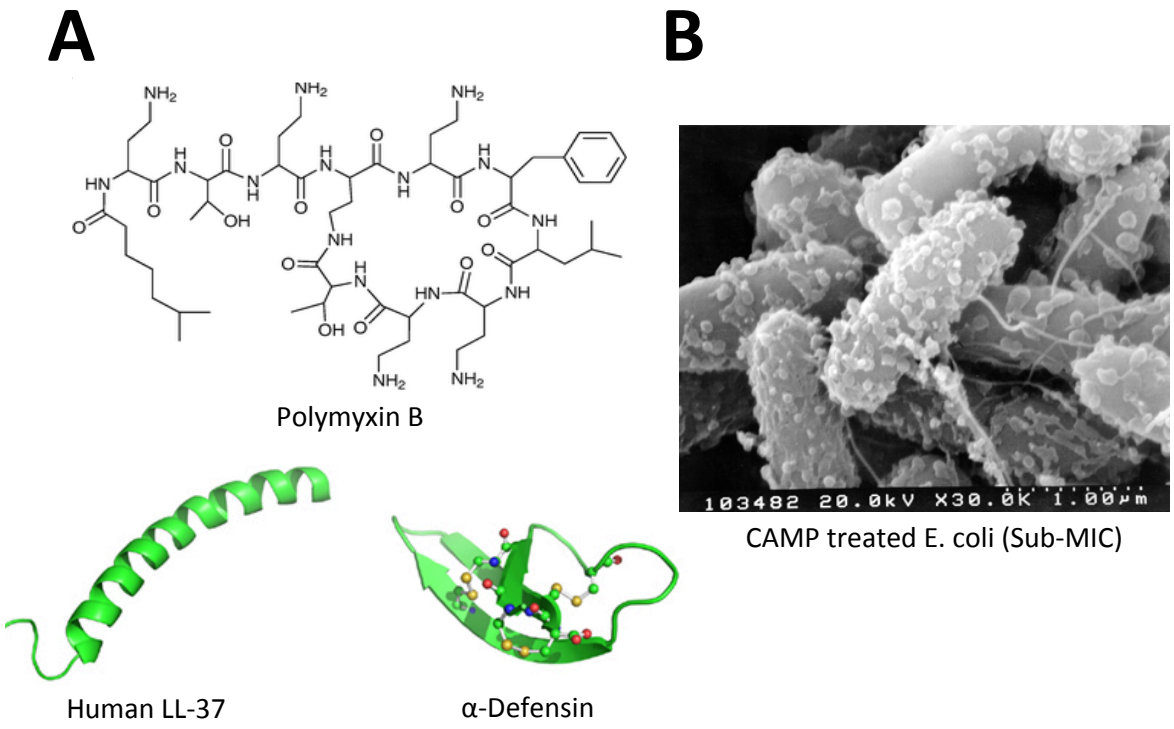


Figure 1.8 Membrane disruption of Gram-negative envelope by cationic antimicrobial peptides (CAMPs).

(A) Structures of representative cationic antimicrobial peptides found in nature. (B) Transmission electron microscopy of *E. coli* membrane disruption, resultant from polymyxin treatment below the minimum inhibitory concentration (sub-MIC).

In general, to evade CAMP-mediated killing, Gram-negative and Gram-positive bacteria have evolved similar strategies aimed at neutralization of the net negative charge of major cell-surface molecules. Phosphatidylglycerols aminoacylated with L-lysine, L-alanine, or D-alanine were recently discovered and confer resistance toward CAMPs in a few species of both Gram-type bacteria (e.g., *Staphylococcus aureus* and *P. aeruginosa*)[111] (Figure 1.9; center purple panel). Gram-positive bacteria specifically neutralize their cell walls by transferring D-alanine to surface teichoic acids, the long

poly-phosphoribitol or poly-phosphoglycerol chains linked to the glycerophospholipids or *N*-acetylmuramic acid groups of surface exposed peptidoglycan[112] (Figure 1.9; left blue panel); whereas Gram-negative bacteria transfer phosphoethanolamine or aminoarabinose to phosphate groups on the lipid A domain of LPS[113,114] (See Figure 1.5; EptA or ArnT respectively). The *mcr-1* allele mentioned previously is a phosphoethanolamine transferase that modifies lipid A[109]. The major focus of this thesis is on the aminoacyl esterification of glycine or diglycine to lipid A of the Gram-negative pathogen *Vibrio cholerae*[115] (Figure 1.9; right red panel). Lipid A glycinylation is a unique Gram-negative strategy necessary for resistance to CAMPs, which resembles features of the Gram-positive pathway for modifying teichoic acids with D-alanine.

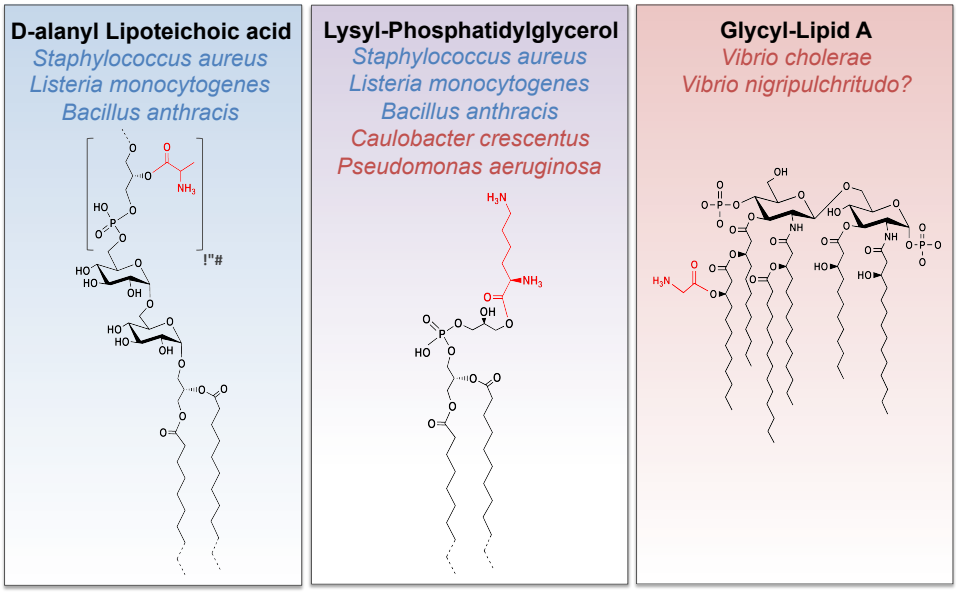


Figure 1.9 Aminoacyl modifications promote CAMP resistance in bacteria.

Examples of lipid species with esterified amino acids from either gram-type bacteria; representative Gram-positives listed in blue text, representative Gram-negative organisms listed in red text.

1.3 The Current pandemic strain of *V. cholerae* resists CAMPs due to a Glycine modified Kdo-lipid A Domain

The gram-negative bacterium *Vibrio cholerae* lives in aquatic reservoirs and enters the human host via an oral route of infection through contaminated food or water. Of more than 200 serotypes of *V. cholerae* found in nature, O1 and O139 are the primary serotypes associated with human disease, where O1 strains are responsible for multiple pandemics recorded throughout epidemiological history[116]. After multiplication and colonization of the small intestine, the bacteria cause a profuse diarrhea characteristic of cholera disease[116]. Cholera toxin, the primary effector responsible for diarrheal pathology, is a heat labile enterotoxin internalized into affected enterocytes in a receptor mediated manner that promotes massive efflux of chloride, and thus water, into the intestinal lumen and also disrupts the barrier function of intestinal epithelial cells[116–118]. The human gastrointestinal tract presents different barriers to impede colonization by pathogenic microorganisms, such as the gastric acid barrier of the stomach, bile acids and cationic antimicrobial peptides[119–121].

1.3.1 Discovery of glycine modified Kdo-Lipid A in *V. cholerae* biotype El Tor

Unlike the O1 Classical biotype of *V. cholerae*, responsible for the first six cholera pandemics, the current pandemic strain of the O1 El Tor biotype is resistant to polymyxin B[122]. A number of other phenotypes distinguish biotype El Tor from Classical pandemic strains: lytic phage susceptibilities, chicken hemocyte agglutination, acetoin

production, and variation in sequence or expression of virulence alleles[123,124]. El Tor strains are responsible for nearly all of 3-5 million cases and 100,000-120,000 deaths associated with *V. cholerae* infection every year (World Health Organization). Previously the Trent lab identified a lipid A late acyl-transferase homolog (Figure 1.4; LpxL/M), renamed LpxN (Vc0212), as important for polymyxin B resistance in El Tor biotype *V. cholerae*[125,126]. LpxN uniquely adds a hydroxyl containing acyl chain. A transposon library screen in biotype El Tor for additional genes that confer resistance to polymyxin B revealed an operon, AlmE (Vc1579), AlmF (Vc1578), and AlmG (Vc1577), is required for polymyxin B resistance in El Tor biotype *V. cholerae* O1[115]. Together, these proteins orchestrate the modification of *V. cholerae* lipid A with glycine or diglycine, where the LpxN transferred hydroxyl acyl chain serves as the site of amino acid attachment[115]. A hypothetical molecular model was proposed (Figure 1.10), based on a four gene pathway that aminoacylates cell surface teichoic acids of Gram-positives[115].

According to the proposed model, AlmE activates glycine through adenylation and then thioesterifies glycine to a 4′-phosphopantetheine group of the proposed carrier protein AlmF (Chapters 2&3). AlmG then transfers the glycine substrate from carrier AlmF to the lipid A anchor of LPS (Chapter 4). The unique hydroxylauryl chain observed in lipid A species from El Tor biotype *V. cholerae* serves as the site of glycine addition[125], thus AlmEFG is the only known lipid A charge remodeling system that does not modify the 1- or 4′- phosphate groups of lipid A. Classical biotype *V. cholerae*, which does not attach glycine to lipid A, and is polymyxin-sensitive, encodes a truncated AlmF lacking the phosphopantetheinylation site required for carrier protein functionality[115] (Chapter 2).

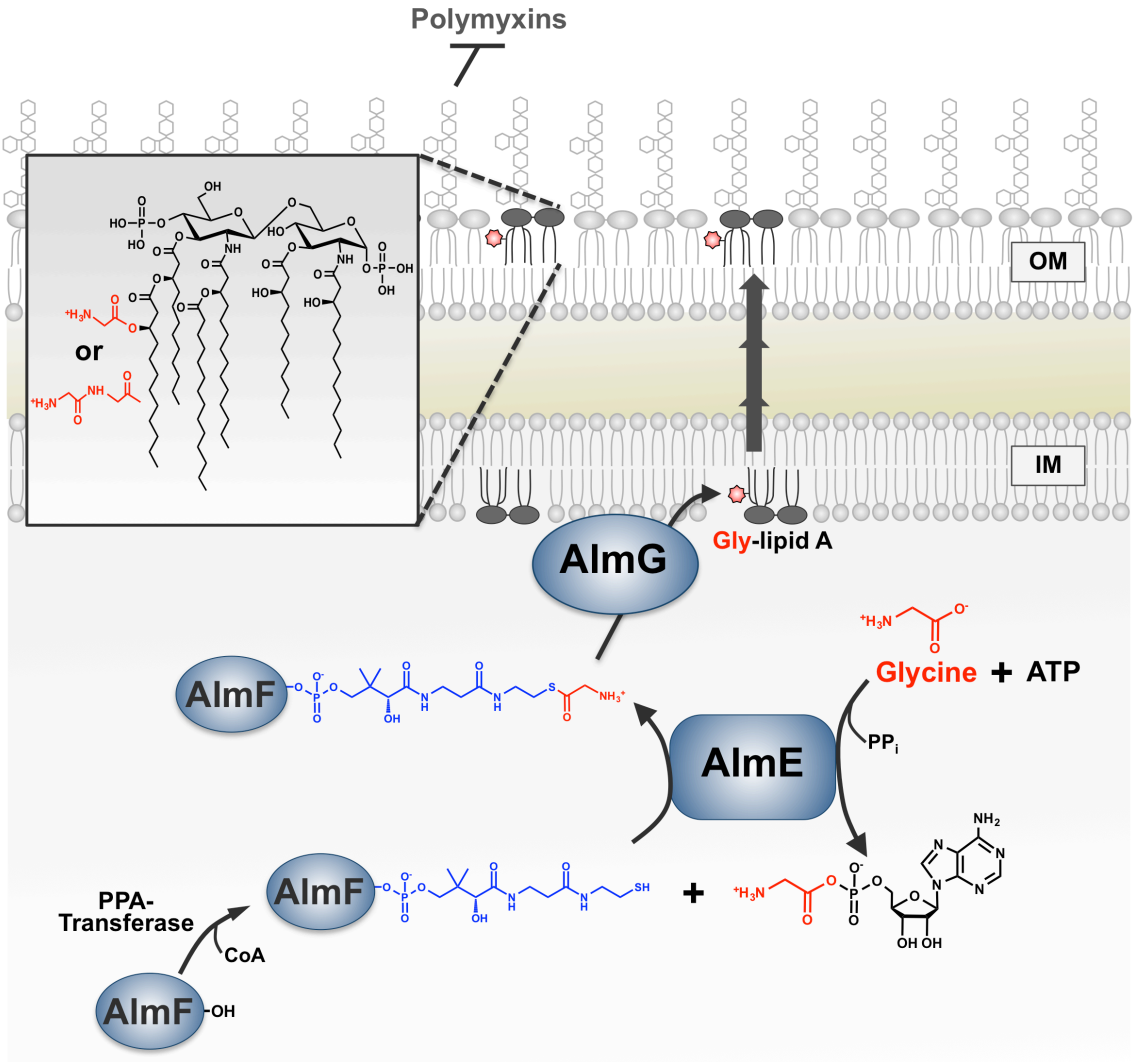


Figure 1.10 Proposed model for the lipid A glycinylation pathway in *V. cholerae*.

AlmE (Vc1579) generates glycyl-AMP and pyrophosphate from glycine and ATP. Glycine is then ligated onto carrier protein holo-AlmF (Vc1578), with concomitant release of AMP. Holo-AlmF is generated after the 4′-phosphopantetheinyl (Ppant) moiety of coenzyme A is transferred onto apo-AlmF by a phosphopantetheinyltransferase. At the inner membrane, glycyl-AlmF then serves as the aminoacyl donor to AlmG (Vc1577), which esterifies glycine onto the secondary hydroxylauryl acyl chain of *V. cholerae* hexa-acylated lipid A. Glycine-modified lipid A is then transported to the bacterial surface to provide resistance against antimicrobial peptides such as polymyxin.

1.3.2 A two component system (VprAB) regulates expression of the *alm* operon, responsible for glycinylating Kdo-lipid A domains in El Tor *V. cholerae*

In addition to LpxN and AlmEFG, a screen of the O1 El Tor non-redundant transposon library for mutants with increased sensitivity to polymyxin B uncovered a two-component system VC1320-19[127]. Renamed VprA-VprB (***V****ibrio* **p**olymyxin resistance **r**egulator) VC1320-19 regulates expression of the *alm* operon (Figure 1.11). Mutant strains lacking VprB, the histidine kinase sensor protein, or VprA, the cognate phosphorylated response regulator, synthesize hexa-acylated lipid A deficient of glycine modification and are highly sensitive to polymyxin. Mutagenesis of conserved amino acids on VprB or VprA were shown essential for the VprAB phosphotransfer signaling cascade[127]. Phosphorylated VprA was further verified to bind a promoter sequence upstream of the *alm* operon, and shown to be a positive regulator of AlmEFG expression, as well as a positive auto-regulator of *vprA* itself[127]. Up-regulation of VprAB, and VprAB-dependent *almEFG* expression, occurs in the presence of sub-lethal polymyxin concentrations[127]. This ensures high levels of glycine modified Kdo-lipid A to promote *V. cholerae* survival in CAMP challenged growth conditions. Finally, mutants deficient in lipid A glycine modification, either through disruption of *vprA* or *almE*, display significantly reduced fitness for intestinal colonization in the suckling mouse model of *V. cholerae* infection[127].

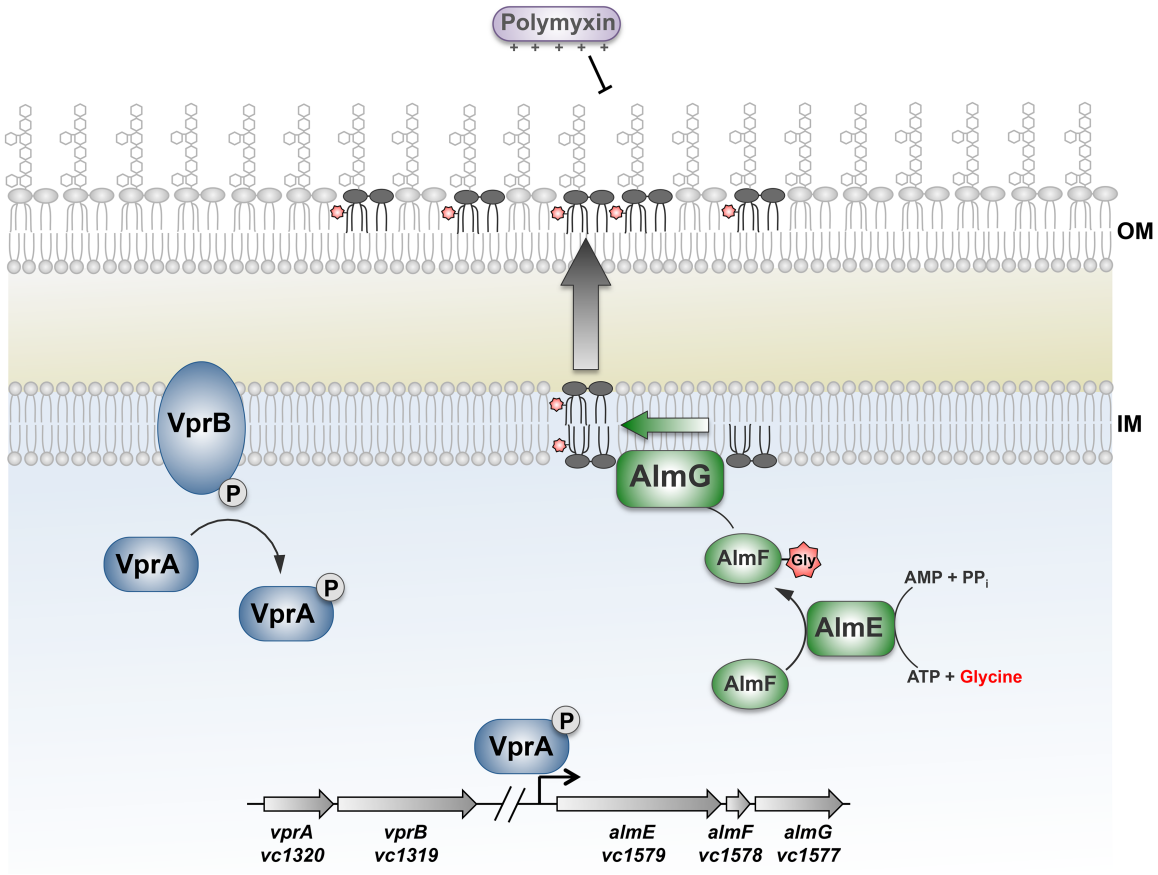


Figure 1.11 The VprAB two component system regulates AlmEFG expression

A model illustrating the transcriptional regulation of glycinyl lipid A modification in *V. cholerae* El Tor. VprB functions as a sensor histidine kinase and activates the transcriptional regulator VprA. Phosphorylated VprA directly promotes expression of the *almEFG* operon. Presumably, glycine addition increases the net positive charge of the bacterial surface providing resistance to the cationic peptide polymyxin. The addition of core oligosaccharide and O-antigen are not shown for simplicity. OM: outer membrane; IM: inner membrane; P: phosphate group.

1.3.3 The importance of glycine modified Kdo-Lipid A species in host colonization

Reduced intestinal colonization of *vprA*, *almE* and *almG* mutants relative to wild-type El Tor strains, suggests that glycine lipid A modification has a protective role for

bacterial membranes against stressful conditions during early stages of *V. cholerae* infection[127]. A pair of studies corroborate with this hypothesis[128,129]. In *V. cholerae* isolated from infant mouse small intestine, the mRNA transcript of the *almEFG* operon is upregulated >10-fold[129]. Moreover, microarray-based comparative gene expression analysis showed a 3.2-fold increase (p-value $1.6x10^{-6}$) in *almE* levels between early- and late-stage infections of humans by *V. cholerae* [128]. In that report, *almE* was one of 42 statistically significant, differentially-regulated genes to be identified. As evidence accumulates to demonstrate the importance of the AlmEFG system during infection, detailed mechanistic studies of AlmEFG have the potential to inform further research on therapeutic drugs and vaccines targeting *V. cholerae* infections. Moreover as *Vibrio* cells might adjust synthesis of VprAB two-component system quickly and efficiently in response to local stressors and host immune factors, further elucidation of *in vivo* signals and host microenvironments to which VprAB responds, will lead to a more sophisticated understanding of glycine modified lipid A on *V. cholerae* pathogenesis.

Chapter 2: Assembly of a Glycine Carrier Protein AlmF[2]

2.1 Introduction

Acyl carrier proteins (ACPs) are important metabolites that participate in primary and secondary metabolic pathways within all domains of life. ACPs share key structural features: three-helix bundle architecture and a post-translational phosphopantetheine modification that serves as the functional group for "carrier" activity. Conventional classification of ACP domains is based on their participation in type I or type II synthases, which biosynthesize compounds like polyketides, non-ribosomal peptides, vitamins or fatty acids. Type I synthases are large, single stranded proteins, with multiple activity domains covalently linked in series. Type II systems assemble from discrete proteins that are not covalently linked. With some exceptions, bacteria use contiguous type I synthases for non-essential processes like poly-ketide production and non-ribosomal peptides, and assembled type II synthases for essential processes like fatty acid biosynthesis. In contrast with bacteria, humans use contiguous type I synthases for fatty acid production[1], so the individual components of bacterial type II fatty acid synthases have been historically viewed as candidate targets in antibacterial design efforts[2–4].

The translated, primary protein sequence of an ACP is referred to as the apo-form (apo-ACP). A 4′-phosphopantetheinyltransferase converts apo-ACPs to holo-ACPs via post-translational attachment of a phosphopantetheine carrier moiety. Carrier

[2] Portions of this chapter have been previously published. Henderson, J.C., Fage, C.D., Cannon, J.R., Brodbelt, J.S., Keatinge-Clay, A.T., and Trent, M.S. Antimicrobial peptide resistance of Vibrio cholerae results from an LPS modification pathway related to non-ribosomal peptide synthetases. *ACS Chem Biol*. **86**, 2138-45 2014. JCH along with equal contribution from co-author CDF contributed data, wrote, and edited this report.

functionality is imparted by the terminal thiol of a phosphopantetheine molecule, that enables ACPs to shuttle substrate molecules as covalently-attached thioesters[5]. A 4′-phosphopanteteinyltransferase appends the phosphopantetheine moiety from Coenzyme A directly onto a conserved, nucleophilic serine on acceptor apo-ACP substrates[6]. Bacterial genomes contain two sub-types of a more general 4′-phosphopantetheinyltransferase superfamily. These are the AcpS subtype, which activates ACPs necessary for primary metabolism like fatty acid biosynthesis, and the surfactin synthetase-activating enzyme (Sfp) subtype, which activates carrier proteins involved in secondary metabolism such as polyketide and vitamin biosynthesis[7].

Most structures deposited in the RCSB protein data bank of 4′-phosphopantetheinyltransferase superfamily members are bacterial proteins (~80%), with a few fungal and human proteins as well (~20%). Structural features strongly inform our understanding of bacterial 4′-phosphopantetheinyltransferase sub-types[3,4]. Members of the AcpS subtype are generally smaller, 140 amino acids in length, form a distinctive α/β fold, and require multimerization into a trimer for 4′-phosphopantetheinyltransferase activity[8,9]. Sfp subtype enzymes are typically twice the size of AcpS members, containing two α/β fold domains, each of which is structurally homologous to an individual AcpS monomer[10,11]. This suggests that Sfp subtypes might have evolved by duplication of an AcpS ancestor. Sfp phosphopantetheinyltransferases do not require trimer formation for activity. The spectrum of acceptor ACP substrates for AcpS subtypes is narrow, whereas Sfp subtypes are generally more promiscuous modifying a broader pool of acceptor ACP

substrates. All members of the 4′-phosphopantetheinyltransferase family bind divalent cations near their respective catalytic active site, where magnesium is most frequently observed in structures and needed for optimum 4′-phosphopantetheinyltransferase activity in functional assays[8,10,11]. Residues required for coordinating divalent cations exhibit the highest degree of conservation among all members of the 4′-phosphopantetheinyltransferase superfamily[4].

In this chapter biochemical evidence supports the hypothesis that *vc1578* (AlmF), of the *vc1577-79* operon required for *V. cholerae* LPS glycinylation, is a glycine carrier protein. A structural model prediction for AlmF, obtained using the Phyre2 algorithm[13], suggests that AlmF forms the diagnostic three-helix bundle observed in all ACPs, with especial similarity to the crystal structure of Gram-positive D-alanyl carrier protein DltC (Figure 2.1; top). Multiple sequence alignment reveals Ser34 of AlmF to be the candidate site for 4′-phosphopantetheine modification. A conserved serine is the only universal residue found in all ACPs, usually found at the N-terminus of helix 2, as predicted for AlmF (Figure 2.1; Ser34 bottom). A 4′-phosphopantetheinyltransferase superfamily member in *V. cholerae*, *vc2457,* a member of the AcpS subtype is shown to encode the dedicated 4′-phosphopantetheinyltransferase that converts apo-AlmF to holo-AlmF in *V. cholerae*. Finally *vc1579* (AlmE), discussed in greater detail in Chapter 3, is shown to be necessary for glycine ligation onto holo-AlmF, and verifies that AlmF functions as a glycine carrier protein in the AlmEFG lipid A glycinylation pathway.

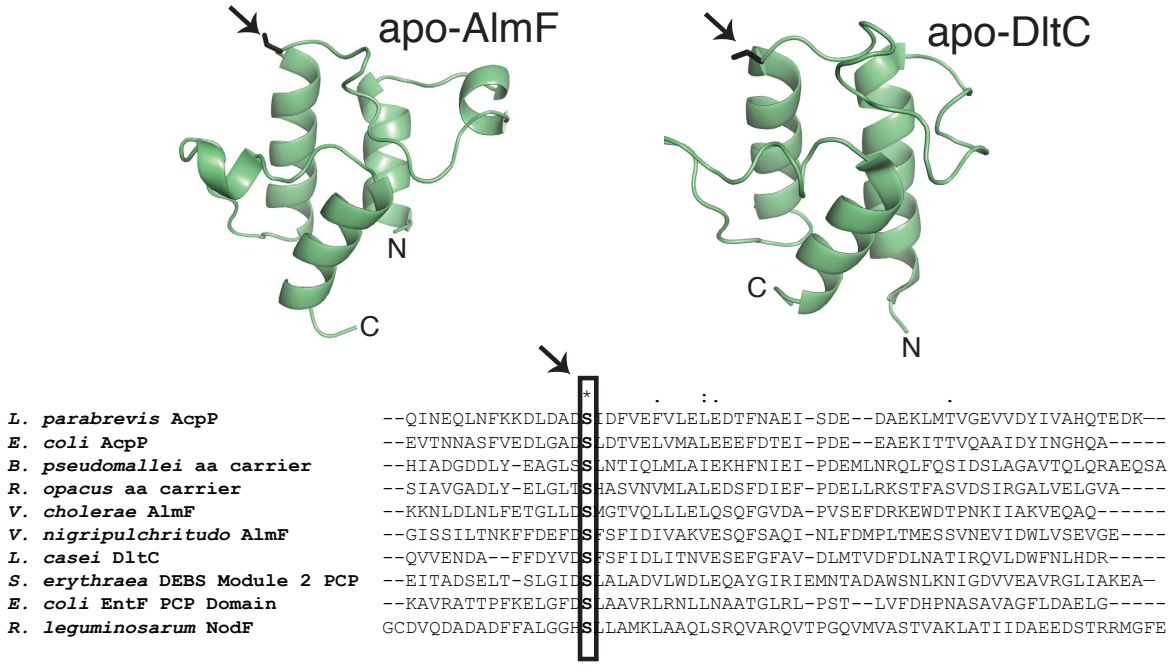


Figure 2.1: Prediction of AlmF Structure & Conserved Serine as Phosphopantetheine Modification Site

(Top) Soluble carrier proteins adopt a three-helix bundle architecture. Apo-AlmF from *V. cholerae* and apo-D-alanyl carrier protein (DltC) from *Lactobacillus casei* (PDB code 1HQB) are depicted as ribbon cartoons. The structural model for AlmF was created using coordinates generated by Phyre2. Both structures are shown without the 4′-phosphopantetheine prosthetic group. A conserved serine residue, represented as black sticks in the model, serves as the point of attachment for 4′-phosphopantetheine, indicated by an arrow. (Bottom) Multiple sequence alignment shows that even across a broad array of carrier protein domains, which can be single domains of multi-domain non-ribosomal peptide synthetases or individual soluble proteins, a conserved serine is readily identifiable (black box). Protein sequences were aligned using the Clustal Omega algorithm (12) and include the following*: Vibrio cholerae* Vc1578/AlmF (NP_231218), *Vibrio nigripulchritudo* A2555/AlmF homolog (YP_008621728), *Lactobacillus parabrevis* AcpP (WP_020088992), *Escherichia coli* AcpP (YP_489362), *Lactobacillus casei* DltC (AAB17659), *Saccharopolyspora erythraea* module 2 carrier domain of 6-Deoxyerythronolide B Synthase (DEBS; PDB code 2JU2), *Rhizobium leguminosarum* NodF (AFQ99236), *Burkholderia pseudomallei* aminoacyl carrier protein (AHK63940), *Rhodococcus opacus* aminoacyl carrier protein (AHK29595), truncated sequence of *E. coli EntF* peptidyl carrier protein domain (ZP_06938669). Symbols below alignment represent the degree of conservation of a particular residue: fully conserved (*), strong conservation (:), and weak conservation (.) respectively.

2.2 RESULTS

2.2.1 Vc2457 is an AcpS subtype 4′-phosphopantetheinyltransferase that converts apo-AlmF to holo-AlmF.

To determine if AlmF is a *bona fide* carrier protein, AlmF was co-expressed with either of the predicted phosphopantetheinyltransferases of *V. cholerae*, Vc2457 (AcpS subtype) or Vc0780 (Sfp subtype), and subsequently purified and examined for phosphopantetheinylation. To obtain quantities of AlmF suitable for analysis, protein production was optimized for overexpression in *E. coli* BL21(DE3)pLysS using the pQlink expression system and purified by a two-step protocol yielding >95% pure AlmF as evaluated by SDS-PAGE (Figure 2.2; e.g. lanes 2 and 3). To assess AlmF phosphospantetheinylation, holo- and apo- carrier proteins could be differentially resolved using destabilizing-urea PAGE, a non-denaturaing conformational gel[14–16]. AlmF co-expressed with Vc2457 (AcpS subtype) displayed a migration shift consistent with phosphopantetheine modification, when compared to AlmF expressed alone (Figure 2.3; compare lanes 1 & 3). Co-expression of AlmF with Vc0780 did not result in a migration shift relative to AlmF expressed alone, indicative that Vc0780 (Sfp subtype) does not phosphopantetheinylate AlmF (Figure 2.3; compare lanes 1 & 2).

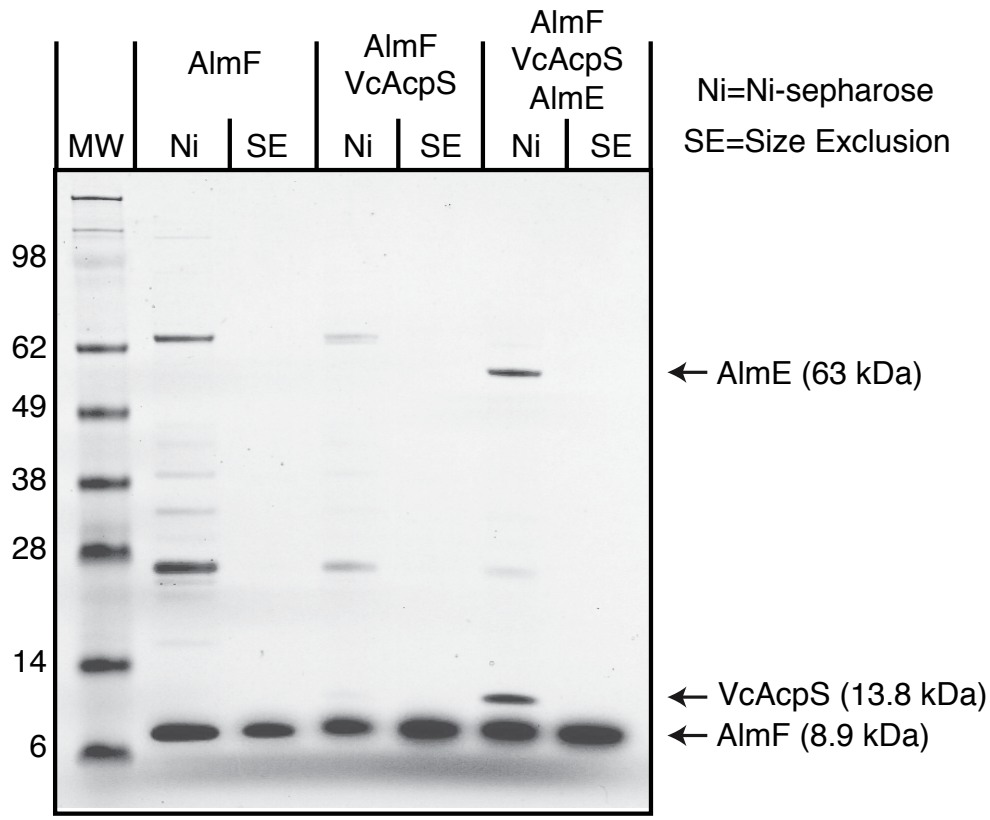


Figure 2.2: SDS-PAGE of Two-Step AlmF Purifications from Various Co-expression Strains Shows Highly Pure AlmF Yields

Due to low protein yields of recombinant $(His)_7$-AlmF, even when over-expressed in *E. coli*, a two-step purification of $(His)_7$-AlmF was required: Ni-sepharose (Ni) followed by size-exclusion (SE). An SDS-PAGE gel shows concentrated purification fractions from N-terminal $(His)_7$-AlmF over-expressed individually (AlmF), co-expressed with Vc2457, or co-expressed with both Vc2457 and AlmE. Purification fractions from AlmF co-expressed with Vc0780 is not shown. 10 μg of total protein was loaded per lane. Purified samples were further characterized by destabilizing-urea PAGE (Figure 2.3 & 2.4) or by UVPD-LC-MS/MS to assess phosphopantetheinylation and subsequent glycine modification status (Figure 2.5).

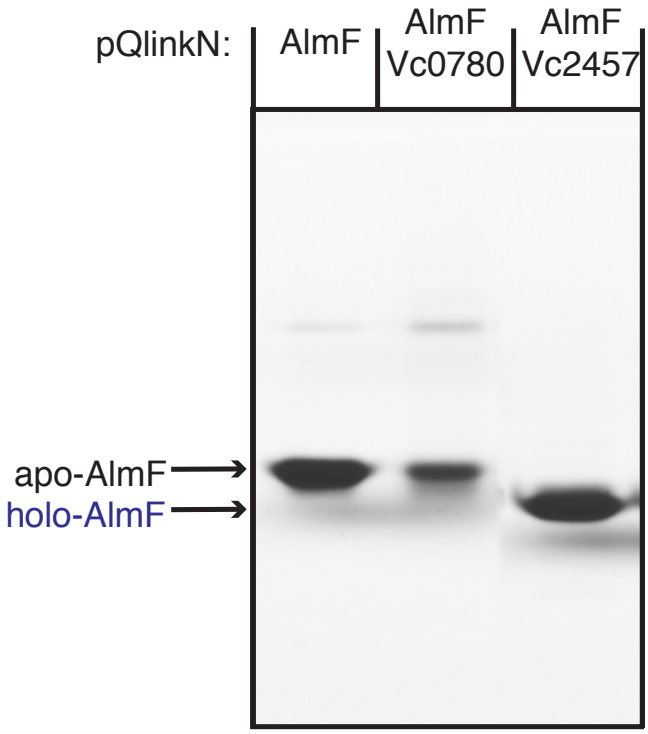


Figure 2.3: Destabilizing-Urea PAGE Resolves holo-AlmF from apo-AlmF, and Reveals Vc2457 is Phosphopantetheinylates AlmF *in vivo*

Co-expression of Vc2457 (VcAcpS) with AlmF (apo-AlmF) in *E. coli* results in the production of phosphopantetheinylated AlmF (holo-AlmF) as assessed by destabilizing urea-PAGE. Consistent with previous reports, holo-AlmF migrates further in the gel than apo-AlmF as a result of increased protein packing after phosphopantetheine modification. Vc0780 does not exhibit phosphopantetheinyltransferase activity towards AlmF.

Vc2457 is a homolog of AcpS from *E. coli*, with 56% identity and 74% sequence similarity. However even with this degree of similarity AcpS, present in our background *E. coli* AlmF overexpression strain, did not appear to convert apo-AlmF to holo-AlmF *in vivo* (Figure 2.3; lane 1). Previous reports have shown an increase in Coenzyme A pools can result in increased nonspecific 4′-phosphopantetheinyltransferase activity by *E. coli* AcpS [17]. To test this, growth media of *E. coli* strains expressing AlmF alone, was supplemented with millimolar concentrations of pantothenate, a condition previously shown to increase Coenzyme A substrate pools[17]. Supplementation with pantothenate resulted in nearly 50% conversion of apo-AlmF to holo-AlmF, the expected and likely result of non-specific *E. coli* AcpS activity (Figure 2.4; lanes 1 & 2). By contrast Vc2457 does not require pantothenate supplementation for complete phosphopantetheinylation of AlmF (Figure 2.4; lanes 3 & 4). This result convincingly demonstrates that Vc2457 is the dedicated, specific 4′-phosphopantetheinyltransferase for holo-AlmF production; a requisite step in carrier protein functionality.

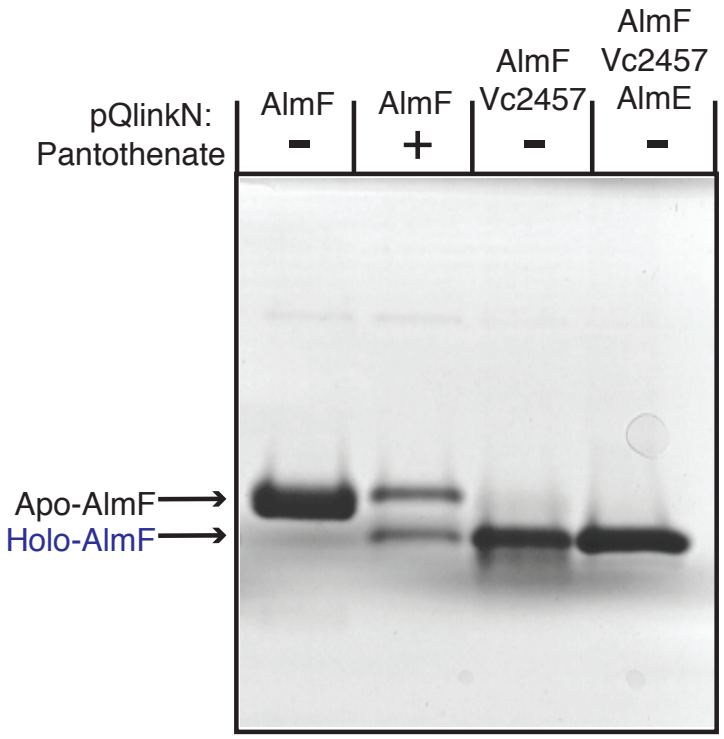


Figure 2.4: Destabilizing-Urea PAGE Shows Vc2457 is a Specific AlmF Phosphopantetheinyltransferase & Glycine-AlmF does not Differentially Migrate from holo-AlmF.

As Vc2457 is a homolog of AcpS, Ca^{+2}-pantothenate (1 mM) supplementation of the *E. coli* AlmF over-expression strain results in modest conversion of apo-AlmF to holo-AlmF. Thus, although the *E. coli* AcpS is not nearly as efficient as Vc2457 at holo-AlmF production, supplementation with pantothenate increases the substrate Coenzyme A pools to push the non-specific reaction forward (Fig. 1). Incubating holo-AlmF with AlmE did not result in a shift in the destabilizing urea-PAGE migration assay (far right lane). However, deeper analysis with UVPD-MS/MS shows that AlmE transfers a glycine to holo-AlmF (Fig. 3).

2.2.2 Fine structural analysis of holo-AlmF reveals Ser34 is the site of post-translational phosphopantetheine modification.

Unequivocal determination of AlmF phosphopantetheine-modification was assessed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) of trypsin-digested AlmF. Digests of unmodified (apo-AlmF; expressed alone) or phosphopantetheine-modified (holo-AlmF; co-expressed with Vc2457) AlmF were analyzed using the same purified AlmF material assessed by destabilizing urea-PAGE (from Figure 2.3). Three key peptides, each covering the first 41 amino acids of AlmF, were identified from LC–MS/MS analysis (Figure 2.5). The AlmF peptide (FFDEFDSFSFIDIVAK) containing the proposed site of modification, Ser34, eluted around 20 min in the apo-AlmF sample, whereas the same Ser34-containing peptide eluted 2 minutes later in the holo-AlmF sample (Figure 2.5; with a 114 *m/z* shift for the $[M+3H^+]^{3+}$ peptide). An increase in overall hydrophobicity is expected upon phosphopantetheine-addition, consistent with the delayed elution time observed for the holo-AlmF peptide relative to the same unmodified sequence. The electrospray ionization mass spectrum of the holo-AlmF Ser34-containing peptide showed a mass shift of ~340 Da relative to the apo-AlmF peptide, indicative of phosphopantetheine-modification (Figure 2.6). The sequence of the unmodified FFDEFDSFSFIDIVAK peptide ($[M+2H^+]^{2+}$) was readily confirmed based on the presence of the y_3 to y_{15} fragment ions (Figure 2.6, upper spectrum). MS/MS analysis of the peptide from the holo-AlmF sample resulted in a series of diagnostic *y* sequence ions as well as singly-charged product ions corresponding to pantetheine (Pant, *m/z* 261.1) and phosphopantetheine (Ppant, *m/z*

359.1) observed in the low-mass region (Figure 2.6, middle spectrum). Generation of pantetheine and phosphopantetheine product ions during MS/MS analysis of other phosphopantetheine-modified proteins has been previously reported[18].

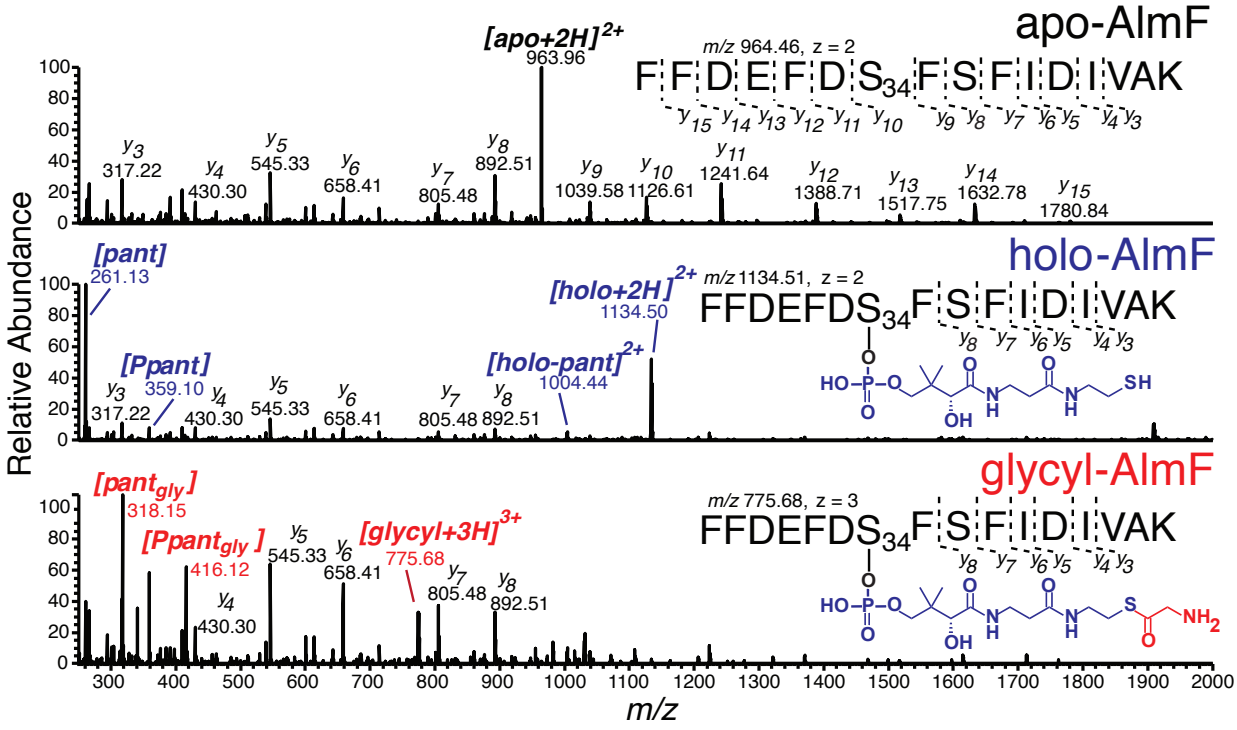


Figure 2.5: UVPD-MS/MS confirms AlmF phosphopantetheinylation by Vc2457 and subsequent glycylation by AlmE when co-expressed in *E. coli*.

Shown are MS^2 spectra resulting from 193 nm ultraviolet photodissociation of precursor ions corresponding to the trypsinized AlmF peptide FFDEFDS$_{34}$FSFIDIVAK: *(top)* doubly charged, unmodified (apo-AlmF; precursor *m/z* 964), *(middle)* doubly charged, phosphopantetheine (Ppant)-modified (holo-AlmF; precursor *m/z* 1134), and *(bottom)* triply charged, Ppant$_{Gly}$-modified (glycyl-AlmF; precursor *m/z* 776). These samples were obtained from purified AlmF, AlmF co-expressed with Vc2457, or AlmF co-expressed with Vc2457 and AlmE, respectively. Ser34 is the Ppant modification site as the *y*-type ion series through Ser36 (y_3-y_8) is conserved for both Ppant-modified AlmF *(middle)* and Ppant$_{Gly}$-modified AlmF *(bottom)*. Cleavage on either side of the phospho group results in diagnostic Ppant modification ions in the low mass region: *(middle)* pantetheine (Pant; *m/z* 261) and Ppant (*m/z* 359) or *(bottom)* Pant$_{Gly}$ (*m/z* 318) and Ppant$_{Gly}$ (*m/z* 416). The most abundant fragment ions are labeled.

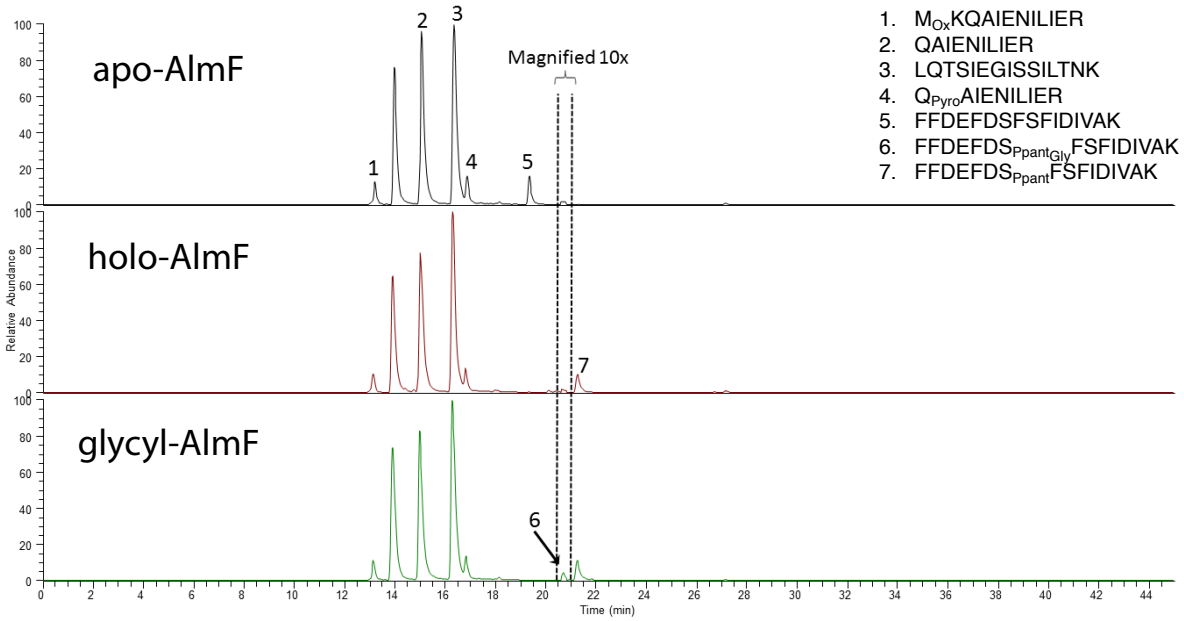


Figure 2.6: HPLC reverse phase chromatography of three independent trypsinized AlmF samples

Stacked comparison of the base peak chromatograms for HPLC reverse-phase chromatographic profiles of trypsinized AlmF samples as analyzed and structurally verified in Figure 2.6: unmodified (Apo-AlmF), phosphopantetheine-modified (holo-AlmF), and glycyl-phosphopantetheine-modified (glycyl-AlmF) samples. Peptide sequence assignments are shown in the legend in the upper right.

It should be noted that BLAST predicted Ser34 of AlmF to be the conserved site of phosphopantetheine attachment, however due to its close proximity, Ser36 could potentially serve as the phosphopantetheine modification site (FFDEFDS$_{34}$FS$_{36}$FIDIVAK). The observed *y*-type product ion fragments containing Ser36 could only be matched to unmodified masses and thus did not exhibit incorporation of phosphopantetheine (Figure 2.6, middle spectrum). As predicted, low-abundance *y*-type product ions (from y_{10} up to y_{15}) that implicated Ser34 as the sole

phosphopantetheine modified site were observed (Figure 2.7). Together with results from destabilizing urea-PAGE experiments, these findings convincingly support our structural homology-based prediction that AlmF is a phosphopantetheine-containing carrier protein, and that Vc2457 is specifically responsible for the conversion of apo-AlmF to holo-AlmF.

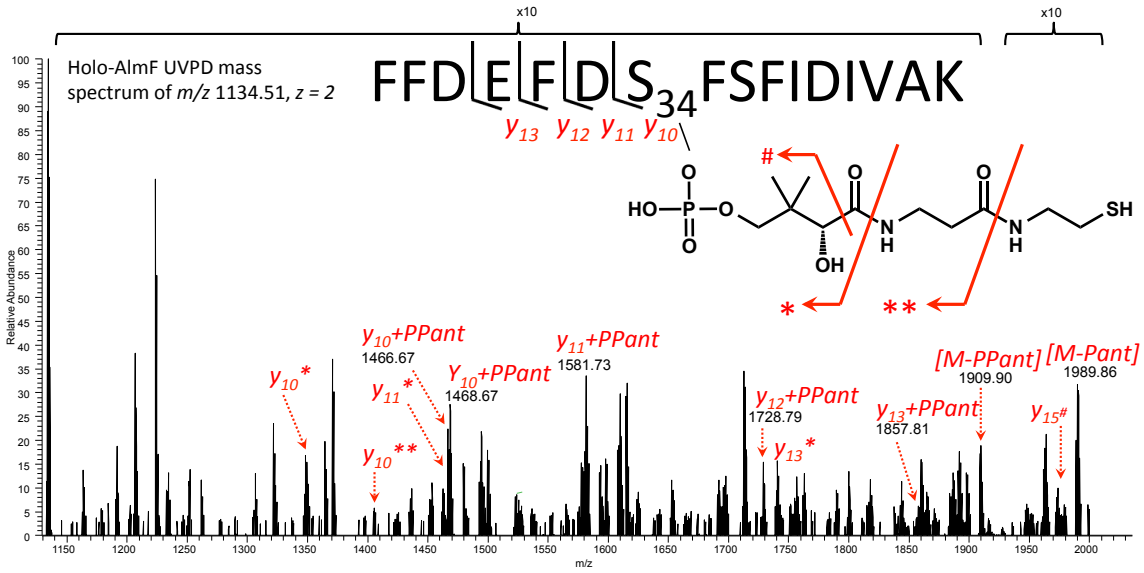


Figure 2.7: Expanded region of UVPD mass spectrum that confirms phosphopantetheinylation of AlmF.

This spectrum shows the higher-mass portion of the MS/MS spectrum from Figure 2.5 *(middle)* resulting from 193 nm ultraviolet photodissociation of the doubly-charged phosphopantetheinylated (Ppant) precursor (holo-AlmF; precursor m/z 1134). The ions containing the phosphopantetheine modification have been highlighted. 10x magnification was used for all ions with *m/z* values greater than the precursor ion with the exception of the ion of *m/z* 1909.90, which corresponds to the peptide after loss of the Ppant moiety. The asterisks and hash symbol are used to indicate the partial losses of the Ppant modification, which accompany the formation of certain *y* ions.

2.2.3 Vc1579 (AlmE) transfers glycine onto holo-AlmF to form the aminoacyl carrier protein Gly-AlmF.

To verify the role of AlmF as a glycyl carrier protein, we engineered holo-AlmF-producing *E. coli* strains to express Vc1579 (AlmE) the predicted aminoacyl-adenyltransferase/ carrier protein ligase. Unexpectedly, destabilizing urea-PAGE revealed no apparent shift in the migration of holo-AlmF when coexpressed with AlmE (Figure 2.4). However, deeper analysis of this sample by the same LC-MS/MS protocols described above demonstrated that the phosphopantetheine-containing peptide was modified with glycine (*m/z 775.8, z = 3,* Figure 2.5 bottom spectra) and, similarly, the glycyl-pantetheine and glycyl-phosphopantetheine singly-charged fragment ions show the 57-Da increase expected from glycine addition (Figure 2.5; bottom). The glycylated phosphopantetheine-containing peptide eluted 30 seconds earlier than the phosphopantetheine-containing peptide, consistent with the addition of a charged chemical moiety (Figure 2.6). By comparing the peak areas from extracted ion chromatograms of the doubly-charged phosphopantetheine-containing and glycylated phosphopantetheine-containing peptides (m/z 1134 vs. *m/z* 1163), we estimate that at minimum ~5% of the total holo-AlmF pool was glycylated *in vivo* when purified from *E. coli* co-expressing AlmE heterologously (Figure 2.8). However comparative ion abundance is qualitative, as differences in ionization efficiency between these versions of the AlmF peptide prohibit truly quantitative measurements. These results validate genetic and *in silico* predictions that AlmE activates glycine for transfer to the phosphopantetheine-containing carrier protein AlmF.

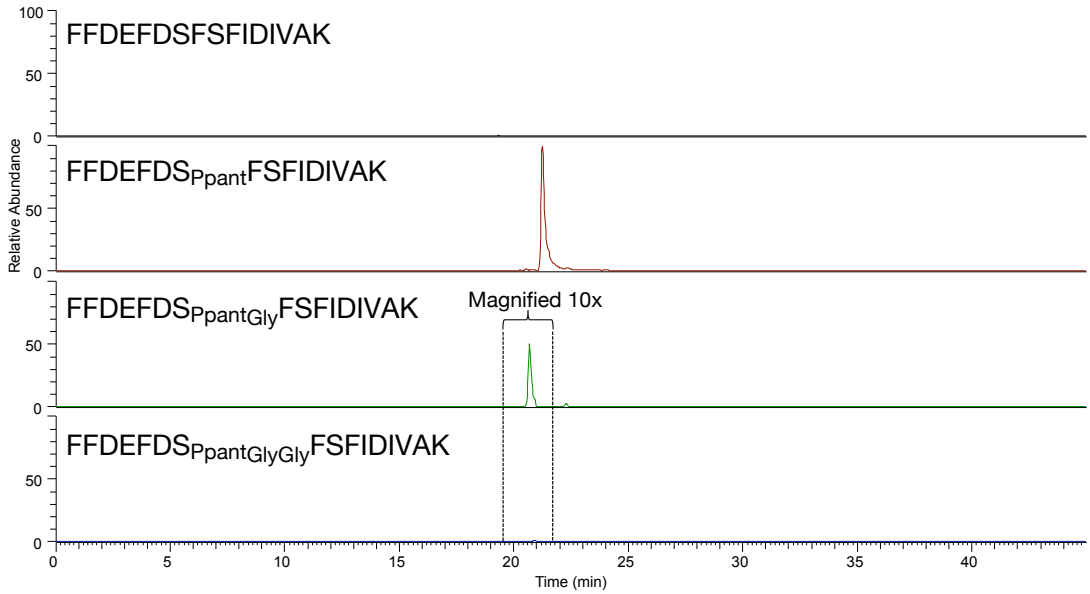


Figure 2.8: Extracted ion chromatogram of trypsinized AlmF purified from *E. coli* strains overexpressing Vc2457 and AlmE.

Shown from top to bottom are extracted ion chromatograms for doubly-charged molecular ions corresponding to unmodified (*m/z* 964), phosphopantetheine-modified (Ppant; *m/z* 1134), glycyl-phosphopantetheine-modified ($Ppant_{Gly}$*; m/z* 1163), and diglycl-phosphopantetheine modified ($Ppant_{GlyGly}$*; m/z* 1191) peptide, $FFDEFDS_{34}FSFIDIVAK$, from the tryptic digest of AlmF co-expressed with Vc2457 and AlmE. The extracted ion chromatogram for the $Ppant_{Gly}$-modified species has been magnified to better show the peak demonstrating glycylation of the Ppant-modified $FFDEFDS_{34}FSFIDIVAK$ AlmF peptide.

2.3 DISCUSSION

2.3.1 Overcoming historical obstacles: efficient over-expression of carrier proteins in *E. coli*.

Overexpression of carrier proteins in *E. coli* host expression strains, with high enough homology to *E. coli* apo-AcpP, a component of the essential type II *E. coli* fatty acid synthase, is inhibitory[19]. With 19% identity to *E. coli* AcpP, Vc1578 (AlmF) yields relatively low protein levels (<1 mg/L of *E. coli*), suggesting that high level expression of apo-AlmF is somewhat toxic in *E. coli*, albeit less than overexpression of *E. coli* AcpP. Optimization of AlmF production required a background strain known to improve expression of toxic proteins (BL21(DE3)pLysS). Similarly, co-expression of AlmF with Vc2457, which converts apo- to holo-AlmF, consistently improved protein yields by more than 3 fold. Conversely co-expression with Vc0780, which does not convert apo- to holo-AlmF did not improve overall yields. The low level expression of AlmF in *E. coli* contributed to the need to design a two-step AlmF purification scheme (See Chapter 5 Methods).

How apo-ACPs contribute to growth inhibition is indicated by structural data obtained between *Bacillus subtilis* AcpS bound to its cognate AcpP[8]. The coenzyme A binding pocket of AcpS is occluded when AcpP carrier protein is bound. Thus coenzyme A, the phosphopantetheine donor substrate, must bind before carrier protein for productive phosphopantetheinyltransferase activity. The *B. subtilis* and *E. coli* proteins are highly conserved. Tying anabolic cues, like sufficient coenzyme A levels, to fatty

acid production, used in de novo synthesis of membrane phospholipids, ensures the energetic costs associated with membrane assembly can be sustained by a cell's current metabolic state.

Key amino acid differences between *V. cholerae* AlmF and *E. coli* AcpP explain why overexpression of AlmF in *E. coli* results in minor toxicity. Again these insights come from co-crystal structures between *B. subtilis* AcpP and AcpS proteins. Key methionine and leucine residues of AcpP each bind within distinct, conserved hydrophobic pockets contained within AcpS[8]. Similarly a number of semi-conserved charged amino acids on an AcpP alpha helical region stabilize association with AcpS through salt bridges and hydrogen bonding interactions[8]. These interactions help orient the conserved carrier protein serine for phosphopantetheinyltransfer (Figure 2.1). Amino acids important for these interactions are conserved between AcpP and AcpS homologs of *B. subtilis* and *E. coli*. However AlmF lacks the methionine, replaced with the much smaller alanine, minimizing the hydrophobic pocket interaction with *E. coli* AcpS. Similarly, half of the conserved charged residues that stabilize the interaction between AcpP and AcpS helices are substituted for uncharged, polar residues in AlmF. These structural details provide reasonable hypotheses for why AlmF overexpression is less toxic: as a result of less tight binding to *E. coli* AcpS.

2.3.2 Unexpected substrate specificity of Vc2457, an AcpS subtype of the 4′-phosphopantetheinyltransferase super-family

The observation that supplementation of apo-AlmF overexpression strains with pantothenate results in~50% conversion to holo-AlmF, indicates that *E. coli* AcpS can be pushed to non-specifically phosphopantetheinylate AlmF (Figure 2.4), as reported for other heterologous carrier proteins in *E. coli*[17]. The genome of *E. coli* also encode for two putative Sfp subtype phosphopantetheinyltransferases, AcpT and EntD[20]. However AcpS is the most likely candidate for non-specific AlmF phosphopantetheinylation, with higher sequence identity to Vc2457 the specific and dedicated phosphopantetheinyltransferase from *V. cholerae* (AcpT 21%, EntD 27%, AcpS 56%; see Figure 2.3). Structrual overlay of Vc2457 and *E. coli* AcpS, onto the *B. subtilis* AcpS structure shows Vc2457 has improved specific association towards AlmF. Vc2457 has a shallower hydrophobic pocket, a leucine replaces a bulkier phenylalanine in *E. coli* AcpS, that would better bind the previously discussed AlmF alanine (See 2.3.1; methionine in AcpP). Residues important to orient the AlmF helix that ends in the conserved serine, site of phosphopantetheinylation, also differ between *E. coli* AcpS and Vc2457. Features in addition to these are also likely to contribute to Vc2457-AlmF specificity and require further investigation.

The discovery that an AcpS subtype, Vc2457, modifies AlmF, and not an Sfp subtype, like Vc0780, means the generalization that AcpS subtypes have more specific and limited substrate pools relative to Sfp subtypes needs revisiting. Before our study revealed their true functions, genome annotation suggested Vc1579 (Chapter 3, AlmE)

and Vc1578 encode proteins similar to type I synthase components involved in enterobactin production. Based on this annotation, an Sfp subtype like EntD would be the predicted 4′-phosphopantetheinyltransferase required for converting Vc1578 (AlmF) into a functional carrier protein. Sfp subtypes function as monomers, with two domains that each look like an individual AcpS subtype monomer. Sfp subtypes modify carrier proteins covalently linked to a series of type I synthase domains. The wide shallow surface area of Sfp subtypes provides a more general binding surface for substrate carrier protein domains[12]. Especially to carrier domains that can be buried within large type I synthase complexes. On the contrary trimerization of AcpS subtypes provides a more narrow surface area to contact substrate carrier proteins, where each has been experimentally shown to bind an individual monomer subunit[8]. Compared to the Sfp subtype, fidelity of AcpS subtypes is maintained by a narrower, deeper interaction surface[8]. Our study of AlmF, a predicted Sfp substrate, that is actually modified by an AcpS subtype, Vc2457, suggests a nomenclature that considers both the structures of 4′-phosphopantetheinyltransferases and modification of type I or II carrier domains, more accurately reflects our current understanding of carrier protein-phosphopantetheinyltransferase interactions. Under this proposed reorganization, structural analogs of the Sfp subtype are more accurately described as type I phosphopantetheinyltransferases that modify type I synthase carrier domains, and AcpS subtypes are type II phosphopantetheinyltransferases that modify type II carrier proteins.

2.3.3 Potential barriers to the efficient glycylation of holo-AlmF in *E. coli*

Our observation that total glycyl-AlmF pools are very low in *E. coli* heterologously co-expressing AlmF, Vc2457, and AlmE is surprising (Figure 2.8). In the following chapter efficient glycylation of holo-AlmF, at levels >70%, is routinely observed *in vitro,* in reaction mixtures that contain pure holo-AlmF and AlmE (Chapter 3). Thus it is unlikely that another factor is required for efficient glycylation of AlmF *in vivo*. More likely are factors within *E. coli* that diminish pools of glycyl-AlmF. Thioesterase activity against glycyl-AlmF is the most likely culprit, whereby the thioester bond that "carries" glycine is cleaved to regenerate holo-AlmF. *E. coli* encode for three biochemically verified acyl-coenzyme A thioesterases[21–23], which are possible candidates to convert glycyl-AlmF to holo-AlmF. Based on a survey of the ThYme thioesterase database in *E. coli* a number of other candidate thioesterases exist, and await further study[24]. Thioesterases have been shown to edit type I synthases activity, particularly in yeast and higher order eukaryotes[25]. For this reason thioesterases are of increasing interest to bio-engineers interested in the type I or type II production of commodities (i.e. plastics, fuel, cosmetics, vitamins, antibiotics) from photosynthetic or single-celled organisms[26].

Because *V. cholerae* biotype El Tor contains lipid A modified with not only one, but sometimes two glycine residues, we investigated whether holo-AlmF can shuttle a diglycine molecule. However our highly-sensitive MS/MS protocols were unable to identify peaks associated with more than one glycine attached to holo-AlmF (Figure 2.8). Based on this descriptive result, and in terms of the overall AlmEFG pathway, diglycine

lipid A formation likely results from slight processivity in the proposed glycyl-carrier protein/ lipid A late aminoacyltransferase AlmG, where it may be capable of transferring a second glycine to an already glycine modified lipid A.

Chapter 3: AlmE is an Aminoacyladenyltransferase/ Carrier Protein Ligase with Substrate Specificity Towards Glycine[3]

3.1 INTRODUCTION

Like acyl carrier domains discussed in chapter 2 there are a broad family of acyl adenyltransferase domains that can function within Type I or Type II synthase pathways. Many family members are not involved in either pathway. These domains catalyze the adenylation of substrate acyl compounds, with ATP or nicotinamide adenine dinucleotide (NAD) as adenyl group donors. The most famous examples are so-called adenyltransferase domains (A domains) of non-ribosomal peptide synthetases (NRPS)[1], acyl- or aryl- CoA ligases[2,3], DltA domains in gram-positives[4], and luciferase homology domains[5]. This particular list of adenyltransferase domains have dual functionality as carrier molecule ligases, transferring chemically activated, adenylated substrate molecules onto thioester based carriers like coenzyme A or acyl carrier proteins. Lipoate adenyltransferase/ protein ligases of the LplA class and class I aminoacyl tRNA synthetases catalyze similar two step chemical reactions: adenylation followed by covalent attachment to an acceptor molecule, however final attachment of lipoic acid to cognate apo-protein molecules and aminoacylation of tRNAs, proceeds through distinct mechanisms that do not form thioester products and are structurally distinct from aminoacyladenyltransferases[6–8]. Within families of structurally related

[3] Large portions of this chapter have been previously published. Henderson, J.C., Fage, C.D., Cannon, J.R., Brodbelt, J.S., Keatinge-Clay, A.T., and Trent, M.S. Antimicrobial peptide resistance of Vibrio cholerae results from an LPS modification pathway related to non-ribosomal peptide synthetases. *ACS Chem Biol*. **86**, 2138-45 2014. JCH along with equal contribution from co-author CDF contributed data, wrote, and edited this report.

acyladenyltransferases, differences in active site architecture govern substrate specificity. These nuances are of particular interest to pharmaceutical companies, industry, and basic science researchers interested in development of antibiotics, biofuels, and other natural product commodities through designed type I or type II synthases.

Aminoacyladenyltransferases, are a subset of adenyltransferases, specific for activating amino acids and related compounds. Amino acid selectivity was thought to be governed by a specific sequence of semi-conserved active site residues, referred to as the Stachelhaus code[9,10]. However as amino acid specificity has been determined for an ever-growing list of domains, the idea that a unique code exists for each amino acid loosely holds true for bulky amino acids at best. More divergent active site architectures exist for domains that specify for and activate smaller amino acids[11–13]. Enantiomeric selectivity of either a L- or a D- amino acid is often observed in aminoacyl adenyltransferases, with few examples of non-selective domains, where an associated racemase often determines the final enantiomer of a given residue within product Type I or II synthase compounds[14].

DltA aminoacyladenyltransferase domains and homologs within gram-positive organisms, like methicillin resistant *Staphylococcus aureus,* are enantiomer selective. In combination with other *dlt* operon proteins namely DltBCD, DltA helps modify cell surface wall and lipo-teichoic acid with D-alanine,[4,15]. In addition to conferring resistance to CAMPs, D-alanylation in gram-positives impacts cell division, biofilm formation, virulence and limits autolysis, a phenotype that contributes to horizontal gene transfer[16–18]. Crystallography data on DltA captured in various stages of its two-step mechanism,

adenylation followed by thiol transfer, has so far provided a terrific model for our enzymological understanding of the aminoacyl adenyltransferase family[15,19,20].

In this chapter the enzymology of AlmE is explored, a distantly related DltA homolog, responsible for adenylation and transfer of glycine to the carrier protein AlmF (Model provided in Chapter 1: Figure 1.10; Evidence provided in Chapter 2: Figures 2.5, 2.6, & 2.8). Through *in vitro* activity assays it is determined that AlmE inefficiently uses D-alanine as a substrate, with demonstrably better efficiency when glycine is provided in assay mixtures. Further exploration of substrate specificity is provided by discussion of the X-ray structure of AlmE in complex with a glycyl-adenylate intermediate determined at 2.26-Å resolution. Visualization of the AlmE active site allowed rational design of a single-residue mutant capable of efficiently activating and transferring L-alanine to AlmF. Final observations *in vivo* show that, despite efforts to design flexibility in AlmE amino acid substrate preference, the AlmEFG pathway remains specific for glycine addition to lipid A, where glycine modified LPS is ultimately assembled and deposited on the gram-negative cell surface.

3.2 RESULTS

3.2.1 AlmE has relaxed substrate selectivity for amino acids other than glycine.

To further explore the enzymology of AlmE, a rapid *in vitro* assay was developed that couples AlmE-catalyzed glycyl-AMP formation with indirect detection of pyrophosphate, the other AlmE reaction product (Figure 1.10). Pyrophosphatase is used to convert AlmE generated pyrophosphate to phosphate, where phosphate levels can be quantified using a spectrophotometric-based malachite green assay (Chapter 6 for details). *In vitro* AlmE assay conditions were optimized based on reports of previously characterized aminoacyl adenyltransferase/carrier protein ligases[15,21,22]. As monitored by our pyrophosphatase-coupled assay inclusion of purified holo-AlmF, compared to apo-AlmF, significantly stimulated AlmE activity (Figure 3.1). Among a representative sampling of amino acids, glycine is the preferred substrate for adenylation; however, a surprising level of activity (~25-30%) was observed when D-alanine is provided in AlmE reaction mixtures (Figure 3.2; Panel A). It should be noted that these assay conditions only monitor the ability to adenylate an amino acid and provide no information regarding subsequent ligation to holo-AlmF or other candidate acceptor substrates.

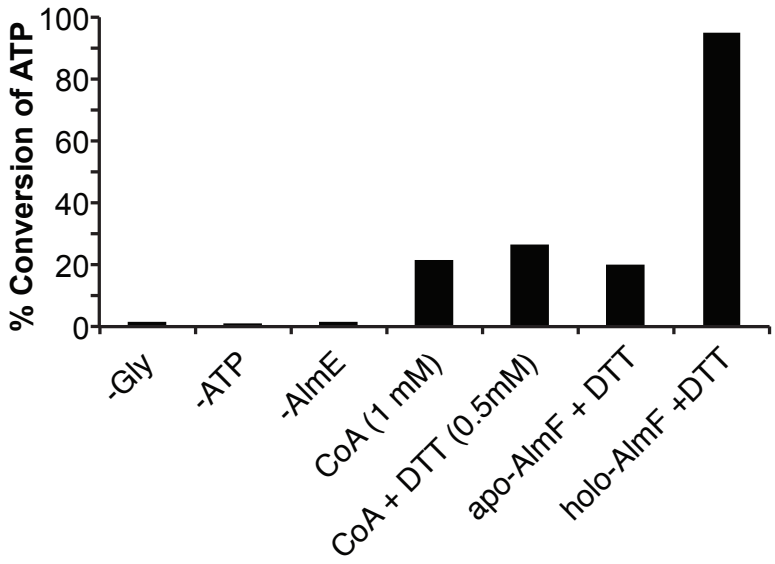


Figure 3.1 *Holo*-AlmF increases AlmE glycine adenyltransferase activity *in vitro*.

Relative to reactions where Coenzyme A (CoA) or apo-AlmF are provided as aminoacyl acceptors, providing holo-AlmF results in a robust increase in AlmE adenyltransferase activity under our standardized pyrophosphatase-coupled assay conditions. AlmE activity requires substrates glycine and ATP. Activity measurements represent the percent conversion of ATP (1 mM) supplied in the assay, where these assays were allowed to proceed to completion (4 h). Error bars denote the standard deviation of reactions as performed in triplicate, and are too small to see on this plot.

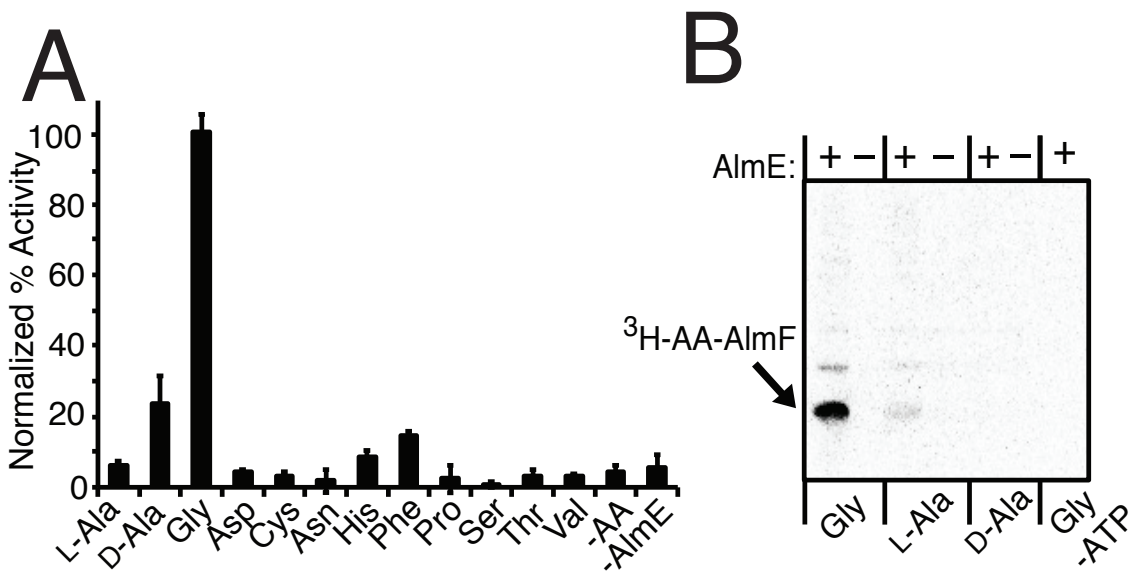


Figure 3.2 AlmE can efficiently adenylate glycine and to a lesser extent D-alanine. However, only glycine is efficiently transferred onto holo-AlmF.

(A) Various amino acids (1 mM) were provided as substrates in pyrophosphatase-coupled AlmE (800 nM) assays. Control reactions are included where no amino acid was provided (-AA) or no enzyme was provided in the reaction (-AlmE). Percent activity was determined within the observed linear range of activity (2 h), and normalized to glycine-containing reactions (~35% activity overall). Activity measurements represent the % conversion of ATP (1 mM) supplied in the assay. Error bars denote the standard deviation of reactions as performed in triplicate. **(B)** Destabilizing urea-PAGE analysis of holo-AlmF incubated with AlmE and the indicated ^{3}H-amino acid *in vitro*, shows that AlmE can efficiently transfer activated glycine onto holo-AlmF, with trace activity observed when ^{3}H-L-alanine is provided and no transferase activity when ^{3}H-D-alanine is provided.

To more comprehensively monitor two-step AlmE activity, we developed an alternative to the pyrophosphatase-coupled assay that could directly monitor formation of ^{3}H-aminoacyl-AlmF formation (See Chapter 6 for details). In this carrier protein focused assay, AlmE, ^{3}H-amino acid, and AlmF are combined in assay mixtures. TCA-precipitated assay products are resuspended and separated by SDS-PAGE and transferred

to a nitrocellulose membrane for final detection of radioisotope incorporation into AlmF carrier proteins. Formation of ^{3}H-glycyl-AlmF is indicated by the presence of the ^{3}H-labeled protein band (Figure 3.2; Panel B), which is not observable when ATP or AlmE are omitted from assay mixtures (Figure 3.2; Panel B lanes with -, or far right lane). We also tested other ^{3}H-labeled amino acids, such as ^{3}H-L-alanine and ^{3}H-D-alanine. Consistent with our pyrophosphatase assays AlmE activated and transferred ^{3}H-L-alanine (<10%; compare panels A and B Figure 3.2), but was unable to transfer ^{3}H-D-alanine to holo-AlmF (Figure 3.2; Panel B). This data suggests that *V. cholerae* preferentially activates and transfers glycine to holo-AlmF, consistent with its participation in the AlmEFG pathway that ultimately produces glycine-modified lipid A.

3.2.2. X-ray crystal structure of AlmE bound to a glycyl-AMP intermediate provides first-of-its kind details on a glycyl adenyltransferase/ carrier protein ligase.

To better understand the features that govern substrate selectivity for glycine we pursued the structural characterization of AlmE. No adenylation domains specific for this smallest amino acid have been deposited in the Protein Data Bank (PDB). AlmE formed narrow rod-shaped crystals with one monomer per asymmetric unit in space group $P3_121$ (Table 3.1). Its structure was determined via molecular replacement to a resolution of 2.26 Å (Table 3.1). In addition to DltA enzymes, AlmE resembles the adenylation domains (A-domains) of nonribosomal peptide synthetases (NRPSs) [1], acyl- and aryl-CoA synthetases [2,3], and firefly luciferases [5]; it is comprised of a large N-terminal body (50.6 kDa, residues 1-446), a smaller C-terminal lid (12.2 kDa, residues 450-556), and a

very short hinge region (residues 447-449) linking the two domains (Figure 3.3; Panel A).

Table 3.1 AlmE X-ray crystallography data and refinement statistics

Data collection	
Space group	$P3_121$
Cell dimensions	
a, *b*, *c* (Å)	116.2, 116.2, 99.6
α, *β*, *γ* (°)	90.0, 90.0, 120.0
Resolution (Å)	50-2.26
R_{merge}	0.107 (0.916)
$I/\sigma(I)$	16.6 (3.6)
No. of reflections	36632 (1800)
Completeness (%)	99.9 (99.5)
Redundancy	10.7 (10.0)
Refinement	
Resolution (Å)	50-2.26
R_{work}/R_{free}	0.166/0.193
No. of atoms	
Protein	4214
Ligand	27
Water	157
Average *B* factors (Å^2)	
Protein	33.9
Ligand	28.6
Water	33.5
RMS deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.085

(Values in parentheses refer to the highest resolution shell 2.30-2.26 Å)

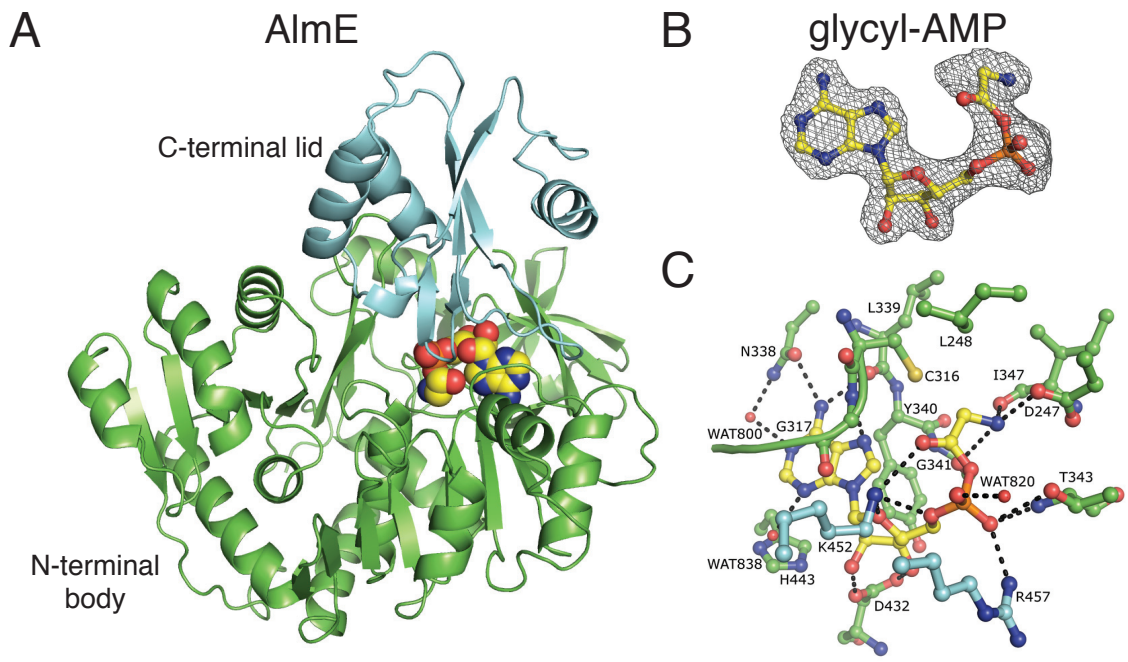


Figure 3.3 Crystal structure of AlmE with bound intermediate, glycyl-AMP.

(A) Cartoon representation of AlmE with glycyl-AMP (yellow spheres) bound between the N-terminal body (green) and C-terminal lid (cyan). AlmE is in the thioester-forming conformation in which the C-terminal lid has rotated ~140° after forming the aminoacyl-adenylate intermediate. In this conformation, AlmE is prepared to transfer glycine from glycyl-AMP to the Ppant arm of carrier protein AlmF. **(B)** F_o-F_c electron density omit map of glycyl-AMP, contoured at 3 RMSD. **(C)** Binding mode of glycyl-AMP (yellow). The adenine ring rests between His443, Ile365 (not pictured), and conserved Tyr340 on one side, and the Gly317-containing loop on the other. Residues contributed by the N-terminal body are green and those by the C-terminal lid are cyan. Hydrogen bonds are shown as dashed black lines. Some main chain atoms are hidden for clarity.

Formation of the glycyl-adenylate intermediate was achieved by incubating AlmE with glycine, $MgCl_2$, and ATP prior to crystallization (Figure 3.3; Panel B). No electron density is apparent for the His-tag, residues 1-14, residues 25-29 that connect the N-terminus to the main body of the enzyme, residue 189, residues 533-536, or 551-556 of

the C-terminus. Density for residues 15-24 is visible adjacent to the C-terminal lid of a symmetry mate, with residues 19-23 forming a short β-strand that runs antiparallel to the last β-strand (residues 523-527) of the neighboring lid.

Several AlmE residues make favorable contacts general to all aminoacyl groups, while others may play a greater role in enforcing glycine selectivity (Figure 3.3; Panel C). The Lys452 amine forms a charged hydrogen bond with the glycyl carbonyl, while the carboxylate of Asp247 as well as the carbonyls of Gly341 and Ile347 make charged interactions with the positively-charged glycyl amine. The residues of AlmE nearest the C_α of the glycyl-adenylate, Leu248 (3.9 Å from $C_{\delta 2}$) and Cys316 (3.4 Å from S_γ), may select against amino acids larger than glycine by sterically clashing with their α-substituents - Leu248 with the side chains of D-amino acids, and Cys316 with those of L-amino acids. The conformation of Leu248 is determined through interactions with the side chains of the neighboring residues Phe246, Ser242, and Leu287. Structural evidence supports the hypothesis that Leu248 appears to contribute to stereoselectivity against D-amino acids, while Cys316 is positioned to sterically select against L-amino acids with large C_α-substituents.

3.2.3 Structural data informs design of an AlmE variant, Cys316Ala, that can efficiently adenylate and transfer L-alanine to holo-AlmF *in vitro*.

By comparing the active site architecture of AlmE with that of D-alanyl carrier protein ligase DltA, we predicted how glycine is selected over other small amino acids such as L- or D-alanine. Based on this analysis, as well as our observation that AlmE

inefficiently adenylates D-alanine and poorly ligates L-alanine onto AlmF *in vitro* (Figure 3.2; Panel B), we decided to investigate the specificity of AlmE through rational mutagenesis. Single-residue variants of Leu248 and Cys316 were prepared with the hypothesis that less bulky side chains would relax the specificity of AlmE for D- or L-amino acids (Figure 3.4; Panel A). Using site-directed mutagenesis, we prepared variants Cys316Ser, Cys316Ala, Leu248Val, and Leu248Ala and tested their specificity for glycine, D-alanine, L-alanine, L-serine, L-cysteine, and L-valine.

As expected Leu248Val and Leu248Ala mutants did not improve AlmE specificity toward L-amino acids since Leu248 is positioned to sterically exclude side chains of D-amino acids (Figure 3.4; Panel B). However, these mutants did not show expected activity toward D-alanine, or even the typical AlmE substrate, glycine. This could indicate that Leu248 substitution significantly alters the geometry of the amino acid binding site, to disrupt a number of the favorable binding interactions described above (Figure 3.3 Panel C; also see Section 3.3.1 for more details).

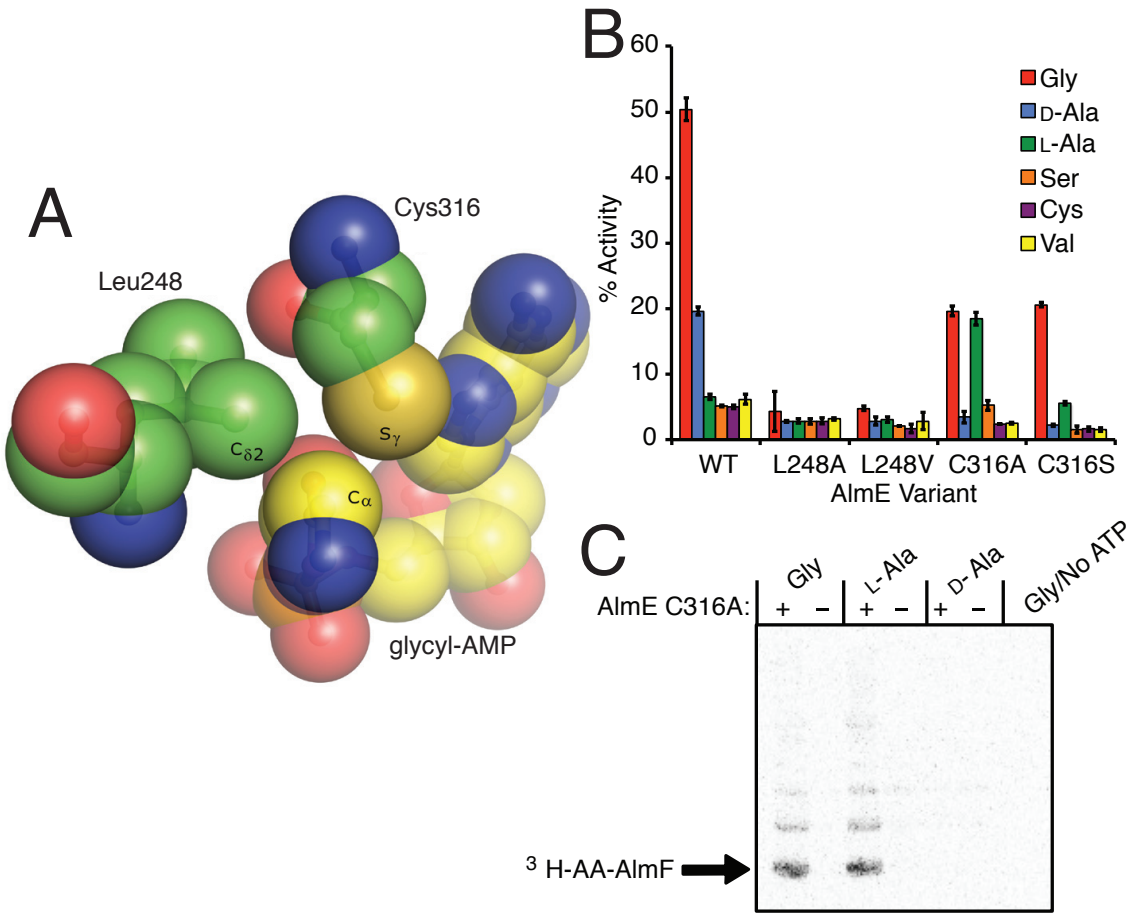


Figure 3.4 A rationally designed AlmE Cys316Ala variant shows relaxed specificity for L-alanine.

(A) The side chains of Cys316 and Leu248, positioned above C_α of the glycyl-adenylate in the image, are hypothesized to exclude large amino acid chains via steric hindrance. Mutation of Cys316 and Leu248 to smaller amino acids might allow for relaxed specificity toward L- or D- amino acids, respectively. **(B)** All single-site AlmE variants display modest reduction in activity towards glycine; however, AlmE Cys316Ala shows relaxed specificity for adenylation of L-alanine. AlmE variants (800 nM) were incubated with small amino acid substrates (1 mM) under standard pyrophosphatase-coupled assay conditions. Activities were measured in the linear range of wild-type AlmE for glycine. Activity measurements indicate the percent conversion of ATP (1 mM) supplied in the reaction, taken at 2.5 h. Error bars represent the standard deviation of reactions, measured in triplicate. **(C)** Destabilizing urea-PAGE analysis of holo-AlmF incubated with AlmE Cys316Ala and the indicated ^{3}H-amino acid *in vitro* shows that the Cys316Ala variant can efficiently transfer activated glycine or L-alanine to holo-AlmF.

The prediction that adequate space could be provided for an L-amino acid by mutating Cys316 appears valid. Replacement of the Cys316 sulfur atom with the relatively smaller oxygen atom, as in Cys316Ser, or removal of the thiol group, as in Cys316Ala, resulted in L-alanine adenylation. The Cys316Ser variant displayed reduced but modest activity for glycine and L-alanine relative to wild-type AlmE, while the Cys316Ala variant was equally able to adenylate glycine or L-alanine at levels below wild-type AlmE with glycine (Figure 3.4; Panel B). Furthermore, the AlmE Cys316Ala variant was equally able to form ^{3}H-glycyl- and ^{3}H-L-alanyl-AlmF, as determined by the destabilizing PAGE/radioisotopic assay described above (Figure 3.4; Panel C). As both Cys316Ser and Cys316Ala variants were less active for glycine adenylation than wild-type AlmE, it appears that substituting smaller side chains removes optimal packing interactions maintained between native AlmE and the C_{α} of the glycyl group (Figure 3.4; Panel A). These results further indicate that the Cys316 thiol of wild-type AlmE sterically occludes all L-amino acids from the active site.

3.3 DISCUSSION

3.3.1 Detailed comparison of factors that contribute to differential amino acid selectivity of DltA and AlmE.

The *B. subtilis* and *B. cereus* DltA enzymes, *bona fide* D-alanyl adenyltransferase/ carrier protein ligases, share respective 32% and 31% overall sequence identity with glycine specific AlmE (Protein BLAST[23]). AlmE has more homology with DltA domains, than with even glycine specific A domains of type I NRPS. On one hand DltA and AlmE participate in type II synthase related pathways, so perhaps it's not surprising that they exhibit such a high degree of conservation. On the other hand, it is remarkable that the active site specificity determining residues are more related between DltA-AlmE than between glycine A domains-AlmE (Figure 3.5). This observation contributes to ever-compounding evidence that a Stachelhaus-like code, or a conserved, specificity determining sequence within aminoacyladenyltransferases, likely does not exist for small amino acids especially[12].

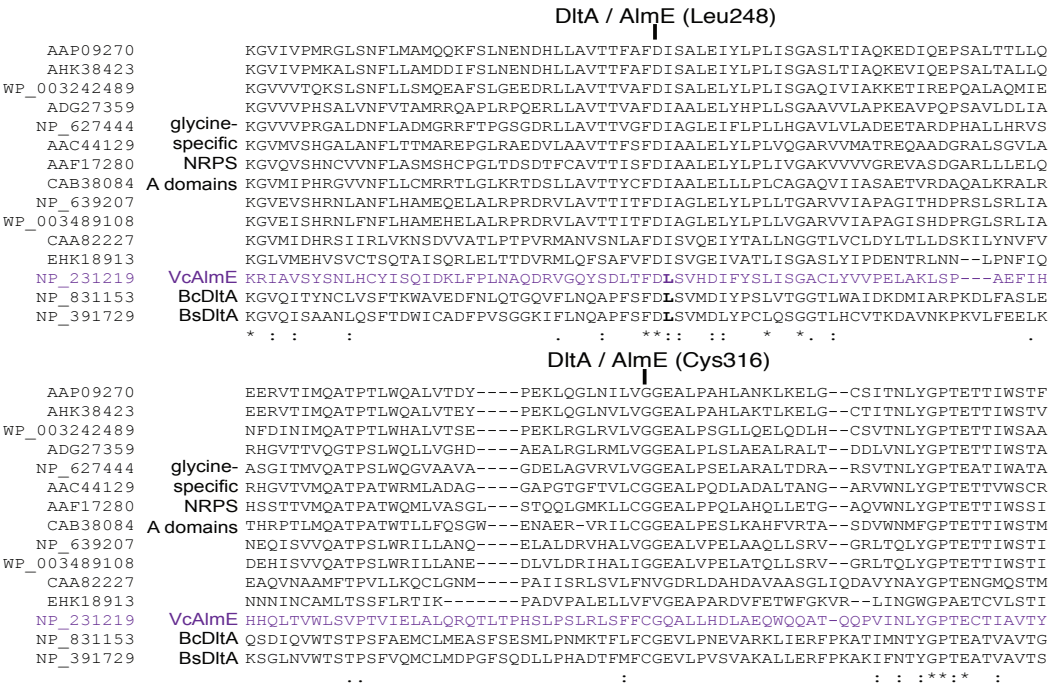


Figure 3.5 Alignment of specificity-encoding regions of Glycine specific A domains from NRPS or AlmE, and D-alanine specific DltA domains

Sequence alignment of specificity-encoding region of glycine-specific A-domains[12], AlmE (purple), and DltAs (*B. subtilis* and *B. cereus*). AlmE is more closely related to DltAs than glycine A-domains of NRPSs. The two residues closest to the C_α of the aminoacyl-adenylate in glycine-specific AlmE, Leu248 and Cys316, are conserved in the D-alanine-specific DltA but are replaced by isoleucine and glycine in glycine-specific A-domains. Symbols below alignment represent degree of conservation of particular residues: fully conserved (*), strong conservation (:), and weak conservation (.).

Due to their similarity with type II synthases, and as AlmE adenylates and thiolates a small amino acid, it may be unsurprising that the atomic coordinates of AlmE are more similar to D-alanine specific DltA enzymes, than adenylation domains specific for larger amino acids[19,20] (*B. subtilis* DltA, PDB code 3E7W; *B. cereus* DltA, 3DHV). The N-terminal bodies of AlmE and DltA display remarkable overlap, with a root means squared deviation (RMSD) of 1.41 Å over 278 C_α atoms for *B. subtilis* DltA and 1.59 Å

RMSD over 289 C_α for the *B. cereus* enzyme (PyMOL). This conservation in architecture is important as most of the residues important for stabilizing the aminoacyladenyl intermediate are associated with the N-terminal body, not the C-terminal lid domain (Figure 3.3; Panels A & C). For example, a unique distortion is seen in DltA or AlmE N-terminal body residues, Val301 and Ile347 respectively, whose backbone carbonyls hydrogen bond with the amine terminus of aminoacyladenyl intermediates. Despite being different residues altogether AlmE Ile347 and DltA Val301 show distorted geometries that are Ramachandran outliers from conformations typically observed for these hydrophobic, branched amino acids in most protein structures (See Figure. 3.6)

There are two distinguishing features between AlmE and DltA enzymes that may explain the preference of AlmE for adenylation of glycine versus D-alanine (Figure 3.2; Panel A). Ser242 and Leu287 of AlmE are substituted with smaller side chains Ala192 and Thr240 in *B. cereus* DltA. The smaller side chains in DltA help position the C_δ atoms of DltA Leu198, spatially equivalent to Leu248 of AlmE (close-up in Figure 3.4 Panel A), farther from C_α of the aminoacyladenylate. In other words, positioning the leucine further from the aminoacyladenylate helps accommodate more room for the methyl group of D-alanine in DltA, than in AlmE where Leu248 is much closer to the C_α of the aminoacyl-adenylate (Figure 3.6). Similarly, the helix containing Leu248 in AlmE is straight, in contrast to the same helix in DltA, which has a slight bend. The bent geometry of the DltA helix creates more space for the D-alanine methyl group (Figure 3.6).

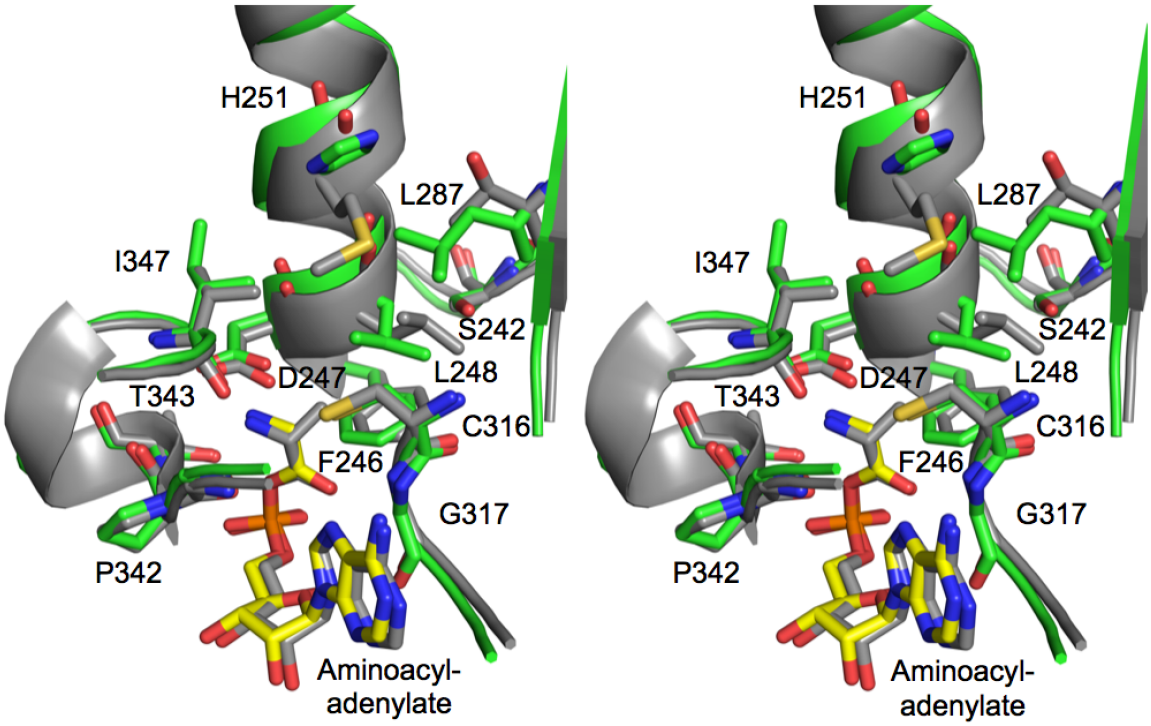


Figure 3.6 Stereoview of the active sites of AlmE (green, with the glycyl-adenylate in yellow) and *B. cereus* DltA (grey, with the D-alanyl-adenylate also in grey) superimposed.

Small differences in the positions and identities of residues near the aminoacyl groups presumably result in the different specificities of the two enzymes. For example, S242, F246, and L287 affect the position of L248, which sterically clashes with side chains of D-aminoacyl groups in AlmE. S242 and L287 are substituted with A192 and T240 in DltA; additionally, H251 is replaced with M201. For clarity, only residues and secondary structural features near the aminoacyl groups are shown. AlmE residues are labeled; corresponding DltA residues include A192 (S242), F196 (F246), D197 (D247), L198 (L248), M201 (H251), T240 (L287), C269 (C316), G270 (G317), P296 (P342), T297 (T343), and V301 (I347).

Single-residue variants of Leu248 and Cys316 were prepared with the hypothesis that less bulky side chains would relax the specificity of AlmE for D- or L-amino acids respectively (Figure 3.4 Panels A & B). This hypothesis is rationally motivated. Atomic coordinate overlap of AlmE and DltA, with NRPS A-domains that act on bulkier

residues, show smaller, more accommodating amino acids are substituted near the C_α of the aminoacyladenylate intermediate. Comparison was made with L-phenylalanine-specific PheA (PDB code 1AMU), L-leucine-specific SrfA-C (PDB code 2VQS), and L-valine-specific CytC1 (PDB code 3VNS). AlmE Leu248 substitutions reduced activity below the detection limit of our assays, however Cys316Ala substitution successfully allowed room for the methyl group of L-Alanine (Figure 3.4; Panel B). Concomitant loss in enzymatic efficiency upon Cys316 substitution, also reinforces the overall importance of the Cys316 thiol in stabilizing the adenylglycl intermediate, as predicted from the crystal structure (Figure 3.3: Panel C or Figure 3.4: Panel A).

3.3.2 A conformational snap-shot of multi-step aminoacyladenyltransferase/ carrier protein ligase catalysis: a model for AlmE through the lens of structural data on DltA and A domains of NRPS.

As mentioned in Section 3.1 crystallographic data on DltA domains, captured in various stages along the two-step adenylation and carrier thiolation chemical reactions, provide foundational structural models for our understanding of the aminoacyladenyltransferase/ carrier protein ligase family[15,19,20]. An overview model is presented in Figure 3.6, originally published by Yonus and colleagues[19]. The starting, open conformation is modeled from type II firefly luciferase and A domain of the termination module in the Type I Surfactin synthase[5,24] (Figure 3.6; top left "open"). A major conformational change occurs between the open conformation and substrate bound states, where a domain rearrangement brings the C-terminal domain and N-terminal body

closer together (Figure 3.6 compare “open” to “adenylation”). These domains are separated by a hinge region, which also exhibits a dramatic conformational shift (Figure 3.6; orange loop). The active-site also displays moderate rearrangement, especially in a region that contains the catalytic lysine residue essential for adenyltransferase activity (Figure 3.6; green region). Pyrophosphate release brings the C-terminal domain and N-terminal body even closer together, with a ~140° rotation of the C-terminal lid relative to its position in the adenylation conformation, to completely limit solvent accessibility to the aminoacyladenylate (Figure 3.6 compare “adenylate” to “thiolation”). The structure of AlmE was captured in this putative thiolation ready state (Compare Figure 3.3 Panel A to Figure 3.6; “transfer”). In addition to limiting solvent accessibility this conformation provides a binding pocket for the 4′-phosphopantheteinyl arm of a cognate holo-carrier protein.

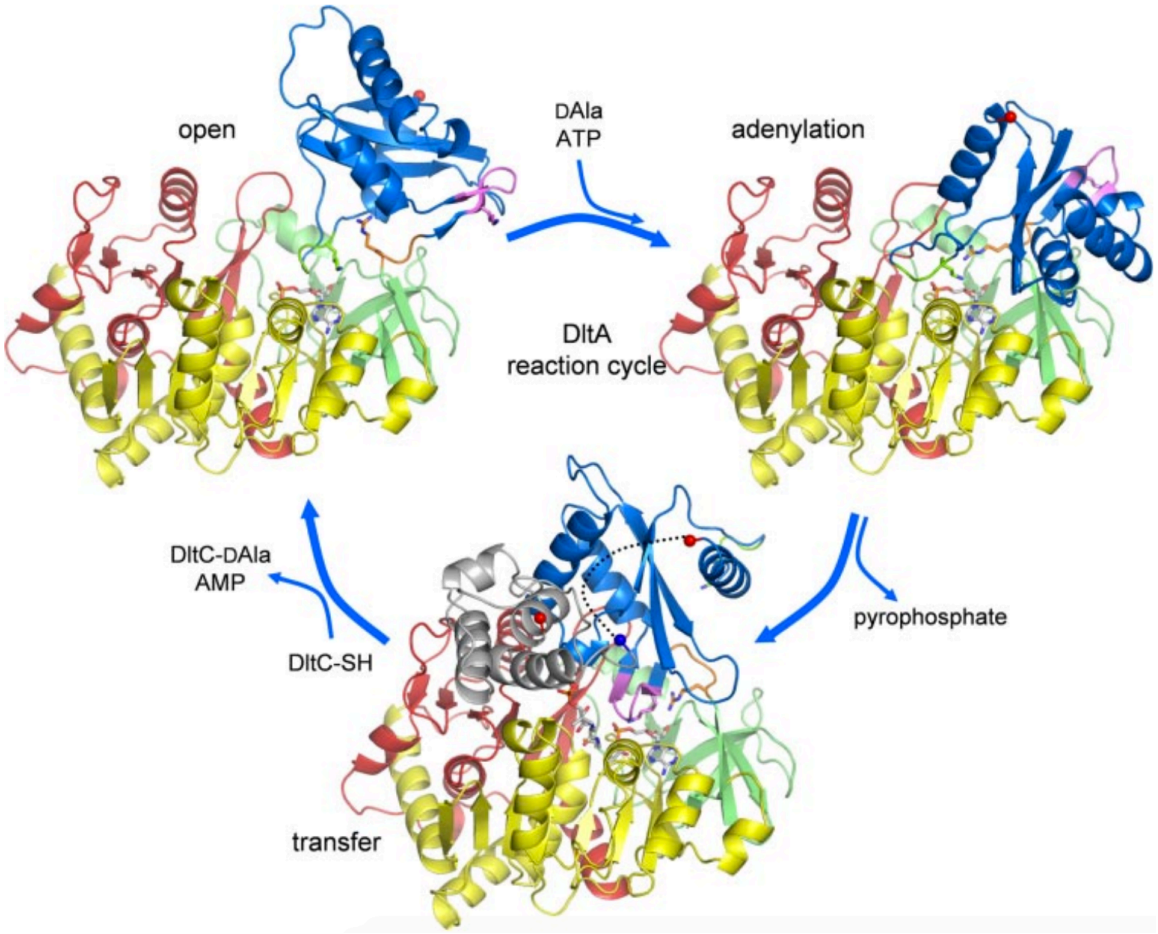


Figure 3.6 Proposed DltA catalytic reaction cycle modeled on available structures of aminoacyladenyltransferase domains.[4]

Color coded regions include the C-terminal domain (blue; with a red sphere indicating the C-terminus), N terminal body (yellow), residues of the hinge region between N- and C-domains (orange), the loop containing Lys-491 (green; necessary for the adenylation half-reaction), and the region containing Lys-402 (lilac; implicated in the thiolation half-reaction). Starting in an open conformation, binding of ATP could facilitate a mutual approach of the large and small domains to complete the active site for D-Ala adenylation. Upon loss of pyrophosphate, the small domain rotates ~140°, bringing residues into place for the thiolation reaction. Presentation of the phosphopantheinyl side chain (SH) of DltC (white) in the active site of DltA would then complete the reaction. In NRPSs, the C-terminal of the adenylation domain would be tethered to the N terminus of the peptidyl carrier protein; this is indicated by the dotted line in the bottom-most model.

[4] Originally published in Yonus *et al*. (2008) Crystal structure of DltA: Implications for the reaction mechanism of non-ribosomal peptide synthetase domains. *J Biol Chem*. 283(47)32484-91.

Relatively few X-ray structures of NRPS adenylation domains in complex with their cognate carrier proteins exist, although some have been obtained with irreversible inhibitor and fusion proteins[25,26]. These structures reveal the major sites of interaction between acyladenylation domain/carrier protein pairs. Two such structures were used to model AlmE and AlmF interaction (Figure 3.7). Attempts to seed crystals of AlmE in complex with holo-AlmF were not successful. AlmE and AlmF were modeled according to the EntB-EntE fused pair[25] (PDB code 3RG2), and also to the MatB co-structure with bound phosphopantetheine[27] (PDB code 3NYQ). The AlmE structure was superposed with the adenylation domain EntE (2.43 Å RMSD over 366 C_α atoms) and the Phyre2 model of AlmF was superposed with the carrier protein EntB (2.88 Å RMSD over 57 C_α atoms). Aside from the 4´-phosphopanetheinyl arm, most interactions with AlmE are expected to involve the AlmF nucleophilic helix (e.g., residues Phe35, Ile38, Asp39, Lys43, and Glu45) and portions of the loop N-terminal to it (e.g., residues Gln14-Ser16, Glu18, Phe29-Asp33). The 4´-phosphopantetheine arm of AlmF is expected to bind AlmE in a mode analogous to the pantetheinyl moiety observed bound to the *Streptomyces coelicolor* malonyl-CoA synthetase MatB[27], where MatB also crystallized in the thiolation conformation (Figure 3.6). After superposing AlmE and MatB, the positions of AlmE backbone atoms that form hydrogen bonds with the 4´-phosphopantetheinyl arm are very similar. Details are provided in the figure legend of Figure 3.7. Conserved residues for adenyltransferase/carrier protein pairs have not been identified, in part because carriers tend to interact with multiple adenyltransferase partners. In support of this idea, interactions between known pairs are mostly insensitive

to single site variation[28,29]. Similarly, given an adenyltransferase domain, directed evolution experiments reveal multiple carrier protein surface residues contribute to its final determination as an ideal substrate for acyltransfer[30].

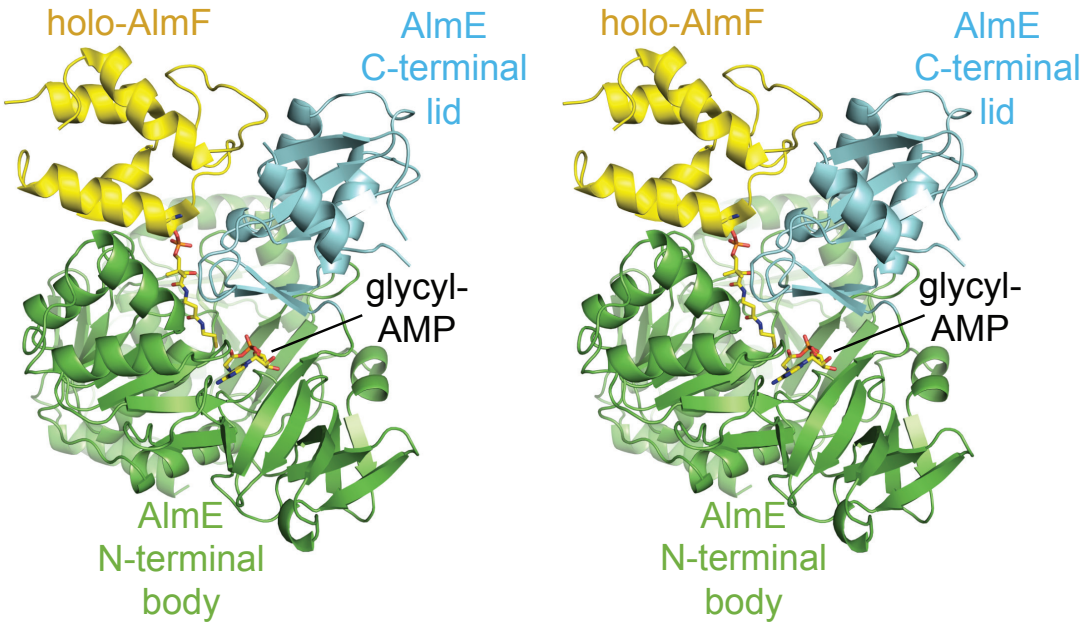


Figure 3.7 Predicted stereo-view structural model of the AlmE-AlmF complex.

The AlmE-AlmF complex is modeled from the EntB-EntE pair, as these pairs have high sequence homology. Interactions between AlmE and the 4′-phosphopantetheine of holo-AlmF are modeled based on the co-crystal structure of MatB with phosphopantetheine. These interactions in detail: Gly455 carbonyl (binds NP1) and Ala454 carbonyl (binds NP2 or OP3). The side chain of AlmE Ser242 may rotate to form a hydrogen bond with OP1 (occupied by water in MatB); those of Asp243 and Gln456 likely rotate for polar interactions. Lys271 of AlmE is forms a hydrogen bond with OP2. Tyr241 and Val289 also contribute steric bulk. For the 4′-phosphopantetheinyl arm to reach the aminoacyladenylate, the phenyl ring of Phe246 must rotate to match the position of the imidazole ring of His187 of MatB, as in its observed position it obstructs the path of the incoming thiol. Phe246 is in this conformation in *B. subtilis* DltA (PDB code 3E7W). AlmE Lys522 likely replaces Arg461 of MatB in binding the 4′-phosphate group of the coenzyme.

3.3.3 Overall selectivity for glycine by the AlmEFG pathway, despite relaxed amino acid specificity by AlmE.

In both types of AlmE assay, which monitored adenyltransferase or thiolation activity, glycine was the preferred substrate relative to other amino acids tested, as measured by total activity (Figure 3.2). Consistent with the proposed model for this family of enzymes, the observation that holo-AlmF increased total AlmE adenyltransferase activity (Figure 3.1), indicates that transfer of glycine to carrier protein efficiently restores AlmE for another round of catalysis, to the open substrate ready conformation (Figure 3.6). It simultaneously provides the most direct evidence to verify the proposed roles of these proteins as glycine carrier protein, AlmF, and glycine adenyltransferase/ carrier protein ligase, AlmE. Despite low-level wild-type AlmE adenyltransferase activity with D-alanine (Figure 3.2; panel A), subsequent lack of ^{3}H-D-alanine transfer onto holo-AlmF (Figure 3.2; panel B), explains how otherwise adenylated D-alanine is excluded from downstream reactions of the AlmEFG lipid A modification pathway. On the other hand, as some adenyltransferase activity was observed with D-alanine the exciting possibility remains that AlmE could transfer adenylated D-alanine onto another carrier protein, or molecule, *in vivo*.

Observations of AlmE activity with L-alanine are unexpected. Although adenyltransferase activity with this amino acid was near background (Figure 3.2; Panel A), some transfer of ^{3}H-L-alanine onto AlmF was detectable in our *in vitro* thiolation assay (Figure 3.2; Panel B). Low level activity of *B. cereus* DltA with L-alanine has been reported[20]. These authors cite the much higher k_{cat}/K_M with D-alanine relative to L-alanine,

in combination with > 20 fold higher intracellular concentration of D-alanine relative to L-alanine in *B. subtilis*[31] as providing at least 60 fold selection against L-alanine[20]. Despite significant effort, the reported intracellular concentration of glycine and L-alanine in *V. cholerae* could not be determined by this author. However, our MS analysis of *E. coli* strains co-expressing AlmE and holo-AlmF did not indicate the formation of L-alanyl-AlmF *in vivo* (as in Figure 2.5). Similarly L-alanyl lipid A has never been observed in *V. cholerae,* thus even if an intracellular pool of L-alanyl-AlmF is generated, it doesn't appear to be a suitable substrate for AlmG, the final step of the AlmEFG pathway. Analogous to the argument above: that AlmE could transfer adenylated D-alanine to another carrier molecule; it also remains possible that L-alanyl-AlmF, if generated, could serve as substrate in hitherto unknown cellular pathways. In sum, despite the minor degree of relaxed specificity with AlmE the overall AlmEFG pathway is strictly specific for glycine. Glycine likely evolved as the amino acid AlmEFG adds to lipid A, as larger amino acids would have deleterious effect on membrane packing, while the small size of glycine provides sufficient electropositive density to effectively resist CAMPs.

Chapter 4: AlmG is a Unique Lipid A Glycyl-transferase of the Bacterial Lysophospholipid Acyltransferase Superfamily

4.1 Introduction

The lysophospholipid acyltransferase superfamily is a massive collection of enzymes with homologs present in all domains of life. Genomes of most organisms encode for multiple superfamily members. Originally classified and enumerated based on observed activity in eukaryotes[1], this superfamily catalyzes transfer of acyl groups from donor molecules, like carrier proteins or coenzyme A, onto acceptor lipids, such as lyso-glycerophospholipids or lipid A, or glycerophospholipid precursors, such as dihydroxyacetonephosphate- or glycerophosphate. Investigations into the asymmetry in hydrocarbon acyl chain distribution found in eukaryotic and eubacterial glycerophospholipids helped uncover this superfamily. On the glycerol backbone of glycerophospholipids, a saturated or monounsaturated fatty acid is usually found at the *sn-1* position, whereas a poly-unsaturated fatty acid is typically found at the *sn-2* position. However biosynthetic enzyme substrate specificity does not adequately explain the observed differences in acyl chain composition found at each glycerol stereogenic center.

In the 1950s Lands' proposed that remodeling enzymes were responsible for turnover and observed heterogeneity particularly at the *sn-2* position of eukaryotic glycerophospholipids[2]. Indeed many lysophospholipid acyltransferases function in this capacity, contributing to the heterogeneity and asymmetry of acyl groups from structurally characterized lipids[1]. The study of diversity in lipid acylation contributes to

our understanding of homeoviscous adaptation and membrane biology. The influence of environmental factors on membrane curvature, fluidity, density and assorted functionality of individual lipid species as secondary biochemical messengers have and will continue to provide exciting insight into mechanisms of molecular communication, transport, and organismal fitness[3–5].

Lipid A late acyltransferases comprise a unique, characteristic subset within the historically named lysophospholipid acyltransferase superfamily, and participate in secondary acylation of bacterial Kdo-lipid A domains. Distinct in sequence and structure from primary lipid A acyltransferases, LpxD or LpxA of the canonical lipid A biosynthesis pathway (Figure 1.4), secondary acyltransferases esterify acyl groups to primary hydroxyl-acyl chains (Figure 4.1; Top Panel A LpxL and LpxM). In *E. coli* these activities are encoded by the genes *lpxL* or *lpxM*[6,7], which respectively transfer lauroyl (C_{12}; 12 carbons) or myristoyl (C_{14}; 14 carbons) fatty acids from acyl carrier proteins to hydroxyacyl fatty acids at the 2′ and 3′ positions of the Kdo-lipid A di-glucosamine backbone (Figure 4.1; Top Panel A upper right structure). Biochemical characterization of lipid A late acyltransferase homologs across bacterial species show nuanced heterogeneity in acyl chain selectivity (e.g. length, saturation, hydroxy versions) as well as positional specificity[8–11].

Lipid A late acyltransferase activity significantly impacts bacterial physiology, where virulence and permeability to antibacterial compounds, particularly of pathogenic organisms, is best studied. Mutants of secondary acyltransferases in clinical isolates of

enterohemmorhagic *Escherichia coli* are attenuated for virulence as assessed by lethality in intraperitoneal infected mice models[12]. *Salmonella* LpxM mutants are also attenuated in models of virulence or inflammation, and display marked reduction in the secretion of virulence factors from type III secretion systems[13–16]. Interestingly, all pathogenic strains of *Shigella flexneri* contain a virulence plasmid that encodes an additional lipid A late acyltransferase homolog[17], where disruption of any of the multiple secondary acyltransferases of *S. flexneri* cause defects in invasion and intracellular replication[17,18].

Resistance to antibacterial compounds, particularly against lipid targeting CAMPs, is sufficiently improved in strains with functional secondary acyltransferases, ascertained from studies on *Klebsiella, Escherichia, Salmonella,* and *Acinetobacter*[11,19–21], among many others. More acyl chains likely reinforce the surface lipid monolayer against permeation by CAMPs through increased density of interaction between hydrophobic acyl chains on associated Kdo-lipid A domains. Interestingly, *Acinetobacter baumanii* has a late acyltransferase homolog that adds to multiple positions of the Kdo-lipid A of this organism. This activity not only provides increased resistance to CAMPs, but also improves survival of this gnarly nosocomial pathogen during desiccation, a phenotype that contributes to spread in hospital settings[11].

Previous interpretations of lipid A structural data obtained with material isolated from pathogenic *V. cholerae* O1 and O139, revealed secondary acylation patterns inconsistent with structures of model gram-negatives[22]. As attenuated *V. cholerae* whole cell have long been the basis for commercial cholera vaccines, the structures of pro-

inflammatory endotoxin, or Kdo-lipid A, a surface component with known adjuvanticity[23], warranted further analysis. Subsequent investigation, with more sophisticated mass spectrometry protocols, convincingly demonstrated that the previously reported lipid A structures contained errors in assigned acyl chain composition and pattern[19,24]; in both primary and secondary acyl chains. In these efforts a new member of the lipid A late acyltransferase family was identified, LpxN, which transfers a hydroxylaurate residue (3-OH C12:0) to the hydroxyacyl chain at the 3′ position of Kdo-lipid A[19,24] (Figure 4.1; bottom panel B).

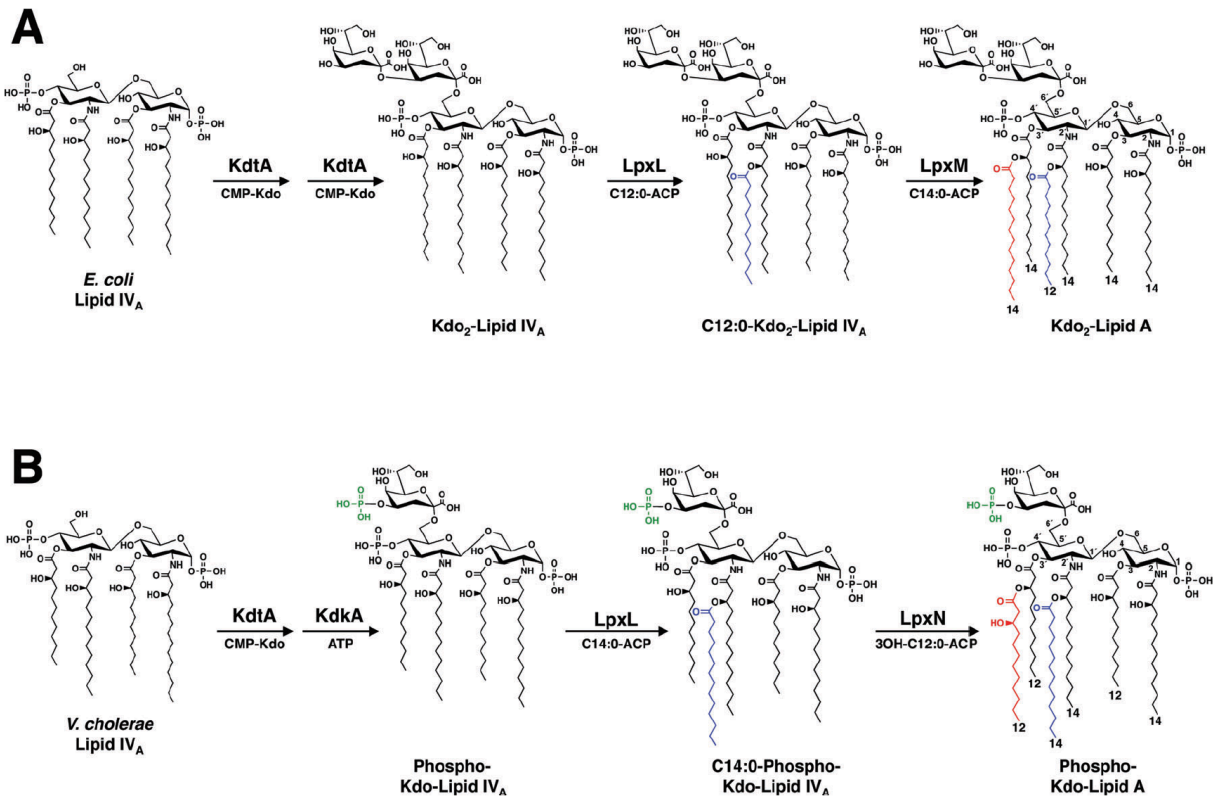


Figure 4.1 Comparison of late steps in the Raetz pathway of Lipid A biosynthesis in *E. coli* compared to the proposed pathway in pathogenic *V. cholerae*.[5]

(**A**) *E. coli* possess a bifunctional Kdo transferase (KdtA, also known as WaaA), which transfers two individual Kdo sugars each to the lipid A precursor lipid IVa, and Kdo-lipid IVa respectively. The lipid A secondary acyltransferases, LpxL and LpxM then catalyze the transfer of a laurate (C12:0) (blue) to the 2′-position hydroxylacyl chain and a myristate (C14:0) (red) to the 3′-position hydroxylacyl chain of the glucosamine disaccharide, creating Kdo_2-lipid A. (B) The *V. cholerae* genome encodes a mono-functional KdtA, which transfers one Kdo residue to lipid IVa. The Kdo kinase (KdkA) then phosphorylates the Kdo sugar (green), resulting in monophosphorylated Kdo-lipid IVa. V. cholerae LpxL (Vc0213) catalyses the transfer of a myristate (C14:0) (blue) to the 2′-position hydroxylacyl chain of the glucosamine disaccharide. LpxN (Vc0212) transfers a 3-hydroxylaurate (3-OH C12:0) (red) to the 3′-position hydoxylacylchain, generating the hexa-acylated *V. cholerae* Kdo-lipid A domain.

---

[5] Figure originally published in Hankins, J.V. *et al*., Elucidation of a novel *Vibrio cholerae* lipid A secondary hydroxyl-acyltransferse and its role in innate immune recognition. *Mol Microbiol*. **81**(5), 1313-1329. (2011).

LpxN is expected to follow LpxL activity in the *V. cholerae* pathway. LpxL is remarkable in that it *requires* a phosphorylated mono-Kdo lipid substrate (Figure 4.1; panel B KdkA phosphorylates Kdo-Lipid IVa), and will not modify Kdo_2-lipid A substrates[24]. Similar specificity is also observed in secondary acyltransferases of *Haemophilus influenzae* and *Bordetella pertussis*[24]. Late acyltransferases of *E. coli*, and other organisms, are not as specific for substrate Kdo-lipid A domains, but with canonical preference for Kdo_2-lipid A[10] (Figure 4.1; compare pathways A to B). *V. cholerae* Biotype El Tor strains that lack a functional LpxN are sensitive to polymyxin B, a phenotype that has also been observed in a number of gram-negatives. However the >100 fold change in CAMP sensitivity is an order of magnitude higher, when compared to other secondary acyltransferase mutants in other organisms like *Klebsiella, Escherichia, Salmonella,* and *Acinetobacter*[11,19–21]. The real reason for sensitivity of LpxN mutants is that it transfers the hydrocarbon acyl chain, which provides the hydroxyl site of glycine esterification by AlmG.

In this chapter the molecular biology of AlmG (Vc1577) is explored, providing evidence to support its hitherto putative role as a glycyl to Kdo-lipid A transferase. Evidence of AlmG glycyl to Kdo-lipid A transferase activity is demonstrated *in vivo* by heterologous expression in a specially engineered *E. coli* strain. As AlmG is the only known aminoacyl transferase in the lysophospholipid acyltransferase superfamily the characterization of its activity *in vitro* was sought. A hexahistidine tagged AlmG construct when expressed in *V. cholerae* biotype El Tor Δ*almG* restores polymyxin B resistance of an otherwise sensitive Δ*almG* mutant, demonstrating *in vivo* functionality of

tagged protein constructs. Despite the lack of predicted transmembrane domains, AlmG most likely functions and associates with its lipid substrate at the inner membrane. Evidence for membrane association and a multi-step purification for hexahistidine tagged AlmG concludes this chapter.

4.2 RESULTS

4.2.1 In vivo evidence for glycyl to lipid A transferase activity upon heterologous expression of AlmG, and requisite glycyl-lipid A pathway enzymes, in an engineered strain of *E. coli*.

As shown in Figure 4.1, there are differences in late steps of Kdo-lipid A biosynthesis upon comparison of *E. coli* and *V. cholerae* pathways, which result in chemically distinct product lipids. Among these differences is acyl chain length (acyl chain lengths are denoted in the product structures in Figure 4.1 pathway A or B), where the primary acyltransferase LpxA homolog in *V. cholerae* is specific for hydroxylauroyl (3-OH C_{12}; Figure 4.1 Panel B, Positions 3 and 3´) in contrast with the *E. coli* LpxA enzyme that prefers longer, hydroxymyristoyl groups (3-OH C_{14}; Figure 4.1 Panel A, Positions 3 and 3´). The secondary acyltransferase LpxL homolog of *V. cholerae* also differs from its counterpart in *E. coli*, adding myristoyl groups (C_{14}; Figure 4.1 Panel B, attached to hydroxyacyl chain at 2´), whereas the *E. coli* LpxL prefers lauroyl acyl chains (C_{12}; Figure 4.1 Panel A, attached to hydroxyacyl chain at 2´). As discussed in section 4.1, LpxL enzymes of *E. coli* and *V. cholerae* also differ in substrate specificity, where Kdo_2-lipid A is the major product in *E. coli,* and phosho-Kdo-lipid A is produced as the major product in *V. cholerae*. Most characterized Kdo-lipid A modification enzymes display some degree of promiscuity in substrate lipid selection, where modest alterations in acyl chain and Kdo domain composition do not appear to dramatically affect activity[10,25,26]. However, it remained possible that AlmG might require idiosyncratic features of the *V. cholerae* pathway, as with the *V. cholerae* LpxL requirement for a

phosphorylated Kdo domain[24]. With this in mind, an *E. coli strain* producing a minimal, model substrate Kdo-lipid A domain was engineered to better assess AlmG glycyltransferase activity.

E. coli strain design began with a Δ*lpxM* acyltransferase mutant that produces a penta-acylated Kdo-lipid A domain, lacking the typical secondary myristoyl acyl chain (Figure 4.1 Panel A, penultimate structure). *V. cholerae* LpxN can replace LpxM activity in *E. coli*[19], where the final lipid product contains a hydroxylauoryl group at the 3″, to provide the substrate hydroxyl group for AlmG glycine addition[27]. Deletion of *lpxT* was also performed, as this modification is dispensable for growth in *E. coli,* is not observed in *V. cholerae,* and can make downstream analysis of ^{32}P-radiolabeled lipids by thin-layer chromatography difficult (Chapter 6 for strain construction). Mutants in the Kdo_2-lipid A heptosyltransferase pathway of *E. coli*, responsible for initiating LPS inner core oligosaccharide synthesis, are viable[28,29], and result in production of a minimal Kdo_2-lipid A product lacking core-oligosaccharide altogether (Figure 4.2 Lipid structure top left). In the *E. coli* Δ*lpxM* Δ*lpxT* mutant we were able to cleanly delete the entire *rfaDFC* operon (Figure 4.2 operon diagram top right and pathway), using a recombineering based approach[30] (Chapter 6 for Methods).

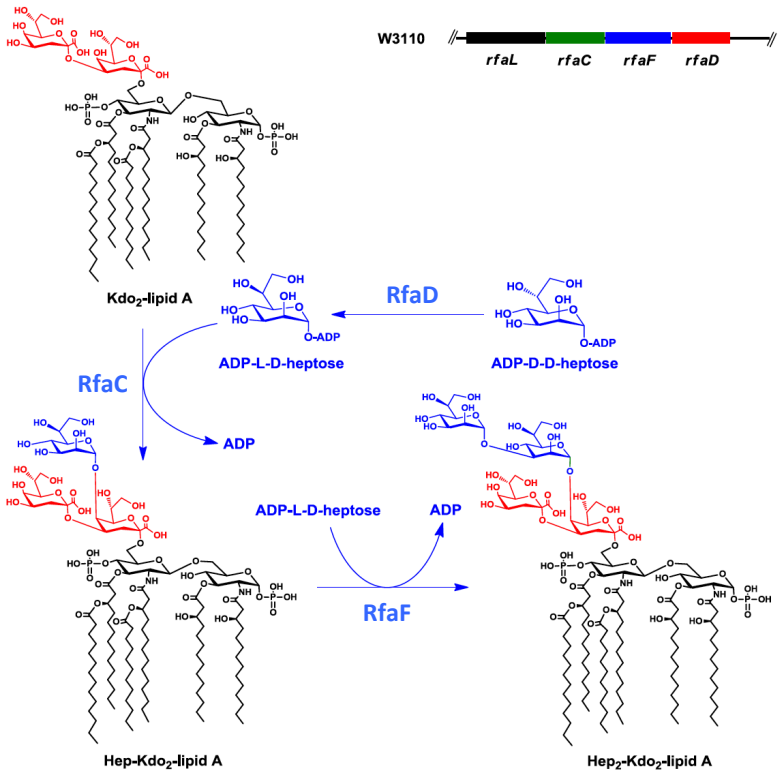


Figure 4.2 Chemical steps in *E. coli* LPS biosynthesis after formation of Kdo-Lipid A domains; inner core oligosaccharide heptose addition by the *rfaDFC* operon.[6]

After synthesis of Kdo_2-lipid A domains in *E. coli* (top left; structure), heptose sugars are added by *rfaDFC* operon encoded enzymes (top right; diagram of operon). RfaD is a heptose-ADP ribosyltransferase, chemically activating heptose sugars for transfer to substrate alcohol groups. In *E. coli* heptosyl addition to Kdo_2-lipid A or heptose-Kdo_2-lipid A are catalyzed by the heptosyltransferases RfaC and RfaF, respectively. Pathway enzymes and chemical steps are highlighted and attached heptose sugars are colored blue, Kdo sugars in red. The oligosaccharide portion of heptosylated-Kdo-lipid A molecules (bottom right structure: Hep_2-Kdo_2-lipid A) is further extended by other enzymes that add additional heptose and hexose sugars to form core-oligosaccharide Kdo-lipid A domains. The RfaDFC pathway and core-oligosaccharide steps occur on the cytoplasmic leaflet of the inner membrane bilayer. Core-oligosaccharide Kdo-lipid A is further processed to mature LPS in the periplasm, upon O-antigen attachment, with ultimate deposition of LPS on the bacterial surface (as in Figure 1.4). *E. coli* K-12 does not attach O-antigen due to a mutation in the *rfb* cluster.

---

[6] Figure adapted from original publication: Wang, J. *et al.*, Construction and characterization of an *Escherichia coli* mutant producing Kdo_2-Lipid A. *Mar. Drugs*. **12**,1495-1511 (2014).

Verification that minimal Kdo_2-lipid A species are produced by the engineered *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC* was initially assessed by comparison of full length versus saccharide truncated LPS species using SDS-PAGE/ProQ Emerald based detection (Chapter 6 for Methods). Evidence for truncation was readily apparent in comparison of LPS isolated from *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC* versus a strain that produces extended "core"-LPS such as *E. coli* Δ*lpxM* Δ*lpxT* (Figure 4.7; compare lane 1 "full length" or lanes 2-4 "core"-LPS to lane 5 "Kdo_2-lipid A" isolated from *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC*).

With observation of minimal Kdo_2-lipid A production it was then hypothetically possible to co-isolate Kdo_2-lipid A with glycerophospholipids using a simple two-phase chemical extraction (Chapter 6 for method details). This would provide a more efficient workflow that circumvents multiple, additional steps typically used to isolate and characterize lipid A material from gram-negatives[31]. Thin layer chromatography analysis of ^{32}P radio-isotopically labeled lipid material, verified this hypothesis, where material from the simpler chemical extraction method showed lipid material in *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC* not present in *E. coli* Δ*lpxM* Δ*lpxT* (Figure 4.3 compare lane 4 to lane 3 respectively). The new lipid material, or "spots", correspond to Kdo_2-penta-acylated(Ac_5)lipid A, or phosphoethanolamine modified variants, as supported by structural analysis of this isolated lipid material by MALDI-MS (Figure 4.4; top two spectra correspond to Figure 4.3 lanes 3 and 4 respectively). Kdo_2-Ac_5-lipid A has an expected singly charged molecular ion mass of 2026.14 *m/z* and is evident in *E. coli*

Δ*lpxM* Δ*lpxT* Δ*rfaDFC* lipid material, but not *E. coli* Δ*lpxM* Δ*lpxT* as indicated in Figure 4.4. The phosphoethanolamine modified lipid A material was also observed by previous groups who generated disrupted heptosyltransferase mutant strains[28], and are consistent with the +123 *m/z* shifted species in the mass spectra of *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC* (Figure 4.4). Molecular ions in the range of 1340-1410 *m/z* correspond to cardiolipin molecular species, an expected glycerophospholipid co-isolated by this method, and also seen near the TLC solvent front as demarcated in Figure 4.3; the major glycerophospholipids phosphatidylglycerol and phosphatidylethanolamine can also be observed near the solvent front (Figure 4.3), but are of molecular masses lower than the range of detection in MS spectra presented in Figure 4.4.

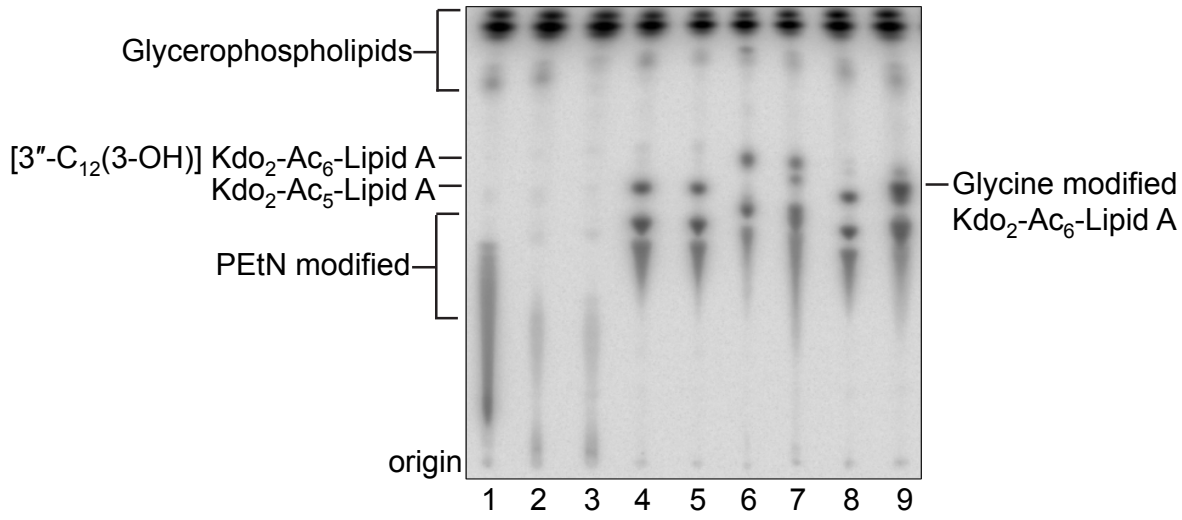


Figure 4.3 TLC analysis of ^{32}P-labeled lipids from deep-rough *E. coli* strains that produce Kdo$_2$-lipid A species and evidence for glycine modification by the *V. cholerae* AlmEFG pathway in *E. coli.*

Phosphate containing lipids of engineered *E. coli* strains were radio-isotopically labeled by supplementing exponential phase bacteria cultures with ^{32}P. All strains were grown in G56, a low phosphate (150 μM) minimal media, supplemented with 1 mM isopropyl β-D-1 thiogalactopyranoside (IPTG). Lipid material was extracted and isolated using the method of Bligh and Dyer, and resolved on silica gel plates by thin-layer chromatography (TLC; solvent system - chloroform: pyridine: 88% formic acid: water, 30:70:16:10, v:v:v:v). Text labels depict origin where lipid material was spotted before TLC separation, as well as proposed lipid species of interest. Lanes represent lipid material from (*1*) *E. coli* K12 W3110, and derived mutant strains: (*2*) Δ*lpxM*; (*3*) Δ*lpxM*Δ*lpxT*; (*4-9*) Δ*lpxM*Δ*lpxT*Δ*rfaDFC*. Strains of W3110 Δ*lpxM*Δ*lpxT*Δ*rfaDFC* (lanes *4-9*) harbor pQlink plasmids (*4*) no plasmid control; (*5*) empty pQlink vector control, or express the following *V. cholerae* genes (*6*) *lpxN*; (*7*) *lpxN, almEFG*; (*8*) *almEFG*; (*9*) *lpxN, almEFG, vc2457*.

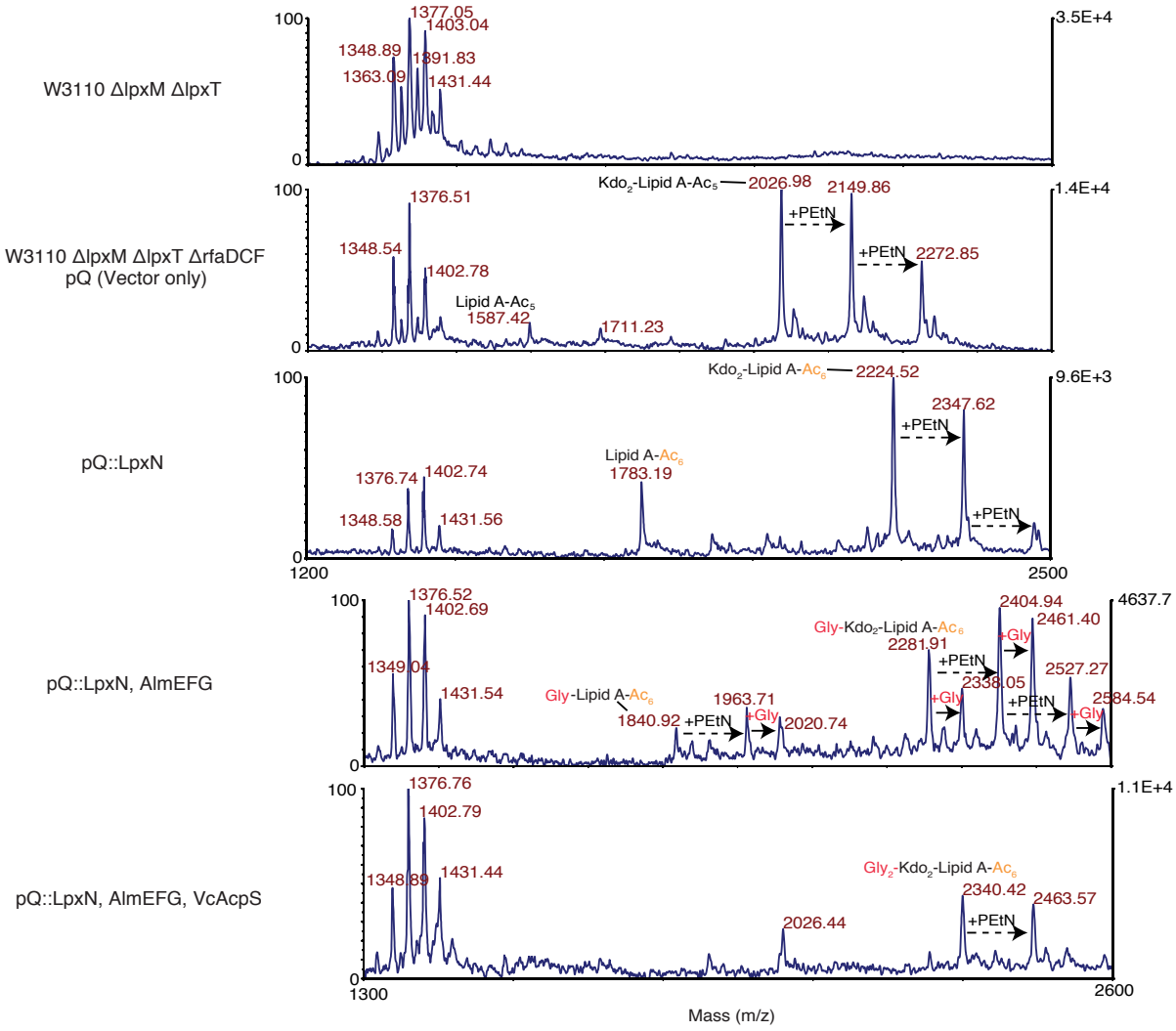


Figure 4.4 MALDI-MS structural analysis of lipids from designer deep-rough *E. coli*, where strains expressing glycine modification machinery show evidence of lipid A glycine modification.

MALDI-MS analyzed lipids were isolated from cultures grown without ^{32}P (as in Figure 4.3 figure legend; see Chapter 6 for details). Strain information is listed to the left of each spectrum, where the bottom four spectra are all from the same W3110 Δ*lpxM* Δ*lpxT* Δ*rfaDFC* background harboring different pQlink expression plasmids. Lipids analyzed here correspond to numbered lanes in Fig. 4.3: (*3*) W3110 Δ*lpxM* Δ*lpxT*; (*5*) W3110 Δ*lpxM* Δ*lpxT* Δ*rfaDFC* pQlink no gene; or W3110 Δ*lpxM* Δ*lpxT* Δ*rfaDFC* with pQlink plasmids carrying *V. cholerae* (*6*) *lpxN*; (*7*) *lpxN, almEFG*; or (*9*) *lpxN, almEFG, vc2457*. Left y-axis denotes relative intensity, right y-axis denotes total counts, and the x axis is of the *m/z* range labeled.

AlmEFG pathway, and other necessary enzymes, were co-expressed in the minimal Kdo_2-lipid A producing *E. coli* strain, to determine whether glycine modification was possible in an organism other than *V. cholerae*. It also provided a great system to examine how specific AlmG was for the particular lipid A structure produced by its host organism. Co-expression constructs were created using the pQlink IPTG inducible plasmid, which enables expression of multiple gene products simultaneously, as described and used in Chapter 2 to produce holo- or gly-AlmF in *E. coli* (Chapter 6 for detailed methods). Expression of LpxN resulted in efficient turnover of penta-acylated Kdo_2-lipid A (Kdo_2-Ac_5-lipid A) to a hexa-acylated hydroxylauroyl containing species in *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC* (Figure 4.3 compare lane 5 and 6 respectively; Figure 4.4 compare spectra pQ versus pQ::LpxN expected mass shift +198 *m/z*). Evidence of LpxN activity has been previously reported in a different strain of *E. coli*[19]. Co-expression of AlmEFG and LpxN in *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC* results in expected chemical migration shifts in TLC analysis of of Kdo-lipid A material material. Glycine results in the addition of a positively charged residue, which in this TLC solvent system results in migration with lower mobile phase retention factors (spot shifts downward); as similarly observed with phosphoethanolamine modified species (Figure 4.3; see TLC legend). The shift in phosphoethanolamine modified species is more dramatic than glycine modification, as might be expected, due to the zwitterionic nature of phosphoethanolamine.

Expression of the trio of AlmEFG proteins, without LpxN, did not affect migration of radiolabeled lipid species (Figure 4.3; lane 8), verifying that the LpxN

transferred hydroxylacyl chain, is the site of glycine esterification, and is required for proposed AlmG glycyltransferase activity. Similarly no apparent change was observed in solvent migration of glycerophospholipids near the solvent front (Figure 4.3; lane 7 or 9), or in MALDI-MS structural analysis of cardiolipin (Figure 4.4; no observed +57 *m/z* shift in 1340-1410 *m/z* peaks). This result indicates that AlmG transfer of glycine to hydroxyl groups on lipids, is specific towards Kdo-Lipid A domains. Co-expression of Vc2457, the AlmF specific phosphopantetheinyltransferase (Chapter 2), appeared to improve overall pathway efficiency, as an increase in glycine modification was evident by TLC (Figure 4.3 compare lane 7 to lane 9) and by MALDI-MS (Figure 4.4 bottom spectra). Interestingly some penta-acylated species is also observed when the pathway flux is apparently pushed towards increased glycine addition (Figure 4.3 lane 9 compared to lane 5; Figure 4.4 bottom spectra; see Section 4.3 Discussion).

Demonstration of AlmEFG pathway functionality and AlmG glycltransfer to Kdo_2-lipid A in *E. coli* provides the first observation of its kind in an organism other than *V. cholerae*. With this in mind, and as isolation of lipid A from *V. cholerae* is incredibly difficult[24,27,31,32]. *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC* was then further engineered to produce "pure" Kdo_2-lipid A substrate, by deleting the phosphoethanolaminetransferases *eptA* and/or *eptB*, for potential use in *in vitro* AlmG assays. The simplest way to generate *E. coli* Δ*lpxM* Δ*lpxT* Δ*rfaDFC* strains with additional Δ*eptA* and/or Δ*eptB* mutation(s), required starting with *E. coli* Δ*lpxM* Δ*lpxT* (Chapter 6 Methods for details). Subsequent deletion of either or both *ept* genes was performed, followed by a final recombineering step to introduce the Δ*rfaDFC* mutation. To completely eliminate phosphoethanolamine

modified Kdo-lipid A species both Δ*eptA* and Δ*eptB* was required (Figure 4.5 Panel A; compare ΔΔA/B to single mutants ΔA or ΔB). Deletion of either *eptA* or *eptB* or both genes did not affect isolation efficiency of truncated Kdo-lipid A species (Figure 4.7 lanes 6-8 compared to lane 5). Further purification of isolated Kdo-lipid A material by removal of contaminant glycerophospholipids was possible via diethylaminoethyl-cellulose chromatography (Figure 4.5 Panel B compare lane QN to QN Pure). Expression of AlmEFG enzymes, and other required components, continued to demonstrate AlmG functionality as a glycyltransferase (Figure 4.6 bottom spectra compared to controls above). In sum this evidence demonstrates the possibility of isolating "pure" Kdo-lipid A AlmG substrate from *E. coli*, of higher purity and in a manner significantly more efficient than can be performed with *V. cholerae*[24,27,31,32].

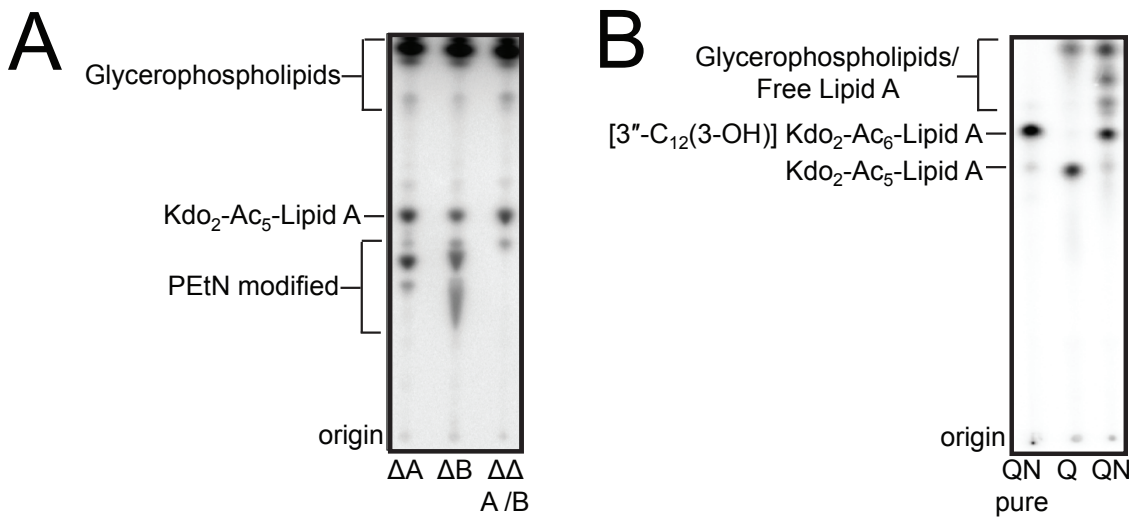


Figure 4.5 Generation of deep rough *E. coli* that produce phosphoethanolamine deficient Kdo_2-lipid A to simplify analysis of TLC separated lipids.

Phosphate containing lipids of engineered *E. coli* strains were radio-isotopically labeled by supplementing exponential phase bacteria cultures with ^{32}P. All strains were grown in G56, a low phosphate (150 μM) minimal media. Lipid material was extracted and isolated using the method of Bligh and Dyer[19], and resolved on silica gel plates by thin-layer chromatography (TLC). Text labels depict origin where lipid material was spotted before TLC separation, as well as proposed lipid species of interest. Results in (**A**) were obtained using the same solvent system in Fig. 4.3 (chloroform: pyridine: 88% formic acid: water, 30:70:16:10, v:v:v:v). Lanes represent lipid material from W3110 Δ*lpxM*Δ*lpxT*Δ*rfaDFC* with additional deletion of the lipid A phosphoethanolamine transferases (*ΔA*) Δ*eptA*; (*ΔB*) Δ*eptB*; or (*ΔΔ A/B*) Δ*eptA*Δ*eptB*. Results in (**B**) were obtained using an alternative solvent system (chloroform: methanol: glacial acetic acid: water, 25:15:4:4, v:v:v:v) and show Kdo_2-lipid A lipid material from different source strains. Lanes represent W3110 Δ*lpxM*Δ*lpxT*Δ*rfaDFC*Δ*eptA*Δ*eptB* strains harboring pQlink plasmids expressing *V. cholerae* (*QN pure*) *lpxN* where lipids were further purified by diethylaminoethyl-cellulose chromatography (Results, and Chapter 6 for details); (*Q*) no gene; (*QN*) *lpxN*. Bacterial cultures analyzed in (**B**) were supplemented with 1 mM isopropyl β-D-1 thiogalactopyranoside (IPTG).

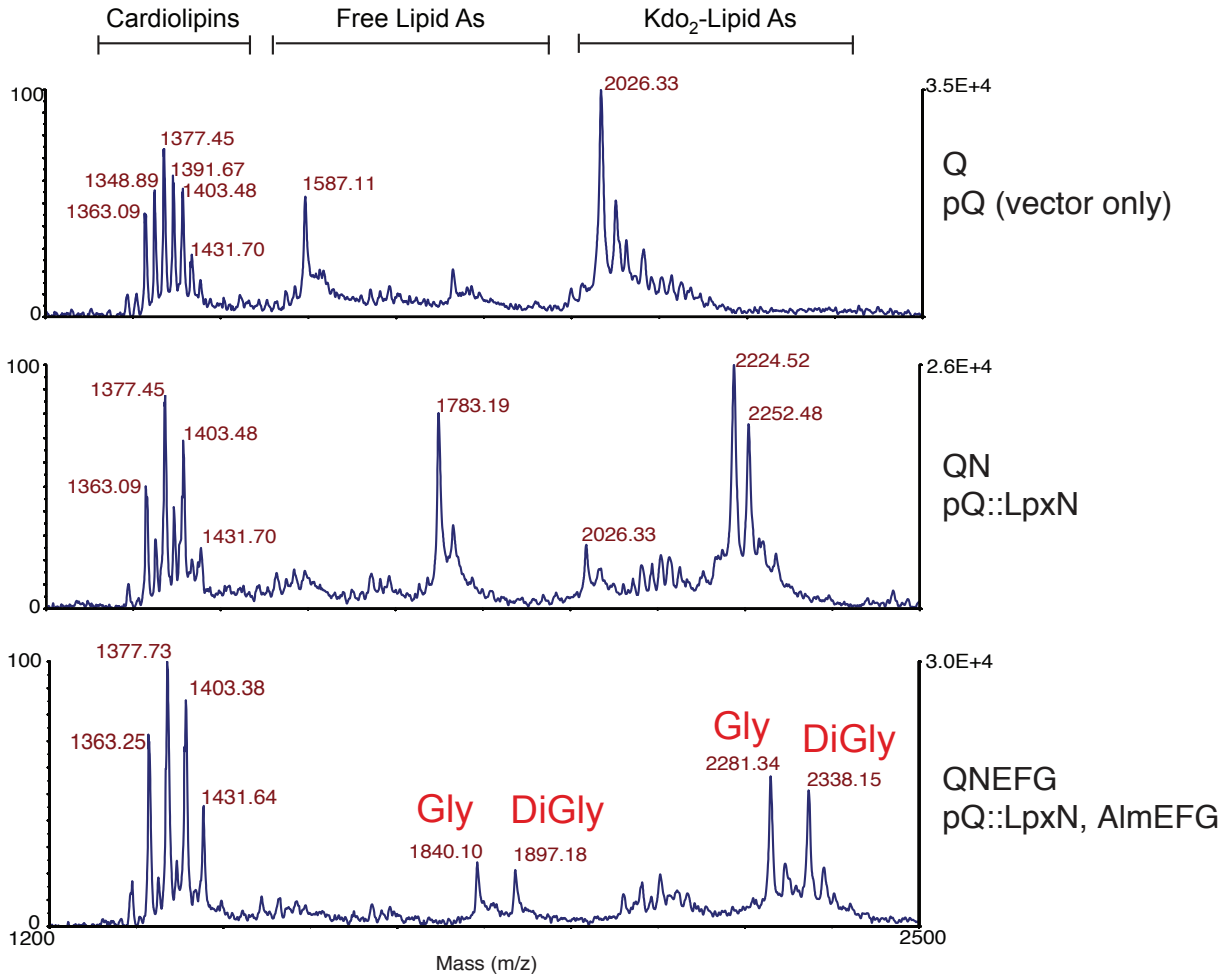
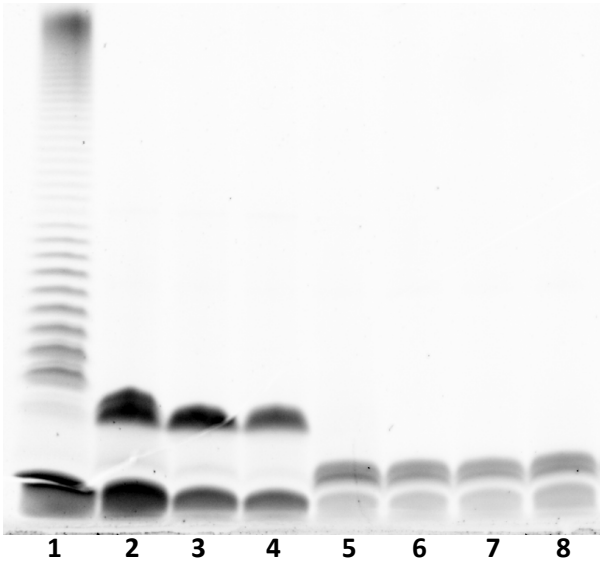


Figure 4.6 Generation of deep rough *E. coli* that produce phosphoethanolamine deficient Kdo_2-lipid A for improved MALDI-MS analysis of glycine modified species.

E. coli strains were grown in G56, a low phosphate (150 μM) minimal media, supplemented with 1 mM isopropyl β-D-1 thiogalactopyranoside (IPTG). Lipid material was extracted and isolated using the method of Bligh and Dyer[33], and analyzed by MALDI-MS (Chapter 6 for details). All spectra are from the same background strain of *E. coli* W3110 Δ*lpxM*Δ*lpxT*Δ*rfaDFC*Δ*eptA*Δ*eptB*. Lipids analyzed here correspond to labeled lanes in Fig. 4.5 Panel B, however, MALDI-MS analyzed lipids were isolated from cultures grown without ^{32}P. Corresponding lanes from Fig. 4.5 Panel B include: W3110 Δ*lpxM*Δ*lpxT*Δ*rfaDFC*Δ*eptA*Δ*eptB* strains harboring pQlink plasmids expressing (*Q*) no gene; (*QN*) *lpxN*; or (*QNEFG*) *lpxN almEFG*. Left y-axis denotes relative intensity, right y-axis denotes total counts, and the x-axis is of the *m/z* range labeled.



 # - Strain
1 - *S.* Typhimurium LT2
2 - *E. coli* K12 W3110
3 - W3110 ΔlpxM
4 - W3110 ΔlpxM, ΔlpxT
5- W3110 ΔlpxM, ΔlpxT, ΔrfaDFC
6 - W3110 ΔlpxM, ΔlpxT, ΔrfaDFC, ΔeptA
7 - W3110 ΔlpxM, ΔlpxT, ΔrfaDFC, ΔeptB
8 - W3110 ΔlpxM, ΔlpxT, ΔrfaDFC, ΔeptA, ΔeptB

Figure 4.7 Engineered deep-rough *E. coli* that produce LPS lacking core oligosaccharide, the minimal predicted AlmG substrate Kdo_2-lipid A.

Each numbered lane below the LPS gel (prepared according to methods in Chapter 6) corresponds to strains listed on the right. *Salmonella enterica* serovar Typhimurium LT2 is included as a control (lane 1), to show a typical O-antigen repeat pattern observed in LPS from many gram-negatives[9]. Laboratory *E. coli* K12 strains, including W3110 (lane 2), synthesize a truncated LPS molecule due to mutations in O-antigen synthesis genes[34]. Smaller LPS molecules migrate further towards the bottom of these gels. All Δ*rfaDFC* mutant strains (compare lanes 2-4 with lanes 5-8) synthesize an even more truncated molecule, corresponding to the lack of core-oligosaccharide (see Figure 4.2).

4.2.2 Histidine-tagged AlmG fusions are functional in vivo, peripherally associate with membranes, and can be purified using a multi-step protocol in *E. coli*.

With the goal of AlmG assay development, to monitor proposed glycine to Kdo-lipid A transfer *in vitro*, overexpression and purification of functional AlmG protein was sought. Hexahistidine tagged AlmG fusions were constructed and assayed for *in vivo*

functionality in *V. cholerae* biotype El Tor Δ*almG* mutants. N-terminal or C-terminal hexahistidine fusions showed minimal disruption of AlmG glycyl-Kdo-Lipid A modification (Figure 4.8), as correlated with rescue of polymyxin B resistance in His_6-AlmG complemented Δ*almG* strains, that are otherwise sensitive to CAMPs (Figure 4.8 middle picture). Subsequent introduction and optimized overexpression of tagged AlmG in host *E. coli* was achieved (Chapter 6 methods).

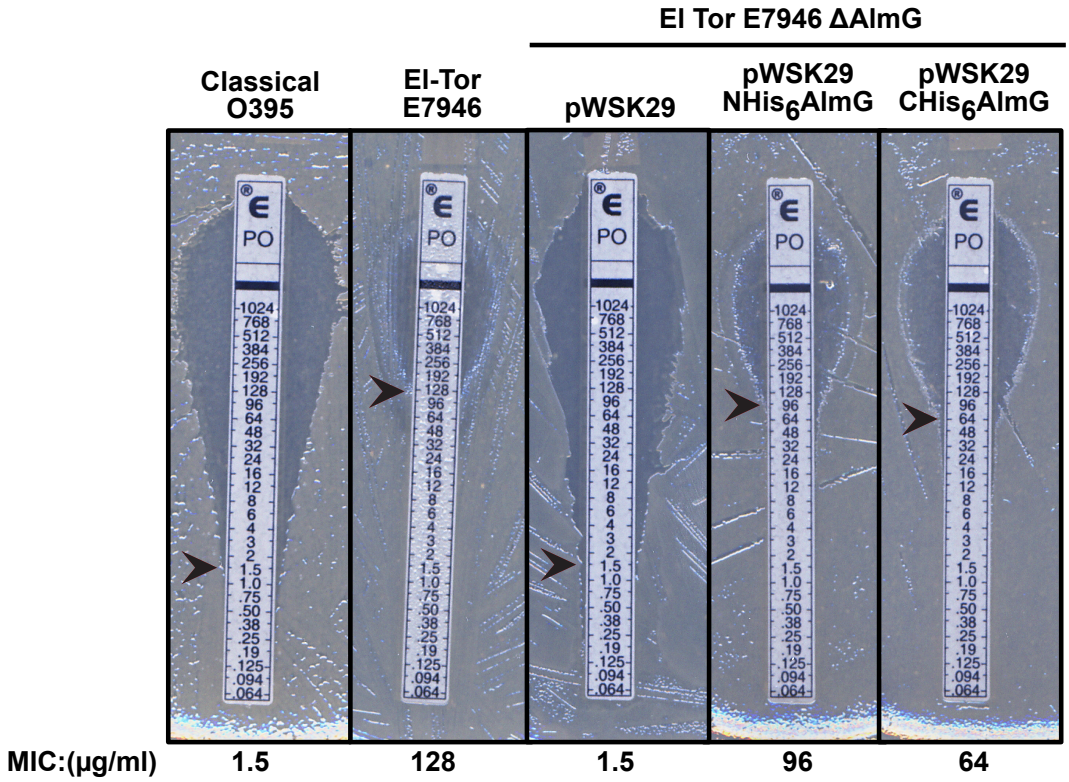


Figure 4.8 Hexahistidine fusion to AlmG does not interfere with proposed glycine lipid A acyltransferase function in *V. cholerae*.

Plasmid construction, *V. cholerae* expression conditions, and details of bacterial growth are detailed in Chapter 6. In brief, *V. cholerae* strains are grown to mid-exponential phase before dilution and sterile application to LB agar plates supplemented with 1 mM IPTG, to induce expression of plasmid encoded AlmG. Pictured above are polymyxin B E-strip® tests with increasing concentrations of CAMP (from bottom of the white indicator strip to the top). Confluent bacterial growth occurs at antibiotic levels below the minimum inhibitory concentration (MIC). Arrows designate the determined MIC value, as reported below each image. Strain information is provided above each image.

AlmG is not predicted to contain transmembrane domains or a canonical signal peptide for periplasmic or extracellular localization[35,36]. Based on this information, in combination with a putative activity on both a lipid substrate and a cytosolic carrier protein, the hypothetical peripheral association of AlmG with the cytoplasmic leaflet of

the inner membrane bilayer was investigated. Membranes from *E. coli* overexpressing N-terminally hexahistidine AlmG fusions (N-His_6-AlmG) were isolated, and washed of non-specific contaminants. Various detergents were used to attempt solubilization of AlmG from membrane material, where membranes were ultimately separated from detergent soluble material via ultracentrifugation (Chapter 6 for methods). Ni-resin chromatography was used to capture any histidine tagged AlmG in detergent solubilized material. As predicted, multiple, but not all detergent solubilized fractions enabled isolation of N-His_6-AlmG, to provide convincing evidence of AlmG association with the bacterial membrane (Figure 4.9; lanes Br, DM, DDM, LD). With this evidence in mind, a multi-step purification protocol was successfully designed and employed to purify N-His_6-AlmG from *E. coli* membranes (Figure 4.10; Pure protein in Lane SP).

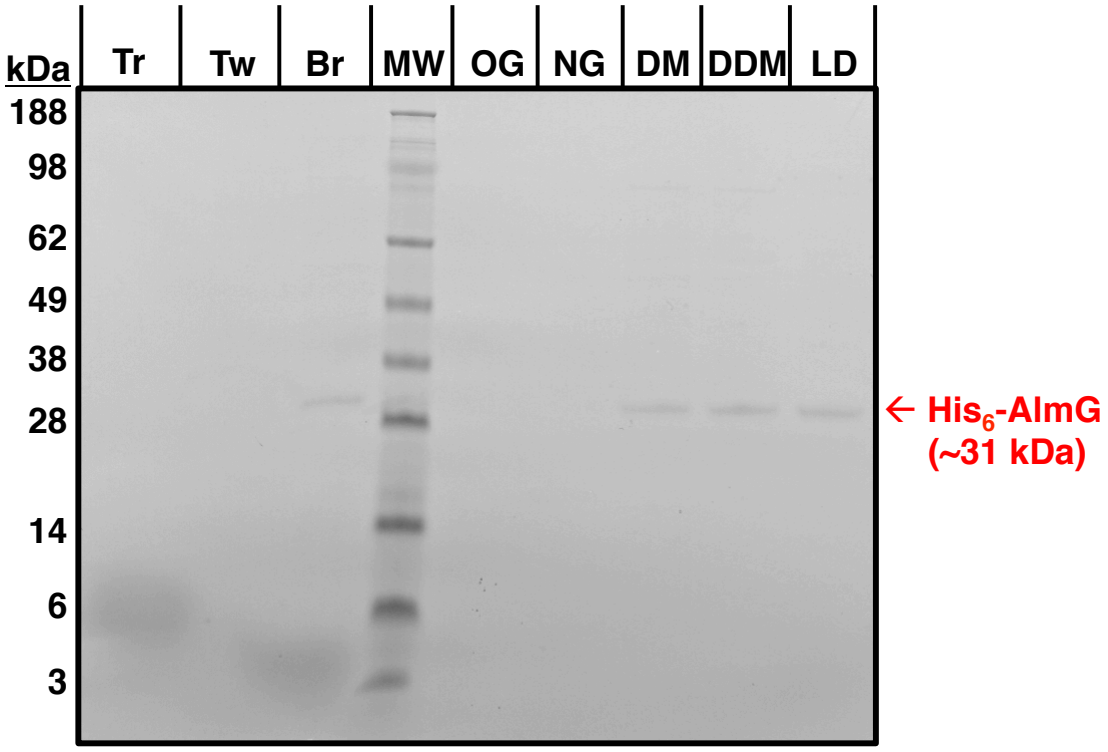


Figure 4.9 AlmG can be extracted by detergent solubilization of washed membrane preparations from host *E. coli* C43(DE3)pLysS.

Membranes from *E. coli* overexpressing N-terminal hexahistidine-tagged AlmG (N-His_6-AlmG) were isolated as detailed in Chapter 6. Detergent exchange and subsequent affinity capture of solubilized N-His_6-AlmG is also described in Chapter 6. Protein samples were analyzed by SDS-PAGE and visualized with coomassie dye. The various detergents used in solubilization conditions are listed in lanes above the pictured gel: (Tr) Triton X-100, (Tw) Tween-20, (Br) Brij-58, (OG) octyl-glucoside, (NG) nonyl-glucoside, (DM) decyl-maltoside, (DDM) dodecyl-maltoside, or (LD) lauryl-dimethylamine oxide. The observed protein band corresponding to the expected mass of N-His_6-AlmG (31 kDa), shows Br, DM, DDM, and LD can solubilize peripherally associated N-His_6-AlmG from host *E. coli* membranes.

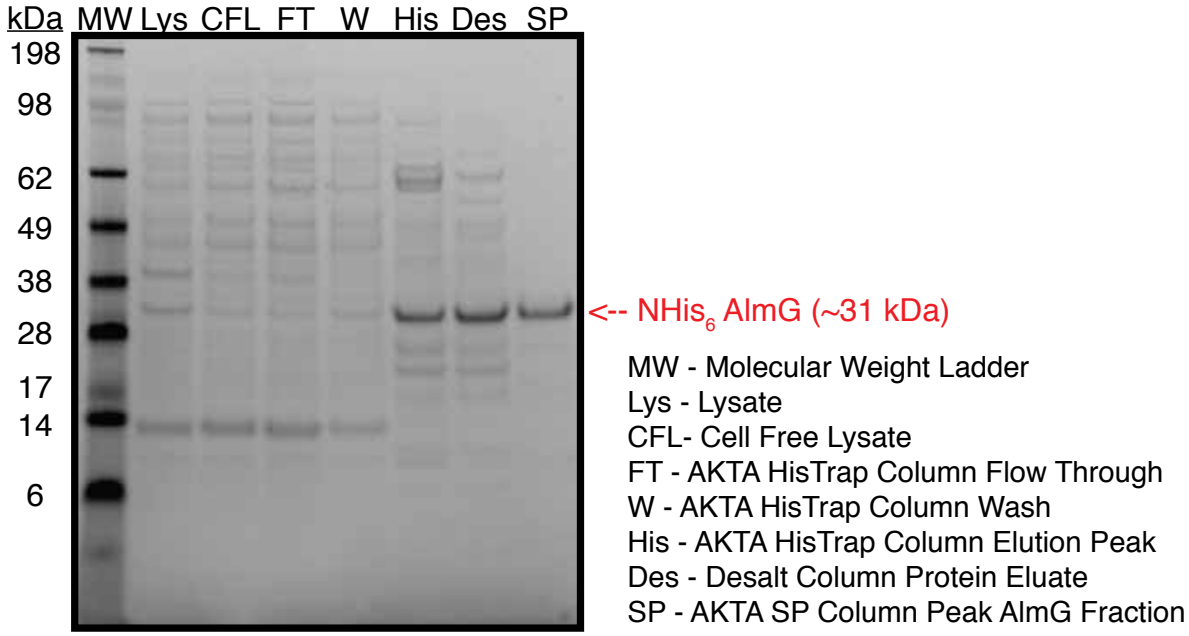


Figure 4.10 Multi-Step Purification of overexpressed $NHis_6$-AlmG from an *E. coli* C43(DE3)pLysS host strain.

N-terminal hexahistidine-tagged AlmG (N-His_6-AlmG) was overexpressed in *E. coli* C43(DE3)pLysS and subjected to a multi-step protein purification protocol (see Chapter 6 for strain construction, protein expression conditions, and detailed purification procedures). Protein aliquots were taken at multiple steps along the purification scheme. Each individual fraction was separated by SDS-PAGE, and visualized in gel by coomassie staining. Each lane contains 10 μg of protein as determined by BCA assay[37]. Lanes are labeled as indicated in the key listed along the right (Chapter 6), to obtain >95% pure $NHis_6$-AlmG.

4.3 DISCUSSION

4.3.1 AlmG is a membrane associated glycyl to Kdo-lipid A transferase, an unexpected activity based on its homology to characterized lipid A secondary acyltransferases.

Evidence in this chapter provides data to support the hypothesis that AlmG functions as glycyl-Kdo-Lipid A transferase. This is a unique and novel activity that cannot be predicted on the basis of currently annotated sequence homology in public databases. Prediction of AlmG membrane association was based on the reasonable assumption that its lipid substrate is a component of the membrane. Yet AlmG does not contain predicted transmembrane helical regions[35], and thus is unlikely to function as an integral membrane protein. In support of its peripheral membrane association, general hydrophathy of AlmG is relatively high, with multiple secondary structure hydrophobic clusters predicted[38], in combination with a very basic theoretical pI (9.49). These features hint at favorable interactions with amphipathic phospholipids (i.e. acidic and hydrophobic) at the inner bacterial membrane. Finally as Gly-AlmF is the putative donor molecule in glycyl-Kdo-Lipid A formation, AlmG's residence on the cytoplasmic face of the inner membrane bilayer is most likely. The detergent exchange assay (Figure 4.9), provides convincing data to support the peripheral membrane hypothesis, but does not provide more detailed membrane topology data. As *E. coli* does not otherwise produce glycine modified Kdo-lipid material, MALDI-MS and TLC analysis of lipid material from *E. coli* hosts expressing AlmG (Figures 4.3, 4.4, and 4.6), and other necessary

enzymes, along with demonstrated AlmG membrane association (Figure 4.9), supports the assignment of *V. cholerae* AlmG as a *bona fide* glycyl-Kdo-lipid A transferase.

Members of the expansive lysophospholipid acyltransferase superfamily typically feature a HX_4D/E catalytic dyad essential for acyl transfer activity[1]. AlmG contains two such candidate active site dyads ($H_{106}X_4D_{111}$ and $H_{215}X_4E_{220}$). The active site dyad in most lyso-phospholipid acyltransferase superfamily members is positioned within a hydrophobic cleft[39–41]. Thus, the dyad at position $H_{106}X_4D_{111}$ is the most likely active site candidate, and aligns similarly in position relative to primary sequences of other lipid A late acyltransferases[42]. Investigation of the function of candidate active sites within AlmG and further characterization of its unusual activity is ongoing. Numerous assay conditions have been sampled for in vitro characterization of AlmG. Purified AlmG protein (Figure 4.10) with purified Kdo-lipid A substrate (Figure 4.5 Panel B), or from source enzyme material isolated from cells where AlmG has been functionally verified (as in material from *V. cholerae* overexpressing AlmG in Figure 4.8), have all been investigated under a variety of parameters: temperature, pH, detergent, or divalent cation concentrations, as required by other, related enzymes[8,24,41]. No *in vitro* activity has *yet* been observed, particularly with TLC based assay detection (as in Figures 4.3 or 4.5), despite clear indication of *in vivo* activity (Figure 4.3 lane 7 compare to lane 6, or Figures 4.4 and 4.6).

Co-expression of Vc2457, the AlmF specific phosphopantetheinyltransferase (Chapter 2), appeared to improve overall pathway efficiency, as an increase in glycine modification was evident by TLC (Figure 4.3 compare lane 7 to lane 9) and verified by MALDI-MS (Figure 4.4 bottom spectra). Vc2457 was previously shown to increase holo-

AlmF formation in *E. coli* (Figure 2.4), which when AlmE is co-expressed should increase the net available Gly-AlmF *in vivo*. An increase in Gly-Kdo-lipid A formation is not unexpected as the chemical law of mass action would predict that an increase in the availability of substrate Gly-AlmF would increase Gly-Kdo-lipid A product formation. Interestingly some penta-acylated species is also observed when the pathway is pushed towards increased glycine addition (Figure 4.3 lane 9 compared to lane 5; Figure 4.4 bottom spectra; see Section 4.3 Discussion). Although not reported in this manuscript, unpublished experiments in a different host *E. coli* background show similar evidence for deacylase activity against glycine-modified Kdo-lipid A domains resulting in penta-acylated Kdo-lipid A species. The source of this activity remains unknown, but does suggest that *E. coli* membranes *only* tolerate some proportion of glycine modified Kdo-lipid A. The biophysical consequence of glycine modification on membrane stability, and effects on associated molecular components, especially in relation to the functions of outer membrane β-barrel proteins, warrants further investigation.

Interestingly a very recent and compelling report on LpxM from *Acinetobacter baumanii* showed evidence for bifunctional activity (*in press* Metzger, LE et al PNAS 2016). This secondary acyltransferase homolog catalyzes not only the transfer of hydrocarbon acyl chains to substrate Kdo-lipid A domains, but also displays thioesterase activity against co-substrate acyl-ACPs. This thioesterase activity *in vivo* would liberate hydrocarbon acyl chains from acyl-ACPs. While somewhat controversial, this activity could indicate that lysophospholipid acyltransferase members, at least related to *A. baumanii* LpxM, liberate fatty acids for redirection into catabolic pathways, under

starvation gconditions for instance, where hydrocarbon acyl chains are metabolized through β-oxidation pathways. Generally acyl-ACPs are fated for anabolic reactions. *A. baumanii* LpxM thioesterase activity occurs when lipid co-substrates are in absentia (i.e. only acyl-ACP is bound), thus metabolic redirection would occur when lipid substrate is limiting as might occur during conditions of nutrient depletion. AlmG functionality in this, or a similar manner, should be explored.

4.3.2. AlmG is distinct, in structure and activity, from other characterized enzymes that transfer amino acids to lipid molecules deposited on bacterial surfaces.

Gram-positive bacteria specifically neutralize their cell walls by transferring D-alanine to surface teichoic acids, the long poly-phosphoribitol or poly-phosphoglycerol chains linked to glycerophospholipids, or *N*-acetylmuramic acid groups of exposed peptidoglycan, on the surface of organisms that contain these modifications[43] (Figure 1.9; left blue panel). Although the genes *dltBD* of the *dltDABC* operon[44] have long been identified as responsible for the final chemical transfer step, from D-alanine linked carrier proteins (*dltC*) to surface molecules, the precise molecular details of the DltBD reaction have not been elucidated. Recent data has clarified that DltC is not exported[45], as some reports have hypothesized. Continued exploration of this system in gram-positives will hopefully soon clarify these important last steps, of D-alanine transfer to teichoic acids, a physiologically important surface modification in pathogenic *Staphylococcus aureus* and other *Firmicutes* like *Listeria, Bacillus, Enterococcus* or *Streptococcus* spp.[45].

Phosphatidylglycerols aminoacylated with L-arginine, L-lysine, L-alanine, or D-alanine have been discovered and confer resistance toward CAMPs in species of both Gram-type bacteria (e.g., *Staphylococcus aureus* and *P. aeruginosa*)[46] (Figure 1.9; center purple panel). Interestingly the MprF family of enzymes responsible for this activity use aminoacylted tRNAs as substrate donor molecules[47]. Any given MprF homolog is usually specific towards one amino acid, although general specificity with multiple, different amino acids has been observed[48]. In addition to CAMP resistance other phenotypes associated with inactive *mprF* homolog mutants exist, such as increased sensitivity to other classes of antibiotics[47] and susceptibility to osmotic stress[49]. Some organisms encode additional hydrolases capable of removing amino acids attached to phosphatidylglycerol, as recently reported in *P. aeruginosa*[50]. This indicates the exciting possibility of maintenance, where the proportions of aminoacylated versus non-modified phosphatidylglycerols are maintained. It is additionally curious that nearly all gram-type organisms with an MprF like family member, have well-reported symbiotic lifestyles, typically as pathogens or commensals.

Kdo-lipid A glycinylation is a unique gram-negative strategy necessary for resistance to CAMPs[27], which more resembles features of the gram-positive pathway for modifying teichoic acids with D-alanine, than MprF-like activity. AlmEFG catalyzes the adenylation of an amino acid (like DltA), subsequent transfer to a carrier protein (like DltA to DltC), and final transfer to a surface component by dedicated transferase machinery (as DltBD). A final important consideration regarding the activity of AlmG is apparent specificity for Gly-AlmF. Detailed molecular investigation of wild-type *V.*

cholerae AlmE, demonstrated that its substrate holo-AlmF could, in addition to glycine, potentially carry L-Alanine as evidenced by *in vitro* AlmEF characterization (Figure 3.2 Panel B). Currently there is insufficient evidence to support the existence of L-Ala-AlmF *in vivo* (Chapter 3). On the other hand, if L-Ala-AlmF is generated *in vivo,* as possible *in vitro*, AlmG certainly does not use L-Ala-AlmF to decorate Kdo-Lipid A domains. L-Ala-Kdo-lipid A is not apparent either in MALDI-MS analysis of isolated lipid material from *E. coli* (Figures 4.4, and 4.6) or *V. cholerae*[27]. Finally, as diglycine modified Kdo-lipid A species are apparent in both *E. coli* (Figures 4.4. and 4.6) and *V. cholerae*[27], it would appear that some flexibility exists in AlmG substrate specificity, which allows a chemistry of at least two rounds of glycine to Kdo-lipid A transfer, first to Kdo-lipid A, and then to Gly-Kdo-lipid A material. Multiple rounds of amino acid transfer to the same molecule has not been reported with DltBD or MprF-like enzymes.

Chapter 5: Final Thoughts and Conclusions

5.1 Potential AlmEFG-like pathways in other organisms

As the first-of-its-kind observation, amino acid addition to Kdo-lipid A domains of other organisms have not yet been reported. Unfortunately as each protein displays similarity to very broad protein families, homologs of AlmE and AlmG can be found in the genomes of most gram-negative bacteria. Concurrence-constrained AlmEFG homology searches within current databases, identify only *Vibrio nigripulchritudo* as an organism with a similar arrangement of a trio of AlmEFG homologs[1]. Dozens of sequenced *V. nigripulchritudo* strains contain this putative operon (personal communication with Frederique le Roux), providing a high degree of confidence that it may function as AlmEFG does in *V. cholerae*. *V. nigripulchritudo* is an emerging shrimp pathogen with multiple reports of devastating effects on aquaculture yields for a number of commercial operations[2–5]. Unpublished preliminary investigation into the functionality of these AlmEFG homologs from *V. nigripulchritudo* indicates AlmF and AlmG can restore polymyxin B resistance in *V. cholerae almF* disruption mutants and clean deletion Δ*almG* backgrounds, respectively. Further verification of glycine-LPS modification and examination of pathway functionality is ongoing as of the publication of this manuscript.

5.2 GLYCINE MODIFIED GLYCOSIDES OF *V. CHOLERAE*, AND OTHER ORGANISMS

Emerging evidence indicates other extracellular compounds, and surface structures, of gram-negative bacteria contain glycine-modified variants. Extracellular polysaccharide, an important component in *V. cholerae* biofilms, apparently contains an unusual glycine modification[6]. Some extracellular polysaccharide material is found covalently attached to a hitherto unidentified surface molecule. The source enzyme(s) responsible for the atypical polysaccharide glycine modification also remain(s) unknown. It is unlikely that AlmG performs this chemistry, as it has high apparent specificity for Kdo-lipid A domains (Chapter 4). However, the participation of AlmE or AlmF in extracellular polysaccharide glycine modification cannot be ruled out. Similarly, glycine modifications to the inner-core polysaccharide portion of LPS have been identified in a number of pathogenic bacteria including *Campylobacter jejuni*, *Haemophilus influenzae,* and *Shigella flexneri*[7–9]. Thus, remodeling of bacterial surface components via glycine addition appears to be a newly observed emergent phenomenon, where in pathogens this modification is likely important for outer membrane maintenance, stability, and perhaps immune evasion.

5.3 THE FULL REPERTOIRE OF LIPID A MODIFICATIONS OBSERVED IN *V. CHOLERAE*

Although this document has focused on glycine modification of *V. cholerae* Kdo-lipid A domains, other modified lipid A structures in this aquatic pathogen have been observed by our laboratory and others (Figure 5.1). Phosphoethanolamine modification has been reported in *V. cholerae* lipid A structures from O1 and O139 serotypes[10]. *V. cholerae* have a functional *eptA* phosphoethanolamine transferase homolog (unpublished results). The factors that regulate expression of this enzyme in *V. cholerae* and other *Vibrio* species are currently being investigated. Similarly, *V. cholerae* contains an LpxR homolog, responsible for the removal of 3′-acyl-oxy-acyl chains in most organisms. However, it does not appear to be functional in *V. cholerae* and may function with an altogether different chemistry, representing the potential discovery of yet another unique lipid processing enzyme of *V. cholerae*[11–14]. Finally, as has been reported in *V. fischeri,* evidence for phosphatidyl, lyso-phosphatidyl, and glycerophosphate addition to lipid A structures of *V. cholerae* have been observed[15]. The acyl chain composition of covalently attached phosphatidyl groups is of the precise pattern observed in *V. cholerae* glycerophospholipids, indicating that glycerophospholipids serve as phosphatidyl donors in synthesis of these lipid A structures. Conditions where this particular set of modifications becomes more prevalent are in minimal media, where GlcNAc or mannose is supplied as the primary carbon source. The biosynthetic origin of these lipid A species await discovery, and may represent the actual structures of previously reported octa-acylated lipid A material[10]. Additionally further investigation of the biological role of these lipid A modifications, and associated regulatory networks, will surely add

significant evidence for the importance of surface remodeling on *V. cholerae* fitness, pathogenesis, and other phenotypes.

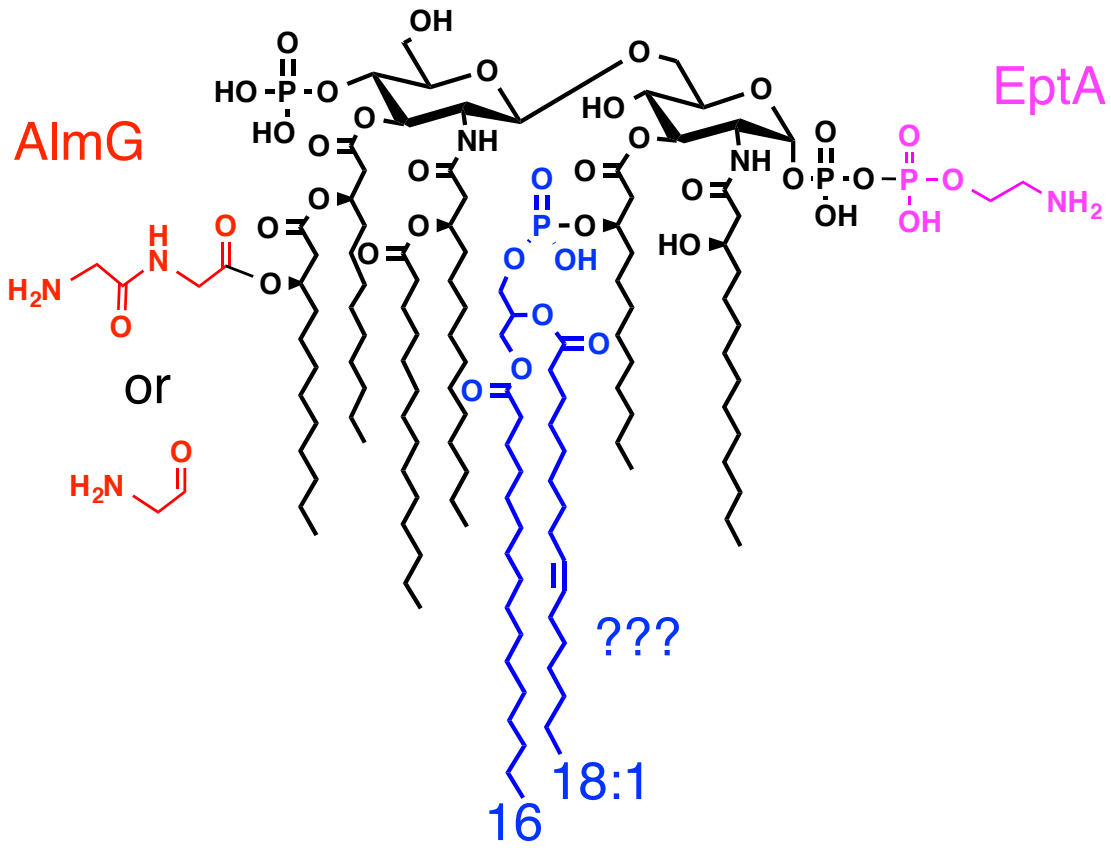


Figure 5.1 The full repertoire of lipid A modifications observed in *Vibrio cholerae*.

In *V. cholerae* El Tor strains, possible modifications include the addition of phosphoethanolamine (pink) by a putative EptA homolog, the addition of glycine or two glycine residues (red) by AlmG, or the presence of phosphatidyl, lyso-phosphatidyl, or phosphoglycerol (blue) moieties of unknown biosynthetic origin. Lipid A species containing all 3 modifications simultaneously have not been observed. The most prevalent modifications observed in bacteria cultured in LB or pH 7.5 minimal media are glycine or diglycine modified species.

5.4 Recent outbreaks of *V. cholerae* biotype El Tor in Kolkata appear to be polymyxin B sensitive

The pathogen *V. cholerae* still affects millions of people around the world every year (WHO), where it is still endemic to regions of sub-Saharan Africa and the Ganges river delta. Recent bacterial sampling in Kolkata, of outbreak strains from infected individuals, have uncovered the emergence of biotype El Tor strains that have apparently lost resistance to polymyxin B[16]. These strains were verified and biotyped based on genetic markers[16]. At the time of this publication, some of these strains are currently under investigation by our lab (gift Dr. Asish Mukhopadhyay). The AlmEFG operon appears intact. Especially important is that *almF* is full length in these polymyxin B sensitive El Tor strains, as assessed by PCR comparison of the entire *almF* open-reading frame, with matched polymyxin B resistant strains. Recall that in the previous pandemic strain, Biotype Classical, *almF* is N-terminally truncated resulting in lack of glycine modified LPS and polymyxin B sensitivity (Chapters 1 and 2). Future characterization of these strains will certainly contribute more sophisticated insight into regulation of glycine modified LPS in this organism, and its affect on human pathogenesis.

5.5 CONCLUSIONS

Biological investigations of *V. cholerae* have long contributed to many, diverse areas of science, as in the specialized fields of ecology, epidemiology, pathology, vaccinology, bacteriology, genetics, cell and molecular biology, and biochemistry. Originally isolated by the Italian anatomist Filippo Pacini in 1854, and rediscovered 30 years later by the German physician Robert Koch, the comma shaped bacilli has remained a provocative organism ever since. The epidemiologist John Snow and Reverend Henry Whitehead, combined statistical demographics with scientific observation during a cholera outbreak in London during 1854, to form a model case, the Broad Street [Pump] Epidemic, foundational to contemporary epidemiology. Their reports also served as evidence, and certainly influenced, the more sophisticated articulation of the germ theory of disease as particularly demonstrated by Louis Pasteur. Thereafter, Dr. Jaime Ferran developed and demonstrated the therapeutic potential of an attenuated oral cholera vaccine preparation in 1885. Despite this initial result, the long-term efficacy of cholera vaccines has always been problematic. Recent recommendation guidelines issued by the WHO call for vaccine stockpiles, particularly for use in situations like what occurred in Haiti (2010), a welcome, long overdue development. Fast-forward to contemporary molecular biology and *V. cholerae* has been used to further explore bacterial toxins, phage-bacteria dynamics, particularly as they impact *Vibrio* virulence, bacteria-host relationships, in humans and with chitinous aquatic organisms, and very recently provided the organism in discovery of the injective type VI secretion apparatus, found and used by numerous bacterial pathogens. The presumption that continued examination of *V. cholerae* will contribute to ever-more exciting biological insights is certainly not unfounded.

Recently, *V. cholerae* has emerged as a leading model organism for membrane biologists, mostly as a consequence of its variable lifestyle. It inhabits multiple aquatic ecological niches and as an enteric pathogen, it's exposed to equally heterogeneous environments after ingestion, before ultimate colonization of the intestinal niche. Exogenous fatty acids in bile, especially polyunsaturated fatty acids, were found to be selectively uptaken by *V. cholerae* and incorporated into their own glycerophospholipid structures, a phenotype not observed in other enteric pathogens[14]. Although evidence suggests this phenomenon occurs independent of the traditional virulence regulon, the additional observation that ocean sediment can serve as a source of subsequently incorporated fatty acids has hitherto unknown, yet likely important, impacts on *V. cholerae* fitness of in all of its ecological niches[14]. The exciting discovery of an unusual, surface exposed lipoprotein, that functions in the degradation of environmental lysophospholipids, provides yet another example of peculiar lipid utilization and membrane metabolism exhibited by *V. cholerae*[12,13].

Causing the disease cholera, which has long plagued human civilization, *V. cholerae* has been extensively studied as a paradigmatic bacterium in pathogenesis. Many investigations have examined the complex regulatory pathways that control the production and timed expression of virulence factors (e.g. cholera toxin, *ctxAB* or toxin co-regulated pilus, *tcp*), used in colonization of the human intestine, and in other aspects of cholera disease pathology. The regulation of the AlmEFG pathway, a *bona fide* pathogenic colonization factor, by the VprAB two component system[17] adds to our deep understanding of the virulence regulon of *V. cholerae*. Recent evidence with the ToxR master virulence regulator of *V. cholerae,* which notably controls, among other genes, cholera toxin and toxin co-regulated pilus, makes this connection even more intimate. Unpublished results from the Trent lab initially showed evidence that *V. cholerae*

Biotype El Tor TcpP mutants do not produce glycine-modified lipid A. However deficient mutants of other members of the virulence cascade, such as *toxR* or *toxT,* had no effect on glycine-modification. Similarly, but somewhat in contrast with these intial observations, a recent publication showed evidence that *toxR* affects VprAB expression, through another regulatory gene *leuO*[18], with impact on polymyxin B resistance. Details of lipid A modification, in these mutant backgrounds, and other salient phenotypes are an ongoing area of investigation. However, these observations further support the hypothesis that glycine modification of lipid A is important for *V. cholerae* pathogenesis in humans.

In conclusion, evidence in this manuscript contributes a thorough molecular mechanism for proteins of the AlmEFG pathway, that modify *V. cholerae* LPS with glycine or diglycine residues. The AlmEFG pathway is important not only for resistance to polymyxin B, but other CAMPs as well, such as the human cathelicidin LL-37, present in the intestinal lumen[19]. It also provides resistance against tachyplesin an arthropod produced CAMP, important in innate immunity of these organisms, where numerous reports show significant evidence for the association of *V. cholerae* with the chitinous exoskeletons of arthropod hosts[20,21]. It is likely that CAMP mediated selection with associated hosts contributed to the evolution of the AlmEFG pathway. Evidence continues to accumulate that strongly indicates that *V. cholerae* AlmEFG dependent membrane remodeling is important during human infection[17,18,22,23]. As such, further detailed mechanistic studies of AlmEFG, and associated regulatory networks, have the potential to inform further research on therapeutic drugs and vaccines targeting *V. cholerae* infections.

The scientific method provides a rigorous, robust approach to a philosophical investigation of existence. Mechanistic, basic science studies, like this one, have the power to enlighten, expanding in depth and sophistication our awareness of the natural

universe. On a more practical level, basic science also has the potential to inform the development of technologies and therapeutics aimed at improving the human experience, even if advances in knowledge do not *immediately* or *obviously* translate into contemporary applications. It is hubris, to assert or judge, that a given scientific investigation has no or little value. All science is part of a continuum, part of a dialogue with the material veil of reality, and should be valued necessarily as a vital community resource. Continued support of basic science is paramount to the progressive amelioration of biological life on this planet. Who knows? Maybe one day these endeavors will expand in impact, to other, farther reaches of our universe. Through detailed understanding of the structural composition and functional forms of existence we may be better able to craft more ideal states of being. Cheers.

Chapter 6: Experimental Methods[7]

6.1 Bacterial growth

Bacterial strains, plasmids, and oligonucleotides used in results sections of this document are listed in Tables 6.1-6.4. *V. cholerae* and *E. coli* were routinely grown at 37 °C on LB or LB agar unless otherwise indicated. *V. nigripulchritudo* was grown in LB supplemented with 0.5 M NaCl, or in Marine Broth 2216 or Marine Broth 2216 agar media (BD Difco). As appropriate antibiotics were supplemented in growth media at the following concentrations: 30 μg/ml chloramphenicol, 50 μg/ml kanamycin, 100 μg/ml ampicillin and 10 μg/ml streptomycin.

[7] Portions of this chapter have been previously published. Henderson, J.C., Fage, C.D., Cannon, J.R., Brodbelt, J.S., Keatinge-Clay, A.T., and Trent, M.S. Antimicrobial peptide resistance of Vibrio cholerae results from an LPS modification pathway related to non-ribosomal peptide synthetases. *ACS Chem Biol*. **86**, 2138-45 2014. JCH along with equal contribution from co-author CDF contributed data, wrote, and edited this report.

Table 6.1 Strains and plasmids used in Chapters 2 and 3.

Strain or Plasmid	Genotype or Description	Source or Reference
Strain		
V. cholerae E7946	O1 El Tor biotype	Hankins *et al*., 2012
V. cholerae N16961	O1 El Tor biotype	S. Payne, University of Texas, Austin, TX
E. coli XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI*^{q}*ZΔ*M15::Tn*10* (Tet^r)]	Stragene
E. coli BL21(DE3)	Protein expression strain, F$^-$ *omp*T *hsd*S$_B$ (r_B^- m_B^-) *dcm*, *gal* λ(DE3)	Novagen
E. coli BL21(DE3)pLysS	Protein expression strain, F$^-$ *omp*T *hsd*S$_B$ (r_B^- m_B^-) *dcm gal* λ(DE3) pLysS Cam^r	Novagen
Plasmid		
pET28b	IPTG-inducible expression vector, Kan^r	Novagen
pQlinkN	IPTG-inducible expression vector, Amp^r	C. Scheich *et al*., 2007
pQlinkH	IPTG-inducible expression vector, N-terminal His_6 tag, Amp^r	C. Scheich *et al*., 2007
pQlinkHVc1578	pQlinkH containing *almF*	This study
pQlinkNVc0780	pQlinkN containing *vc0780*	This study
pQlinkNVc2457	pQlinkN containing *vc2457*	This study
pQlinkNVc1579	pQlinkN containing *almE*	This study
pQVc1578-Vc0780	pQlinkH for co-expression of N-terminal-His_6-AlmF & non-tagged Vc0780	This study
pQVc1578-Vc2457	pQlinkH for co-expression of N-terminal-His_6-AlmF & non-tagged Vc2457	This study
pQVc1578-Vc2457-Vc1579	pQlinkH for co-expression of N-terminal-His_6-AlmF & non-tagged Vc2457 and Vc1579 (AlmE)	This study
pETAlmE	pET28b containing *almE*	This study
$pETAlmE_{C316S}$	C316S mutant derivative of pETAlmE	This study
$pETAlmE_{C316A}$	C316A mutant derivative of pETAlmE	This study
$pETAlmE_{L248V}$	L248V mutant derivative of pETAlmE	This study
$pETAlmE_{L248A}$	L248A mutant derivative of pETAlmE	This study

Table 6.2 Primers utilized in Chapters 2 and 3.

Restriction sites are underlined and nucleotide substitution mutations are lower case.

Name	**Sequence**
FpQlinkHVc1578	5’- GCGGCCGGATCCATGAAACAAGCCATTGAAAATATCC -3’
RpQlinkHVc1578	5’- GCCGCGAAGCTTTTATTCCCCTACTTCGCTCAC -3’
FpQlinkNVc0780	5’- GCGGCCGGATCCTTGTGTCGTTATGCCCGATCC -3’
RpQlinkNVc0780	5’- GCCGCGAAGCTTTCAAGTAGGTAAGGTGGGGAC -3’
FpQlinkNVc2457	5’- GCGGCCGGATCCATGATCGTTGGACTCGGAAC -3’
RpQlinkNVc2457	5’- GCCGCGAAGCTTTTAACGGCGTTCAAGGATCAC -3’
FpQlinkNVc1579	5’- GCGGCCGGATCCatgccatacaattctgcg -3’
RpQlinkNVc1579	5’- GCCGCGAAGCTTtcatcaatggcttgtttcatatg -3’
F-AlmE	5’- atgcttgcagctagcatgccatacaattctgcg -3’
R-AlmE	5’- gtacgaacggaattctcatcaatggcttgtttcatatgt -3’
F-AlmE$_{C316S}$	5’- gactgagttttttctccggacaagccttgct -3’
R-AlmE$_{C316S}$	5’- agcaaggcttgtccggagaaaaaactcagtc -3’
F-AlmE$_{C316A}$	5’- cgactgagttttttcgccggacaagccttgc -3’
R-AlmE$_{C316A}$	5’- gcaaggcttgtccggcgaaaaaactcagtcg -3’
F-AlmE$_{L248V}$	5’- gatttgacttttgatgtttcggtacatgata -3’
R-AlmE$_{L248V}$	5’- tatcatgtaccgaaacatcaaaagtcaaatc -3’
F-AlmE$_{L248A}$	5’- atttgacttttgatgcttcggtacatgatat -3’
R-AlmE$_{L248A}$	5’- atatcatgtaccgaagcatcaaaagtcaaat -3’

6.2 RECOMBINANT DNA TECHNIQUES.

Genomic DNA was routinely isolated using the Easy DNA kit (Life sciences). Plasmids (Tables 6.1 and 6.3) were isolated using the QIAPrep Spin Miniprep Kit (Qiagen). Custom primers were obtained from Integrated DNA Technologies (Tables 6.2 and 6.4). PCR reagents were purchased from Takara, New England Biolabs, or Stratagene, and PCR products were routinely isolated using the QIAquick PCR Purification Kit (Qiagen). DNA fragments were isolated from 0.7% agarose gels using the Qiaquick Gel Extraction Kit (Qiagen) or sometimes isolated using the DNA clean and concentrator kit (Genesee). All other DNA-modifying enzymes were purchased from New England Biolabs and used according to manufacturer's protocols. When necessary, purity and yield of extracted DNA were monitored using a NanoDrop instrument (GE). To create pQlink plasmids used for co-expression of multiple genes, reagents purchased from Invitrogen were used for ligation-independent cloning (LIC) and as described in Scheich et al., 2007 [1]. All plasmids constructed in this study are initially transformed into chemically competent *E. coli* XL-1 Blue storage strain before transformation into the particular strain for a given experiment (Tables 6.1 and 6.3). Plasmid promoter and gene-insert sequences were verified by DNA sequencing at the ICMB Core Facility at the University of Texas at Austin or with Genewiz.

6.3 METHODS SPECIFIC TO CHAPTER 2 (ALMF)

6.3.1 Cloning, Expression, and Purification of N-terminal His_6-AlmF or Co-expression Constructs.

Gene *almF* was PCR-amplified from El Tor biotype *V. cholerae* O1 strain N16961 genomic DNA (gDNA). *Bam*HI or *Hind*III restriction sites were engineered to flank the resultant PCR product using primers FpQlinkHVc1578 and RpQlinkHVc1578 (Table 6.2). After restriction digest of both plasmid pQlinkH [1] and PCR-amplified *almF* insert, plasmid DNA was treated with Antarctic phosphatase before ligation reactions were used to produce pQlinkHVc1578. The same cloning strategy was used to obtain pQlinkNVc0780, pQlinkNVc2457, and pQlinkNVc1579, where each gene was amplified using the corresponding primers by PCR with *V. cholerae* N16961 gDNA template. Co-expression constructs pQVc1578-Vc0780, pQVc1578-Vc2457, and pQVc1578-Vc2457-Vc1579 (Table 6.1) were created using the standardized protocol described in Scheich et al., 2007 [1].

Escherichia coli BL21(DE3)pLysS cells containing the appropriate expression plasmid were grown in LB (1.5 L, 37 °C, 180 rpm shaking). Upon reaching approximately OD_{600} = 0.6, cultures were briefly chilled on ice to ~18 °C and expression was induced by addition of a sterile IPTG solution (100 mM) to a final concentration of 1 mM. After 20 h (18 °C, 180 rpm) cells were pelleted, resuspended in ice-cold Buffer A (500 mM NaCl, 50 mM HEPES pH 7.5, 10% (v/v) glycerol, 40 mM imidazole, 1 mM DTT) supplemented with 10 μM PMSF, and lysed via French press (20k psi). Lysate was clarified by centrifugation (4 °C, 20k x *g*, 30 m) and the soluble portion was syringe-

filtered (0.4 μm) before injection over an AKTA-compatible Ni-sepharose HiTrap FF column (GE) pre-equilibrated with Buffer A. The column was washed with Buffer A (10 column volumes), before gradient elution over 30 column volumes to 100% Buffer B (500 mM NaCl, 50 mM HEPES pH 7.5, 10% (v/v) glycerol, 400 mM imidazole, 1 mM DTT). 1 ml fractions were collected where the peak Abs_{260} fraction typically eluted between ~130-210 mM imidazole. Corresponding fractions with >80% AlmF, as determined by SDS-PAGE analysis (AlmF MW 8.9 kDa), were pooled and concentrated via filter centrifugation (Amicon Ultra, 3 kDa cutoff). To further purify AlmF, size-exclusion chromatography was performed using a Superdex S75 column (GE) equilibrated with 1:5 dilution of Buffer A (no imidazole). Concentrated Ni-sepharose-purified AlmF was diluted 1:5 in water before addition to S75 column. Peak fractions were analyzed by SDS-PAGE, and fractions containing >95% AlmF, were pooled and concentrated via filter centrifugation (Amicon Ultra, 3 kDa cutoff).

6.3.2 Destabilizing-Urea PAGE of Purified AlmF Samples.

Urea-PAGE gels were manually cast using a mini-PROTEAN casting stand and frame (BioRad). Separating gel consisted of 13% polyacrylamide (19:1; Sigma), 2.5 M urea, 375 mM Tris pH 9.5, 0.05% w/v APS, and 0.05% v/v TEMED. Stacking gel consisted of 4% polyacrylamide (19:1; Sigma), 2.5 M urea, 125 mM Tris pH 6.8, 0.05% w/v APS, and 0.1% v/v TEMED. 20 μg of total protein was loaded per lane. Gels were run with prechilled running buffer (25 mM Tris Base, 192 mM glycine, 15% w/v urea, dissolved in ddH_2O) in a 4°C cold room for ~3-4 h at a constant 100 V. Visualization of

protein bands was performed by Coomassie-staining for 24 h (45% v/v methanol, 10% v/v glacial acetic acid, 45% v/v ddH_2O, and 3g/L Coomassie R-250) and 24-48 h, due to the high % polyacrylamide, with destain solution (40% v/v ethanol, 10% v/v glacial acetic acid, and 50% v/v ddH_2O). Gels were stored in ddH_20 until photographically imaged.

6.3.3. UVPD-MS Analysis of Tryptic Digested AlmF Samples.

Apo-AlmF, holo-AlmF (AlmF + Vc2457), and glycyl-AlmF (AlmF + Vc2457 + AlmE) samples were buffer exchanged three times into 100 mM ammonium bicarbonate using 3 kDa molecular weight cutoff filters. Samples were then digested with trypsin (1:50 enzyme-to-substrate ratio) overnight and desalted. Peptides were then evaporated to dryness and reconstituted in LC solvent A (98/2/0.1, water/acetonitrile/formic acid, v/v/v). For LC profiling, peptides were separated on a Dionex RSLCnano (ThermoDionex) using a gradient of 0-15% solvent B (100/0.1, acetonitrile/formic acid) over the course of 30 minutes, and then up to 70% solvent B over 15 minutes. Flow rate was set to 750 nL/min through a 150 x .075 mm C18 analytical column packed in-house. The three most abundant multiply-charged ions were selected for fragmentation using collision-induced dissociation (NCE 35) and subsequent detection in a Velos Pro dual linear ion trap mass spectrometer (ThermoFisher, San Jose, CA). Ultraviolet photodissociation spectra of the same peptides were acquired on an Orbitrap Elite mass spectrometer (ThermoFisher) after elution using the same gradient described previously on an Eksigent NanoLC Ultra (Eksigent, Redwood City, CA) at a flow rate of 300

nL/min. UVPD was performed using 3 pulses of a 193 nm excimer laser (Coherent Excister XS) (at 3 mJ per pulse). The Orbitrap Elite mass spectrometer was modified and adapted for UVPD as described previously [2]. All spectra were manually interpreted.

6.4 METHODS SPECIFIC TO CHAPTER 3 (ALME)

6.4.1 Cloning, Expression, and Purification of N-terminal His_6-AlmE or C316 & L248 single amino acid variants.

The *almE* gene was PCR-amplified from pVc1579 [3] with primers F- and R-AlmE (Table 6.2), cleaved with *Nhe*I and *Eco*RI, and ligated into pET28b (Novagen) to yield pETAlmE. C316S and L248V mutants were prepared by carrying out site-directed mutagenesis on pETAlmE with the associated primers (Table 6.2). C316S and L248V were further mutated to generate C316A and L248A with the associated primers (Tables 6.1 and 6.2).

Escherichia coli BL21(DE3) cells containing the appropriate expression plasmid were grown in LB (1 L) with 50 mg/L kanamycin at 37 °C (180 rpm shaking). Upon reaching approximately OD_{600} = 0.5, cultures were briefly chilled on ice to ~18 °C and expression was induced by addition of a sterile IPTG solution (100 mM) to a final concentration of 1 mM. After 18 h (18 °C, 180 rpm) cells were pelleted, resuspended in ice-cold Buffer A + 10 μM PMSF (500 mM NaCl, 50 mM HEPES pH 7.5, 10% (v/v) glycerol, 40 mM imidazole), and lysed via French press (20k psi). Lysate was clarified by centrifugation (4 °C, 20k x *g*, 30 m) and the soluble portion was syringe-filtered (0.4 μm) before injection over an AKTA-compatible Ni-sepharose column (GE) pre-equilibrated with Buffer A (no PMSF). The column was washed with Buffer A (10 column volumes), before gradient elution over 30 column volumes to 100% Buffer B (500 mM NaCl, 50 mM HEPES pH 7.5, 10% (v/v) glycerol, 400 mM imidazole). 1-ml fractions were collected where the peak Abs_{260} fraction typically eluted between ~150-200 mM

imidazole. Corresponding fractions with >95% AlmE, as determined by SDS-PAGE analysis (AlmE MW 63 kDa), were pooled and concentrated via filter centrifugation (Amicon Ultra, 3 kDa cutoff). For crystallization studies, Ni-sepharose-purified AlmE was further purified on a Superdex 200 gel filtration column (GE Healthcare Life Sciences) equilibrated with 150 mM NaCl, 10 mM HEPES pH 7.5, 10% (v/v) glycerol.

In vitro AlmE Assays. Glycine containing AlmE/ pyrophosphatase coupled assay conditions are as follows: HEPES (50 mM, pH 7.5), KCl (100 mM), $MgCl_2$ (10 mM), ATP (1 mM), glycine (1 mM), yeast inorganic pyrophosphatase (Sigma, 0.5 U), purified holo-AlmF (8 μM), DTT (100 μM), and purified AlmE (800 nM) incubated at 25°C for 2.5 or 3 h as indicated in Results and/or Figure Legends. Percent activity corresponds to the percent conversion of [ATP] during AlmE catalyzed aminoacyl-AMP formation, where detected [phosphate]/2 equals [ATP] consumed. AlmF ^{3}H-aminoacylation assays with AlmE were performed under identical reaction conditions as described above with omission of DTT and pyrophosphatase, inclusion of 0.3 μCi ^{3}H-amino acid, and were incubated for 18 hours at 18 °C. AlmF in assay mixtures was concentrated by TCA precipitation and resolved by SDS-PAGE. Resolved protein was transferred to a nitrocellulose membrane, which was exposed to a ^{3}H-sensitive screen for 5 days.

6.4.2 Crystallization of AlmE.

Protein was concentrated to 10 mg/mL in 10% glycerol, 25 mM NaCl, 10 mM HEPES pH 7.5 and incubated with 10 mM glycine, 10 mM ATP, and 1 mM $MgCl_2$ for 10

min before preparing crystal trays. Narrow rod-shaped crystals grew in 50 mM $MgCl_2$, 100 mM HEPES pH 7.0 by sitting-drop vapor diffusion at 22 °C. Prior to flash freezing in liquid nitrogen, crystals were soaked for 10 min in cryosolution consisting of 20% glycerol, 50 mM $MgCl_2$, 100 mM HEPES pH 7.0. While crystals also grew in the absence of glycine and ATP, their quality and size were significantly diminished.

Data Collection, Processing, Phasing, and Refinement. Data were collected at Advanced Photon Source Beamline 23-ID-D and processed in HKL2000 [4] (Table 3.1). One crystal diffracted to 2.26 Å. Molecular replacement (MR) computations were performed in Phaser [5,6], with the N-terminal body of one monomer of PheA (PDB code 1AMU, residues 17-428) as a search model [7]. One AlmE monomer was found per asymmetric unit in space group $P3_121$. This solution was submitted to the ARP/wARP server for automated refinement [8] before iterative refinement was carried out in PHENIX [9], REFMAC [5,10], and Coot [11].

6.5 METHODS SPECIFIC TO CHAPTER 4 (ALMG)

6.5.1. Generation of deep-rough and other *E. coli* mutant strains

Construction of gene deletion strains of *E. coli* (Table 6.3), except Δ*rfaDFC*, were performed via P1 *vir* phage transduction according to previously published methodologies[12]. In brief, Keio collection *E. coli* deletion mutants, where a kanamycin resistance cassette replaces the open reading frame of a given gene, were used to generate donor phage for transduction into a given host *E. coli* background. Lambda red recombineering was used to generate *ΔrfaDFC* mutants according to previously published protocols[13]. Modifications to the general procedure include supplementation of outgrowth cultures with 1 mM arabinose and incubation at 30°C. Markerless deletion mutants were subsequently generated according to published protocols with yeast Flp recombinase machinery encoded on pcp20 a high temperature curable vector[13]. In multiple deletion strains P1 transduction was followed by marker removal, before subsequent deletions were introduced. Recombineering was always the last step, as *ΔrfaDFC* severely limits binding of P1 *vir* phage; phage tail proteins specifically interact with the core-oligosaccharide portion of LPS[14]. Correct strain construction was verified by PCR, TLC analysis of ^{32}P radiolabeled or MALDI-MS analysis of non-radiolabeled lipid products, and ProQ Emerald based detection of gel separated LPS.

Table 6.3 Strains and plasmids used in Chapter 4.

Strain or Plasmid	Genotype or Description	Source or Reference
Strain		
V. cholerae E7946	O1 El Tor biotype	Hankins *et al*., 2012
V. cholerae Classical	O395 Classical biotype	Karl Klose, University of Texas San Antonio, San Antonio, TX
V. cholerae E7946Δ*almG*	Δ*almG* clean deletion, replaced with Kanamycin^R cassette	Hankins *et al*., 2012
V. cholerae N16961	O1 El Tor biotype	S. Payne, University of Texas, Austin, TX
E. coli BW25113	Laboratory *E. coli* K12 strain, host strain of Keio non-essential gene deletion collection	GE, Dharmacon
E. coli BW25113 Δ*lpxT*	Kanamycin^R, host for P1 *vir* transduction	GE, Dharmacon
E. coli BW25113 Δ*eptA*	Kanamycin^R, host for P1 *vir* transduction	GE, Dharmacon
E. coli BW25113 Δ*eptB*	Kanamycin^R, host for P1 *vir* transduction	GE, Dharmacon
E. coli W3110	Common Laboratory *E. coli* K12 strain	Yale, CGSC
E. coli MLK1067	*E. coli* K12 W3110 Δ*lpxM,* Chloramphenicol^R	Karrow and Georgopoulos (1992)
E. coli JCH0010	*E. coli* MLK1067 Δ*lpxT::*Kanamycin^R, Chloramphenicol^R	This study
E. coli JCH0011	*E. coli* MLK1067 Δ*lpxT,* Chloramphenicol^R	This study
E. coli JCH0020	*E. coli* MLK1067 Δ*lpxT,* Δ*rfaDFC::* Kanamycin^R, Chloramphenicol^R	This study
E. coli JCH0021	*E. coli* MLK1067 Δ*lpxT,* Δ*rfaDFC*, Chloramphenicol^R	This study
E. coli JCH0100	*E. coli* MLK1067 Δ*lpxT,* Δ*eptA::*Kanamycin^R Chloramphenicol^R	This study
E. coli JCH0101	*E. coli* MLK1067 Δ*lpxT,* Δ*eptA* Chloramphenicol^R	This study
E. coli JCH0200	*E. coli* MLK1067 Δ*lpxT,* Δ*eptB::*Kanamycin^R Chloramphenicol^R	This study
E. coli JCH0201	*E. coli* MLK1067 Δ*lpxT,* Δ*eptB* Chloramphenicol^R	This study
E. coli JCH0300	*E. coli* MLK1067 Δ*lpxT,* Δ*eptA,* Δ*eptB::*Kanamycin^R, Chloramphenicol^R	This study

Table 6.3 Strains and plasmids used in Chapter 4. (continued)

E. coli JCH0301	*E. coli* MLK1067 Δ*lpxT*, Δ*eptA*, Δ*eptB*, Chloramphenicol^R	This study
E. coli JCH0302	*E. coli* MLK1067 Δ*lpxT*, Δ*eptA*, Δ*eptB*, Δ*rfaDFC::*Kanamycin^R, Chloramphenicol^R	This study
E. coli JCH0303	*E. coli* MLK1067 Δ*lpxT*, Δ*eptA*, Δ*eptB*, Δ*rfaDFC*, Chloramphenicol^R	This study
E. coli XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI^{q}ZΔ*M15::Tn*10* (Tet^r)]	Stragene
E. coli C43(DE3)pLysS	Protein expression strain, F – *ompT hsdSB (rB-mB-) gal dcm* (DE3) pLysS (CmR)	Lucigen

Plasmid

pET28b	IPTG-inducible expression vector, Kan^r	Novagen
pET21b	IPTG-inducible expression vector, Kan^r	Novagen
pQlink	IPTG-inducible expression vector, Amp^r	C. Scheich *et al*., (2007)
pWSK29	Low-copy IPTG-inducible expression vector, Amp^r	Wang and Kushner (1991)
pET28AlmG	pET28b containing *almG*	This study
pET21AlmG	pET21b containing *almG,* N-terminal His tag coding region is removed	This study
pKD4	Plasmid used in recombineering, encodes for FRT flanked Kanamycin^R cassette	Datsenko and Wanner (2000)
pcp-20	Arabinose inducible, Temperature curable Plasmid, helper plasmid to remove FRT flanked Kanamycin^R cassettes from target host DNA, to generate markerless deletion strains	Datsenko and Wanner (2000)
pQLpxN	pQlink containing *lpxN*	This Study
pQAlmE	pQlink containing *almE*	Henderson *et al*. (2014)
pQAlmF	pQlink containing *almF*	Henderson *et al*. (2014)
pQAlmG	pQlink containing *almG*	This study
pQVcAcpS	pQlink containing *vc2457*	Henderson *et al*. (2014)
pQLpxN, AlmE	pQlink containing *lpxN* and *almE*	This study
pQLpxN, AlmEF	pQlink containing *lpxN, almE* and *almF*	This study
pQLpxN, AlmEFG	pQlink containing *lpxN, almE, almF*, and *almG*	This study
pQLpxN, AlmEFG, VcAcpS	pQlink containing *lpxN, almE, almF, almG*, and *vc2457*	This study

Table 6.3 Strains and plasmids used in Chapter 4. (continued)

pQAlmEF	pQlink containing *almE* and *almF*	Henderson *et al*. (2014)
pQAlmEFG	pQlink containing *almE, almF,* and *almG*	This study
pWSK29::NHis$_6$AlmG	pWSK29 containing NHis$_6$-*almG,* as excised from pET28AlmG	This study
pWSK29::CHis$_6$AlmG	pWSK29 containing CHis$_6$-*almG,* as excised from pET21AlmG	This study

6.5.2. General procedure for isolation of glycerophospholipid and Kdo$_2$-lipid A species.

A single colony of bacteria is used to inoculate 5 ml of media (typically LB or G56 media, supplemented with 100 μM IPTG or antibiotic as in section 6.1 as necessary) and grown overnight at 37 ºC or at 30°C for Δ*rfaDFC* strains. The next day a 10 ml culture is inoculated to a starting OD_{600} of ~0.05, grown in a 20x150mm disposable glass culture tube. Growth media for TLC analysis is supplemented with 10 μCi/ml of ^{32}P. Cells are grown to mid-exponential phase, when an OD_{600} of 0.8-1.0 is reached. Cells are then harvested in 16 x 125mm glass centrifuge tubes with PTFE lined cap using a fixed angle clinical centrifuge at 1,500 x *g* for 12 minutes. Supernatant is removed into appropriate radioactive waste container. Cell pellets are washed with 5 ml of phosphate buffered saline (PBS), and centrifuged again for 10 minutes at 1,500 x *g*. Supernatant is discarded, cells are resuspended in 5 ml of single phase Bligh-Dyer mixture consisting of chloroform-methanol-water (1:2:0.8, v:v:v), mixed by vortex, and incubated at room temperature for 20-30 minutes to ensure complete chemical cell lysis. These mixtures are

centrifuged in a clinical centrifuge at 1,500 x *g* for 15 minutes. Supernatant, which contains glycerophospholipids, isoprenyl lipids, free lipid A, and Kdo-lipid A domains is then transferred into a clean 16 x 125mm glass centrifuge tubes, then converted into a two-phase Bligh-Dyer mixture by adding 1 ml of chloroform and 1 ml of methanol, yielding a chloroform-methanol-aqueous (2:2:1.8 v/v) mixture. This mixture is vortexed for ~30s, then centrifuged for 12 minutes in a clinical centrifuge at 1,500 x *g* to separate the organic and aqueous phases. The lower oganic phase is removed into a clean glass centrifuge tube using a glass Pasteur pipet. A second extraction is performed by adding 2 ml of pre-equilibrated two-phase Bligh-Dyer lower phase (2:2:1.8; v:v:v; chloroform:methanol:PBS pH 7.5) to the remaining upper phase. This material is vortexed and centrifuged as before. The lower phase is combined with the previous lower phase and 7.6 ml of pre-equilibrated upper phase is added to yield a final two-phase Bligh-Dyer solution (chloroform-methanol-water; 2:2:1.8, v/v). Material is vortexed, centrifuged at 1500 x g for 12 minutes, lower phase is transferred to a clean glass tube where isolated lipid material is dried using a nitrogen stream dryer. Dried samples are used right away in subsequent protocols (Section 6.5.3 or 6.5.4), transferred to smaller containers by resuspension in 4:1 v/v chloroform:methanol then re-dried, and/or routinely stored as dried material at -20 ºC until further use.

6.5.3 Thin layer chromatography (TLC) to analyze ^{32}P radio-isotopically labeled phospholipids

E. coli strains were grown in G56, a low phosphate (150 μM) minimal media, supplemented with 100 μM isopropyl β-D-1 thiogalactopyranoside (IPTG) to radiolabel phospholipid material, where isolation of lipid material is performed as in section 6.5.2. TLC mobile phase solvent system is prepared in a clean TLC tank with ~40 cm chromatography paper, to accommodate a 10 x 20 cm or 20 x 20cm silica gel 60 plate. Solvents are allowed to pre-equilibrate for at least 3 hours, most often overnight. Silica gel 60 plates are prepared by scraping excess silica from edges using a razor blade. Origin where lipid material is spotted is demarcated using a graphite pencil 1.5 cm from the bottom of the plate. ^{32}P-labeled sample is typically dissolved in 500 µl of 4:1 chloroform-methanol (v/v), vortexed and bath sonicated to completely dissolve lipid material. 5 µl of lipid sample is added to a scintillation vial containing 5 ml of scintillation cocktail and counted in a scintillation counter to enable calculation of total counts per minute of sample. With a microcapillary pipet, 20,000-50,000 cpm per sample was spotted on the silica plate origin and allowed to air dry (> 15 min). TLC plates are then placed into a pre-equilibrated tank (solvent system indicated in figure legends) and run for ~3 hours until mobile phase reaches the top of the plate. Solvent systems used were chloroform: pyridine: 88% formic acid: water, 30:70:16:10, v/v as in Figure 4.3 and Figure 4.5 panel A or chloroform: methanol: glacial acetic acid: water, 25:15:4:4, v/v as in Figures 4.5 panel B. Plates are removed and dried under a cold air

gun for ~40 min, especially when pyridine was used in the TLC solvent system. TLC plates were wrapped in plastic wrap and exposed to phosphorImager screen in an autoradiography cassette overnight. The next morning, screens were scanned using a phosphorimager and images were subsequently processed and analyzed using Quantity One® software (BioRad).

6.5.4. Matrix assisted laser desorption ionization mass spectrometry to structurally characterize isolated phospholipid species

E. coli strains were routinely grown in G56, a low phosphate (150 μM) minimal media, supplemented with 1 mM isopropyl β-D-1 thiogalactopyranoside (IPTG), to match conditions for isolation of lipid material as analyzed by TLC. Typically dried lipid A samples were transferred in microvials, where material was resuspended in ~50 µl chloroform-methanol (4:1, v/v) and vortexed/sonicated to obtain a ~1-5 µg/µl lipid A solution when extracted from *E. coli*. Azothiothymine matrix was prepared by mixing a saturated 6-aza-2-thiothymine in 50% acetonitrile solution with saturated tribasic ammonium citrate (20:1, v/v). Mixtures were vortexed and centrifuged, before 0.5 μl was spotted on MALDI plate. 0.5 µl of lipid sample was typically deposited onto the previously spotted ATT matrix, and mixed on plate. MALDI-MS data were calibrated using calibration 1 and calibration 2 mixtures from Anaspec peptide mass standards in the negative mode. Spectra were acquired in the linear, negative mode on a Voyager-DE (AB biosystems) by scanning for optimal ion signals, with at minimum 300 laser shots per spectra. Samples were diluted or concentrated as needed to generate high-quality

spectra. Typical voltage parameters included 20 k V, 95% grid voltage, 0.002 guide wire voltage, and a delay time of 150 ns. A low mass gate at 500 Da was also employed.

6.5.5. Procedure to visualize LPS by ProQ® Emerald fluorescent dye

Bacteria were grown in Luria-Bertani (LB) media cultures at 30°C until mid-exponential phase. Cell amounts were normalized based on optical density, harvested, and resuspended in Tricine buffered Laemmli sample buffer (with 4% β-mercaptoethanol), followed by proteinase K digestion overnight at 55°C to remove contaminant protein. Resultant material containing intact LPS is boiled at 100°C before resolution by tricine-sodium dodecyl sulfate–10-20% polyacrylamide gel electrophoresis. Electrophoresis is performed at 4°C, typically for 2 h. In gel oxidation of associated carbohydrate was followed with visualization of LPS using a fluorescence based molecular probe according to manufacturer protocols (Pro-Q® Emerald Life technolgies).

6.5.6. Cloning and expression conditions for N-terminal or C-terminal His_6-AlmG

Host strains for the expression of hexahistidine-AlmG fusions (Table 6.3) were of either *E. coli* or *V. cholerae* origin, as indicated for any specific experiment, and were constructed similarly to material in Section 6.4.1, using different primers specific to *vc1577* (AlmG) (Table 6.4). *V. cholerae* expression was routinely performed from low copy pWSK29 vectors, in LB cultures grown and induced with 1 mM IPTG (unless otherwise indicated) at 30°C or 37°C. For expression in *E. coli* pET vectors were used

(Table 6.3). Initial growth was performed at 37°C in LB, until preparation for induction, when cultures were chilled on ice for at least 15 minutes, typically of cells at OD600 ~ 0.4. 1 mM IPTG was introduced to cell media and growth was allowed to proceed at 18°C in overnight cultures (subsequent cellular material was then used as in sections 6.5.8 or 6.5.9). Large cultures (1.5 l each) were generally used in preparation for multi-step protein purification (section 6.5.9).

Table 6.4 Primers utilized in Chapter 4.

Restriction sites are underlined. Fk_seq are specific to flanking regions of indicated genes for use in sequencing or PCR verification of gene deletions. Lowercase nucleotides signify sequences on target plasmids or genomic DNA (e.g. as in plasmids used in recombineering _pKD4).

Name	**Sequence**
rfaD_ForFlank_pKD4	ATGATTACAGACATTCGTGTCTGAGATTGTCTCTGACTC CATAATTCGAAGGTTACAGTTgtgtaggctggagctgcttc
rfaC_RevFlank_pKD4	TGATTTCAGAGTGTAAGGTTTCAATGAATGAAGTTTAA AGGATGTTAGCATGTTTTACCTcatatgaatatcctccttag
Fk_lpxM_for_seq	GCCTTATCCGAAACTGGAAAAGC
Fk_lpxM_rev_seq	TCTCCTCGCGAGAGGCTTTT
Fk_eptA_for_seq	TGTTCGATGGAAACACCGTG
Fk_eptA_rev_seq	CGTATCGTCTTCAACAATCAGAATTT
Fk_eptB_for_seq	GTTTGTCCAGGGTTTGTTTGC
Fk_eptB_rev_seq	CGAGAAAGTCAGCAGGCCGC
Fk_lpxT_for_seq	GCGTTTTTCAGTAAGATAATTAGAGA
Fk_lpxT_rev_seq	AATAACCCTGATGATGTTAATTACTGTGAG
Fk_rfaC_for_seq	TACAAGAGGAAGCCTGACGG
Fk_rfaC_rev_seq	AAGTTTAAAGGATGTTAGCATGTTTTACCT
Fk_rfaD_for_seq	TGACTCCATAATTCGAAGGTTACAGTT
Fk_rfaD_rev_seq	CACCAGTATTTTCATGCAGAGCTC
Fk_rfaF_for_seq	GCGACGCATAAGAGCTCTGC
Fk_rfaF_rev_seq	ACGATCAAAACCCGCATCCG
FpQLpxN	GCGGCCGGATCCgtgagcgatcaagacaacggc
RpQLpxN	GCCGCGAAGCTTttagctttgatcgccatgatggg
FpQAlmG	GCGGCCGGATCCgtgcgtatttttggttaaaagc
RpQAlmG	GCCGCGAAGCTTttacttaaaacgccgataaagccag
FpET28AlmG	GCGGCCCATATGgtgcgtatttttggttaaaagc
RpET28AlmG	GCCGCGCTCGAGttacttaaaacgccgataaagccag
FpET21AlmG	GCGGCCCATATGgtgcgtatttttggttaaaagc
RpET21AlmG	GCCGCGCTCGAGttacttaaaacgccgataaagccag

6.5.7 Polymyxin B Minimum Inhibitory Concentration Assay.

V. cholerae strains were grown overnight in LB media, then diluted next day 1:100 in fresh LB (1 mM isopropyl β-D-1 thiogalactopyranoside IPTG, 100 μg/ml ampicillin). Cells were grown to an OD_{600} of ~0.5, diluted 1:50 and then applied to LB agar with sterile cotton tip applicator. Quantitative minimum inhibitory concentration (MIC) values were determined using ETest® gradient polymyxin strips (Biomerieux) after ~18 h growth 37°C, using manufacturer's guidelines.

6.5.8. Isolation of membranes and detergent exchange assays with N-terminal His_6-AlmG

N-terminal hexahistidine-tagged AlmG was overexpressed in *E. coli* C43(DE3)pLysS. Bacteria were harvested (washed with 1x PBS), resuspended in lysis buffer (0.2 M NaCl, 50 mM HEPES pH 7.5) and lysed by the hydraulic modification to the method of French and Milner[37]. Lysate was cleared of cell debris and unlysed cellular material, by centrifugation at 20k x g for 30 minutes. Crude membranes are isolated from cell-free lysate by ultracentrifugation at 125 k x g for 1 h, washed in lysis buffer by dounce homogenization[38] to remove weakly or non-specifically associated material, followed by a second ultracentrifugation step to isolate washed membranes. Washed membranes were re-homogenized and exchanged into buffer (0.2 M NaCl, 50 mM HEPES pH 7.5) containing each of several detergents at 1% w/v, and incubated at 4°C for 2 hours. Ultracentrifiguation was used to remove "debris" and insoluble membrane

material at 125 x g for 1 h. Detergent solubilized supernatant was applied to a spin column containing nickel resin to bind hexahistidine tagged AlmG. Bound material was washed in buffer containing 20 mM Imidzaole, and eluted with buffer containing 400 mM imidazole to recover any potentially bound protein. Samples were analyzed by SDS-PAGE. Protein gels were visualized using coomassie dye. Detergents used include as in Figure 4.9: (Tr) Triton X-100, (Tw) Tween-20, (Br) Brij-58, (OG) octyl-glucoside, (NG) nonyl-glucoside, (DM) decyl-maltoside, (DDM) dodecyl-maltoside, or (LD) lauryl-dimethylamine oxide. All detergents were used at 1% w/v well above each respective critical micelle concentrations (CMC; > 5-100 fold).

6.5.9 Multi-step purification of N-terminal His_6-AlmG

Escherichia coli C43(DE3)pLysS cells containing the appropriate pET28 expression plasmid were grown in LB (1.5 L, 37 °C, 180 rpm shaking). Upon reaching approximately OD_{600} = 0.4, cultures were briefly chilled on ice to ~18 °C and expression was induced by addition of a sterile IPTG solution (100 mM) to a final concentration of 1 mM. After ~18 h (18 °C, 180 rpm) cells were pelleted, resuspended in ice-cold Buffer A (500 mM NaCl, 50 mM HEPES pH 7.5, 20 mM imidazole, 1 mM DTT, 0.5% w/v Brij-58) supplemented with 10 μM PMSF, and lysed via cell disruption (20k psi). Lysate was clarified by centrifugation (4 °C, 20k x *g*, 30 m) and the soluble portion was syringe-filtered (0.4 μm) before injection over an AKTA-compatible HisTrap FF column (GE) pre-equilibrated with Buffer A. The column was washed with Buffer A (10 column

volumes), before gradient elution over 30 column volumes to 100% Buffer B (500 mM NaCl, 50 mM HEPES pH 7.5, 400 mM imidazole, 1mM DTT, 0.5% w/v Brij-58). 1 ml fractions were collected where the peak Abs_{260} fraction typically eluted between ~100-160 mM imidazole. Corresponding fractions with >80% AlmG, as determined by SDS-PAGE analysis (AlmG MW 31 kDa), were pooled and concentrated via filter centrifugation (Amicon Ultra, 10 kDa cutoff).

To further purify AlmG, material was desalted before AKTA compatible SP column purification. Concentrated fractions were diluted to a final volume of 1.5 ml before exchange into Buffer C (20 mM NaCl, 20 mM HEPES pH 8.0, 1 mM DTT, 0.05% Brij-58), by direct injection into an AKTA compatible desalt column prepared according to manufacturer's protocols. Protein was recovered as monitored by Abs_{260} and concentrated to a final volume of 0.5 ml using 10 kDa centrifugation filters as with HisTrap FF elution fractions above. This material was directly injected onto a prepared SP HP column (Akta), according to manufacturer's protocols. Sample was washed with 5 column volumes of Buffer C, and then gradient eluted to 100% Buffer D (1 M NaCl, 20 mM HEPES pH 8.0, 1 mM DTT, 0.05% Brij-58). Peak AlmG fractions eluted around 0.4 M NaCl as monitored by Abs_{260} and analyzed by SDS-PAGE.

References

1. Kamio, Y. & Nikaido, H. Outer membrane of *Salmonella* typhimurium: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry* **15,** 2561–2570 (1976).
2. Dowhan, W. Lipids and Extracellular Materials. *Annual Review of Biochemistry* **83,** 45–49 (2014).
3. Kanfer, J. & Kennedy, E. P. Metabolism and function of bacterial lipids. Ii. Biosynthesis of phospholipids in *Escherichia coli*. *J. Biol. Chem.* **239,** 1720–1726 (1964).
4. Raetz, C. R. Molecular genetics of membrane phospholipid synthesis. *Annu. Rev. Genet.* **20,** 253–295 (1986).
5. Lu, Y.-H., Guan, Z., Zhao, J. & Raetz, C. R. H. Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of *Escherichia coli*. *J. Biol. Chem.* **286,** 5506–5518 (2011).
6. Osman, C., Haag, M., Wieland, F. T., Brügger, B. & Langer, T. A mitochondrial phosphatase required for cardiolipin biosynthesis: the PGP phosphatase Gep4. *EMBO J.* **29,** 1976–1987 (2010).
7. Tamura, Y. *et al*. Tam41 is a CDP-diacylglycerol synthase required for cardiolipin biosynthesis in mitochondria. *Cell Metab.* **17,** 709–718 (2013).
8. Tan, B. K. *et al*. Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 16504–16509 (2012).

9. Zhang, J. *et al*. Mitochondrial Phosphatase PTPMT1 Is Essential for Cardiolipin Biosynthesis. *Cell Metabolism* **13,** 690–700 (2011).

10. Bogdanov, M., Xie, J., Heacock, P. & Dowhan, W. To flip or not to flip: lipid-protein charge interactions are a determinant of final membrane protein topology. *The Journal of Cell Biology* **182,** 925–935 (2008).

11. Dowhan, W. & Bogdanov, M. Lipid-dependent membrane protein topogenesis. *Annu. Rev. Biochem.* **78,** 515–540 (2009).

12. Matsumoto, K. Dispensable nature of phosphatidylglycerol in *Escherichia coli*: dual roles of anionic phospholipids. *Mol. Microbiol.* **39,** 1427–1433 (2001).

13. Wang, X., Bogdanov, M. & Dowhan, W. Topology of polytopic membrane protein subdomains is dictated by membrane phospholipid composition. *EMBO J.* **21,** 5673–5681 (2002).

14. Zhang, W., Campbell, H. A., King, S. C. & Dowhan, W. Phospholipids as determinants of membrane protein topology. Phosphatidylethanolamine is required for the proper topological organization of the gamma-aminobutyric acid permease (GabP) of *Escherichia coli*. *J. Biol. Chem.* **280,** 26032–26038 (2005).

15. Castuma, C. E., Crooke, E. & Kornberg, A. Fluid membranes with acidic domains activate DnaA, the initiator protein of replication in *Escherichia coli*. *J. Biol. Chem.* **268,** 24665–24668 (1993).

16. de Vrije, T., de Swart, R. L., Dowhan, W., Tommassen, J. & de Kruijff, B. Phosphatidylglycerol is involved in protein translocation across *Escherichia coli* inner membranes. *Nature* **334,** 173–175 (1988).

17. Garner, J. & Crooke, E. Membrane regulation of the chromosomal replication activity of *E. coli* DnaA requires a discrete site on the protein. *EMBO J.* **15,** 2313–2321 (1996).

18. Lill, R., Dowhan, W. & Wickner, W. The ATPase activity of secA is regulated by acidic phospholipids, secY, and the leader and mature domains of precursor proteins. *Cell* **60,** 271–280 (1990).

19. Xia, W. & Dowhan, W. In vivo evidence for the involvement of anionic phospholipids in initiation of DNA replication in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 783–787 (1995).

20. Braun, V. & Bosch, V. Sequence of the murein-lipoprotein and the attachment site of the lipid. *Eur. J. Biochem.* **28,** 51–69 (1972).

21. Hantke, K. & Braun, V. Covalent Binding of Lipid to Protein. *European Journal of Biochemistry* **34,** 284–296 (1973).

22. Nakayama, H., Kurokawa, K. & Lee, B. L. Lipoproteins in bacteria: structures and biosynthetic pathways. *FEBS J.* **279,** 4247–4268 (2012).

23. Tokunaga, M., Tokunaga, H. & Wu, H. C. Post-translational modification and processing of *Escherichia coli* prolipoprotein in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **79,** 2255–2259 (1982).

24. Wu, H. C., Lai, J. S., Hayashi, S. & Giam, C. Z. Biogenesis of membrane lipoproteins in *Escherichia coli*. *Biophysical Journal* **37,** 307–315 (1982).

25. Smith, W. P., Tai, P. C. & Davis, B. D. *Bacillus licheniformis* penicillinase: cleavages and attachment of lipid during cotranslational secretion. *Proc. Natl. Acad. Sci. U.S.A.* **78,** 3501–3505 (1981).

26. Bohin, J. P. & Kennedy, E. P. Regulation of the synthesis of membrane-derived oligosaccharides in *Escherichia coli*. Assay of phosphoglycerol transferase I in vivo. *J. Biol. Chem.* **259,** 8388–8393 (1984).

27. Jackson, B. J., Bohin, J. P. & Kennedy, E. P. Biosynthesis of membrane-derived oligosaccharides: characterization of mdoB mutants defective in phosphoglycerol transferase I activity. *J. Bacteriol.* **160,** 976–981 (1984).

28. Jackson, B. J. & Kennedy, E. P. The biosynthesis of membrane-derived oligosaccharides. A membrane-bound phosphoglycerol transferase. *J. Biol. Chem.* **258,** 2394–2398 (1983).

29. Miller, K. J. & Kennedy, E. P. Transfer of phosphoethanolamine residues from phosphatidylethanolamine to the membrane-derived oligosaccharides of *Escherichia coli*. *J. Bacteriol.* **169,** 682–686 (1987).

30. Bishop, R. E. *et al*. Transfer of palmitate from phospholipids to lipid A in outer membranes of gram-negative bacteria. *EMBO J.* **19,** 5071–5080 (2000).

31. Needham, B. D. & Trent, M. S. Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nature Reviews Microbiology* **11,** 467–481 (2013).

32. Phillips, N. J. *et al*. The Lipid A from *Vibrio fischeri* Lipopolysaccharide. *J Biol Chem* **286,** 21203–21219 (2011).

33. Raetz, C. R. H., Reynolds, C. M., Trent, M. S. & Bishop, R. E. Lipid A Modification Systems in Gram-Negative Bacteria. *Annual Review of Biochemistry* **76,** 295–329 (2007).

34. Donohue-Rolfe, A. M. & Schaechter, M. Translocation of phospholipids from the inner to the outer membrane of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **77,** 1867–1871 (1980).

35. Bayer, M. E. Zones of membrane adhesion in the cryofixed envelope of *Escherichia coli*. *J. Struct. Biol.* **107,** 268–280 (1991).

36. Kellenberger, E. The 'Bayer bridges' confronted with results from improved electron microscopy methods. *Mol. Microbiol.* **4,** 697–705 (1990).

37. Doerrler, W. T., Reedy, M. C. & Raetz, C. R. An *Escherichia coli* mutant defective in lipid export. *J. Biol. Chem.* **276,** 11461–11464 (2001).

38. Doerrler, W. T., Gibbons, H. S. & Raetz, C. R. H. MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. *J. Biol. Chem.* **279,** 45102–45109 (2004).

39. Tefsen, B., Bos, M. P., Beckers, F., Tommassen, J. & de Cock, H. MsbA is not required for phospholipid transport in *Neisseria meningitidis*. *J. Biol. Chem.* **280,** 35961–35966 (2005).

40. Malinverni, J. C. & Silhavy, T. J. An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 8009–8014 (2009).

41. Yakushi, T., Masuda, K., Narita, S., Matsuyama, S. & Tokuda, H. A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nat. Cell Biol.* **2,** 212–218 (2000).

42. Carpenter, C. D. *et al.* The Vps/VacJ ABC transporter is required for intercellular spread of *Shigella flexneri*. *Infect. Immun.* **82,** 660–669 (2014).

43. Bishop, R. E. Structural biology of membrane-intrinsic β-barrel enzymes: Sentinels of the bacterial outer membrane. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1778,** 1881–1896 (2008).

44. Dekker, N. Outer-membrane phospholipase A: known structure, unknown biological function. *Mol. Microbiol.* **35,** 711–717 (2000).

45. Snijder, H. J. & Dijkstra, B. W. Bacterial phospholipase A: structure and function of an integral membrane phospholipase. *Biochim. Biophys. Acta* **1488,** 91–101 (2000).

46. Hittle, L. E. *et al.* Site-specific activity of the acyltransferases HtrB1 and HtrB2 in *Pseudomonas aeruginosa* lipid A biosynthesis. *Pathogens and Disease* **73,** ftv053 (2015).

47. Trent, M. S., Stead, C. M., Tran, A. X. & Hankins, J. V. Diversity of endotoxin and its impact on pathogenesis. *Journal of Endotoxin Research* **12,** 205–223 (2006).

48. Vorachek-Warren, M. K., Ramirez, S., Cotter, R. J. & Raetz, C. R. H. A Triple Mutant of *Escherichia coli* Lacking Secondary Acyl Chains on Lipid A. *Journal of Biological Chemistry* **277,** 14194–14205 (2002).

49. Whitfield, C. & Trent, M. S. Biosynthesis and Export of Bacterial Lipopolysaccharides*. *Annual Review of Biochemistry* **83,** 99–128 (2014).

50. Rubin, E. J., O'Brien, J. P., Ivanov, P. L., Brodbelt, J. S. & Trent, M. S. Identification of a broad family of lipid A late acyltransferases with non-canonical substrate specificity. *Mol. Microbiol.* **91,** 887–899 (2014).

51. Bainbridge, B. W. *et al.* Acyl chain specificity of the acyltransferases LpxA and LpxD and substrate availability contribute to lipid A fatty acid heterogeneity in *Porphyromonas gingivalis*. *J. Bacteriol.* **190,** 4549–4558 (2008).

52. Bartling, C. M. & Raetz, C. R. H. Crystal Structure and Acyl Chain Selectivity of *Escherichia coli* LpxD, the *N* -Acyltransferase of Lipid A Biosynthesis. *Biochemistry* **48,** 8672–8683 (2009).

53. Buetow, L., Smith, T. K., Dawson, A., Fyffe, S. & Hunter, W. N. Structure and reactivity of LpxD, the N-acyltransferase of lipid A biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 4321–4326 (2007).

54. Raetz, C. R. & Roderick, S. L. A left-handed parallel beta helix in the structure of UDP-N-acetylglucosamine acyltransferase. *Science* **270,** 997–1000 (1995).

55. Chung, H. S. & Raetz, C. R. Dioxygenases in *Burkholderia ambifaria* and *Yersinia pestis* that hydroxylate the outer Kdo unit of lipopolysaccharide. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 510–5 (2011).

56. White, K. A., Lin, S., Cotter, R. J. & Raetz, C. R. A *Haemophilus influenzae* gene that encodes a membrane bound 3-deoxy-D-manno-octulosonic acid (Kdo) kinase. Possible involvement of kdo phosphorylation in bacterial virulence. *J. Biol. Chem.* **274,** 31391–31400 (1999).

57. Jenkins, R. J. & Dotson, G. D. Dual Targeting Antibacterial Peptide Inhibitor of Early Lipid A Biosynthesis. *ACS Chemical Biology* **7,** 1170–1177 (2012).

58. Nayar, A. S. *et al*. Novel Antibacterial Targets and Compounds Revealed by a High-Throughput Cell Wall Reporter Assay. *Journal of Bacteriology* **197,** 1726–1734 (2015).

59. Anderson, M. S., Robertson, A. D., Macher, I. & Raetz, C. R. Biosynthesis of lipid A in *Escherichia coli*: identification of UDP-3-O-[(R)-3-hydroxymyristoyl]-alpha-D-glucosamine as a precursor of UDP-N2,O3-bis[(R)-3-hydroxymyristoyl]-alpha-D-glucosamine. *Biochemistry* **27,** 1908–1917 (1988).

60. Liang, X., Lee, C.-J., Zhao, J., Toone, E. J. & Zhou, P. Synthesis, Structure, and Antibiotic Activity of Aryl-Substituted LpxC Inhibitors. *Journal of Medicinal Chemistry* **56,** 6954–6966 (2013).

61. Barb, A. W. *et al*. Inhibition of lipid A biosynthesis as the primary mechanism of CHIR-090 antibiotic activity in *Escherichia coli*. *Biochemistry* **46,** 3793–3802 (2007).

62. Barb, A. W. & Zhou, P. Mechanism and inhibition of LpxC: an essential zinc-dependent deacetylase of bacterial lipid A synthesis. *Curr Pharm Biotechnol* **9,** 9–15 (2008).

63. Lee, C.-J. *et al*. Species-specific and inhibitor-dependent conformations of LpxC: implications for antibiotic design. *Chem. Biol*. **18,** 38–47 (2011).

64. McClerren, A. L. *et al*. A slow, tight-binding inhibitor of the zinc-dependent deacetylase LpxC of lipid A biosynthesis with antibiotic activity comparable to ciprofloxacin. *Biochemistry* **44,** 16574–16583 (2005).

65. Brown, M. F. *et al*. Potent inhibitors of LpxC for the treatment of Gram-negative infections. *J. Med. Chem.* **55,** 914–923 (2012).

66. Lee, C.-J. *et al*. Structural basis of the promiscuous inhibitor susceptibility of *Escherichia coli* LpxC. *ACS Chem. Biol.* **9,** 237–246 (2014).

67. Liang, X. *et al*. Syntheses, structures and antibiotic activities of LpxC inhibitors based on the diacetylene scaffold. *Bioorg. Med. Chem.* **19,** 852–860 (2011).

68. Mansoor, U. F. *et al*. Design and synthesis of potent Gram-negative specific LpxC inhibitors. *Bioorg. Med. Chem. Lett.* **21,** 1155–1161 (2011).

69. Tomaras, A. P. *et al*. LpxC inhibitors as new antibacterial agents and tools for studying regulation of lipid A biosynthesis in Gram-negative pathogens. *MBio* **5,** e01551-1514 (2014).

70. Birck, M. R., Holler, T. P. & Woodard, R. W. Identification of a Slow Tight-Binding Inhibitor of 3-Deoxy-d-manno-octulosonic Acid 8-Phosphate Synthase. *J. Am. Chem. Soc.* **122,** 9334–9335 (2000).

71. Cipolla, L. *et al*. The Kdo biosynthetic pathway toward OM biogenesis as target in antibacterial drug design and development. *Curr Drug Discov Technol* **6,** 19–33 (2009).

72. Claesson, A., Jansson, A. M., Pring, B. G., Hammond, S. M. & Ekström, B. Design and synthesis of peptide derivatives of a 3-deoxy-D-manno-2-octulosonic acid

(KDO) analogue as novel antibacterial agents acting upon lipopolysaccharide biosynthesis. *J. Med. Chem.* **30,** 2309–2313 (1987).

73. Claesson, A., Luthman, K., Gustafsson, K. & Bondesson, G. A 2-deoxy analogue of KDO as the first inhibitor of the enzyme CMP-KDO synthetase. *Biochem. Biophys. Res. Commun.* **143,** 1063–1068 (1987).

74. Pring, B. G. *et al*. Synthesis of 8-substituted derivatives of the 2-deoxy analogue of 3-deoxy-beta-D-manno-2-octulopyranosonic acid (2-deoxy-beta-KDO) as inhibitors of 3-deoxy-D-manno-octulosonate cytidylyltransferase. *J. Med. Chem.* **32,** 1069–1074 (1989).

75. Xu, X. *et al*. Structure-based design of novel inhibitors of 3-deoxy-D-manno-octulosonate 8-phosphate synthase. *Drug Des Discov* **18,** 91–99 (2003).

76. Eckford, P. D. W. & Sharom, F. J. Functional characterization of *Escherichia coli* MsbA: interaction with nucleotides and substrates. *J. Biol. Chem.* **283,** 12840–12850 (2008).

77. Reyes, C. L. & Chang, G. Structure of the ABC transporter MsbA in complex with ADP vanadate and lipopolysaccharide. *Science* **308,** 1028–1031 (2005).

78. Woebking, B. *et al*. Drug-lipid A interactions on the *Escherichia coli* ABC transporter MsbA. *J. Bacteriol.* **187,** 6363–6369 (2005).

79. Gronenberg, L. S. & Kahne, D. Development of an Activity Assay for Discovery of Inhibitors of Lipopolysaccharide Transport. *J. Am. Chem. Soc.* **132,** 2518–2519 (2010).

80. Narita, S. & Tokuda, H. Biochemical characterization of an ABC transporter LptBFGC complex required for the outer membrane sorting of lipopolysaccharides. *FEBS Letters* **583,** 2160–2164 (2009).

81. Ruiz, N., Gronenberg, L. S., Kahne, D. & Silhavy, T. J. Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. *PNAS* **105,** 5537–5542 (2008).

82. Okuda, S., Freinkman, E. & Kahne, D. Cytoplasmic ATP Hydrolysis Powers Transport of Lipopolysaccharide Across the Periplasm in *E. coli*. *Science* **338,** 1214–1217 (2012).

83. Tran, A. X., Dong, C. & Whitfield, C. Structure and Functional Analysis of LptC, a Conserved Membrane Protein Involved in the Lipopolysaccharide Export Pathway in *Escherichia coli*. *J. Biol. Chem.* **285,** 33529–33539 (2010).

84. Chng, S.-S., Gronenberg, L. S. & Kahne, D. Proteins Required for Lipopolysaccharide Assembly in *Escherichia coli* Form a Transenvelope Complex. *Biochemistry* **49,** 4565–4567 (2010).

85. Freinkman, E., Okuda, S., Ruiz, N. & Kahne, D. Regulated Assembly of the Transenvelope Protein Complex Required for Lipopolysaccharide Export. *Biochemistry* **51,** 4800–4806 (2012).

86. Freinkman, E., Chng, S.-S. & Kahne, D. The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel. *Proceedings of the National Academy of Sciences* **108,** 2486–2491 (2011).

87. Grabowicz, M., Yeh, J. & Silhavy, T. J. Dominant Negative lptE Mutation That Supports a Role for LptE as a Plug in the LptD Barrel. *Journal of Bacteriology* **195,** 1327–1334 (2013).

88. Dong, H. *et al*. Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511,** 52–56 (2014).

89. Gruss, F. *et al*. The structural basis of autotransporter translocation by TamA. *Nat. Struct. Mol. Biol.* **20,** 1318–1320 (2013).

90. Noinaj, N. *et al*. Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501,** 385–390 (2013).

91. Srinivas, N. *et al*. Peptidomimetic Antibiotics Target Outer-Membrane Biogenesis in *Pseudomonas aeruginosa*. *Science* **327,** 1010–1013 (2010).

92. Sherman, D. J., Okuda, S., Denny, W. A. & Kahne, D. Validation of inhibitors of an ABC transporter required to transport lipopolysaccharide to the cell surface in *Escherichia coli*. *Bioorganic & Medicinal Chemistry* **21,** 4846–4851 (2013).

93. Urfer, M. *et al*. A Peptidomimetic Antibiotic Targets Outer Membrane Proteins and Disrupts Selectively the Outer Membrane in *Escherichia coli*. *J. Biol. Chem*. (2015). doi:10.1074/jbc.M115.691725

94. Werneburg, M. *et al*. Inhibition of lipopolysaccharide transport to the outer membrane in *Pseudomonas aeruginosa* by peptidomimetic antibiotics. *Chembiochem* **13,** 1767–1775 (2012).

95. Needham, B. D. *et al*. Modulating the innate immune response by combinatorial engineering of endotoxin. *Proceedings of the National Academy of Sciences* **110,** 1464–1469 (2013).

96. Cullen, T. W. *et al*. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. *Science* **347,** 170–175 (2015).

97. Cullen, T. W. *et al*. *Helicobacter pylori* versus the Host: Remodeling of the Bacterial Outer Membrane Is Required for Survival in the Gastric Mucosa. *PLoS Pathog* **7,** e1002454 (2011).

98. Silipo, A. *et al*. Covalently linked hopanoid-lipid A improves outer-membrane resistance of a *Bradyrhizobium* symbiont of legumes. *Nature Communications* **5,** 5106 (2014).

99. Vamadevan, A. S. *et al*. Regulation of TLR4-associated MD-2 in intestinal epithelial cells: a comprehensive analysis. *Innate Immun* **16,** 93–103 (2010).

100. Vaishnava, S., Behrendt, C. L., Ismail, A. S., Eckmann, L. & Hooper, L. V. Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. *Proc Natl Acad Sci U S A* **105,** 20858–20863 (2008).

101. Coats, S. R. *et al*. The Lipid A Phosphate Position Determines Differential Host Toll-Like Receptor 4 Responses to Phylogenetically Related Symbiotic and Pathogenic Bacteria. *Infect. Immun.* **79,** 203–210 (2011).

102. Ernst, R. K. *et al*. Specific lipopolysaccharide found in cystic fibrosis airway *Pseudomonas aeruginosa*. *Science* **286,** 1561–1565 (1999).

103. Koropatnick, T. A. *et al*. Microbial-Factor Mediated Development in a Host-Bacterial Mutualism. *Science* **306,** 1186–1188 (2004).

104. Rader, B. A., Kremer, N., Apicella, M. A., Goldman, W. E. & McFall-Ngai, M. J. Modulation of Symbiont Lipid A Signaling by Host Alkaline Phosphatases in the Squid-*Vibrio* Symbiosis. *mBio* **3,** e00093-12-e00093-12 (2012).

105. Li, J. *et al*. Colistin: the re-emerging antibiotic for multidrug-resistant Gram-negative bacterial infections. *The Lancet infectious diseases* **6,** 589–601 (2006).

106. Falagas, M. E. *et al*. Colistin therapy for microbiologically documented multidrug-resistant Gram-negative bacterial infections: a retrospective cohort study of 258 patients. *International Journal of Antimicrobial Agents* **35,** 194–199 (2010).

107. Guilhelmelli, F. *et al*. Antibiotic development challenges: the various mechanisms of action of antimicrobial peptides and of bacterial resistance. *Frontiers in Microbiology* **4,** (2013).

108. Falagas, M. E., Rafailidis, P. I. & Matthaiou, D. K. Resistance to polymyxins: Mechanisms, frequency and treatment options. *Drug Resistance Updates* **13,** 132–138 (2010).

109. Liu, Y.-Y. *et al*. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *The Lancet Infectious Diseases* **16,** 161–168

110. McGann, P. *et al*. *Escherichia coli* Harboring *mcr-1* and *bla* $_{\text{CTX-M}}$ on a Novel IncF Plasmid: First Report of *mcr-1* in the United States. *Antimicrobial Agents and Chemotherapy* **60,** 4420–4421 (2016).

111. Roy, H. Tuning the properties of the bacterial membrane with aminoacylated phosphatidylglycerol. *IUBMB Life* **61,** 940–953 (2009).

112. Neuhaus, F. C. & Baddiley, J. A Continuum of Anionic Charge: Structures and Functions of D-Alanyl-Teichoic Acids in Gram-Positive Bacteria. *Microbiology and Molecular Biology Reviews* **67,** 686–723 (2003).

113. Needham, B. D. & Trent, M. S. Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nature Reviews Microbiology* **11,** 467–481 (2013).

114. Raetz, C. R., Reynolds, C. M., Trent, M. S. & Bishop, R. E. Lipid A Modification Systems in Gram-Negative Bacteria. *Annu Rev Biochem* **76,** 295–329 (2007).

115. Hankins, J. V., Madsen, J. A., Giles, D. K., Brodbelt, J. S. & Trent, M. S. Amino acid addition to Vibrio cholerae LPS establishes a link between surface remodeling in gram-positive and gram-negative bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **109,** 8722–7 (2012).

116. Faruque, S. M., Albert, M. J. & Mekalanos, J. J. Epidemiology, Genetics, and Ecology of Toxigenic *Vibrio cholerae*. *Microbiology and molecular biology reviews* **62,** 1301–1314 (1998).

117. Spangler, B. D. Structure and function of cholera toxin and the related *Escherichia coli* heat-labile enterotoxin. *Microbiological reviews* **56,** 622–647 (1992).

118. Guichard, A. *et al*. Cholera Toxin Disrupts Barrier Function by Inhibiting Exocyst-Mediated Trafficking of Host Proteins to Intestinal Cell Junctions. *Cell Host & Microbe* **14,** 294–305 (2013).

119. Zasloff, M. Antimicrobial peptides of multicellular organisms. *nature* **415,** 389–395 (2002).

120. Hancock, R. E. & Diamond, G. The role of cationic antimicrobial peptides in innate immune defences. *Trends in Microbiology* **8,** 402–410 (2000).

121. Ouellette, A. J. Paneth cell α-defensins in enteric innate immunity. *Cellular and Molecular Life Sciences* **68,** 2215–2229 (2011).

122. Gangarosa, E. J., Bennett, J. V. & Boring, J. R. Differentiation between vibrio cholerae and *Vibrio cholerae* biotype El Tor by the polymyxin B disc test: comparative results with TCBS, Monsur's, Mueller-Hinton and nutrient agar media. *Bulletin of the World Health Organization* **36,** 987 (1967).

123. Beyhan, S., Tischler, A. D., Camilli, A. & Yildiz, F. H. Differences in Gene Expression between the Classical and El Tor Biotypes of *Vibrio cholerae* O1. *Infection and Immunity* **74,** 3633–3642 (2006).

124. Raychoudhuri, A. *et al*. Biotyping of *Vibrio cholerae* O1: time to redefine the scheme. *Indian Journal of Medical Research* **128,** 695 (2008).

125. Hankins, J. V. *et al*. Elucidation of a novel *Vibrio cholerae* lipid A secondary hydroxy-acyltransferase and its role in innate immune recognition. *Molecular Microbiology* **81,** 1313–1329 (2011).

126. Hankins, J. V. & Trent, M. S. Secondary acylation of *Vibrio cholerae* lipopolysaccharide requires phosphorylation of Kdo. *J Biol Chem* **284,** 25804–12 (2009).

127. Herrera, C. M. *et al*. The *Vibrio cholerae* VprA-VprB Two-Component System Controls Virulence through Endotoxin Modification. *mBio* **5,** e02283-14 (2014).

128. LaRocque, R. C. *et al*. Transcriptional Profiling of *Vibrio cholerae* Recovered Directly from Patient Specimens during Early and Late Stages of Human Infection. *Infection and Immunity* **73,** 4488–4493 (2005).

129. Mandlik, A. *et al*. RNA-Seq-Based Monitoring of Infection-Linked Changes in *Vibrio cholerae* Gene Expression. *Cell Host & Microbe* **10,** 165–174 (2011).

130. Maier, T., Jenni, S. & Ban, N. Architecture of mammalian fatty acid synthase at 4.5 A resolution. *Science* **311,** 1258–1262 (2006).

131. Jung, J. *et al*. Crystal structure of the essential *Mycobacterium tuberculosis* phosphopantetheinyl transferase PptT, solved as a fusion protein with maltose binding protein. *Journal of Structural Biology* **188,** 274–278 (2014).

132. Vickery, C. R. *et al*. Structure, Biochemistry, and Inhibition of Essential 4′-Phosphopantetheinyl Transferases from Two Species of *Mycobacteria*. *ACS Chemical Biology* **9,** 1939–1944 (2014).

133. Halavaty, A. S. *et al*. Structural characterization and comparison of three acyl-carrier-protein synthases from pathogenic bacteria. *Acta Crystallographica Section D Biological Crystallography* **68,** 1359–1370 (2012).

134. Crosby, J. & Crump, M. P. The structural role of the carrier protein – active controller or passive carrier. *Natural Product Reports* **29,** 1111 (2012).

135. Beld, J., Sonnenschein, E. C., Vickery, C. R., Noel, J. P. & Burkart, M. D. The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Natural Product Reports* **31,** 61 (2014).

136. Fischbach, M. A. & Walsh, C. T. Antibiotics for Emerging Pathogens. *Science* **325,** 1089–1093 (2009).

137. Parris, K. D. *et al*. Crystal structures of substrate binding to *Bacillus subtilis* holo-(acyl carrier protein) synthase reveal a novel trimeric arrangement of molecules resulting in three active sites. *Structure* **8,** 883–895 (2000).

138. Chirgadze, N. Y., Briggs, S. L., McAllister, K. A., Fischl, A. S. & Zhao, G. Crystal structure of *Streptococcus pneumoniae* acyl carrier protein synthase: an essential enzyme in bacterial fatty acid biosynthesis. *The EMBO Journal* **19,** 5281–5287 (2000).

139. Quadri, L. E. *et al*. Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **37,** 1585–1595 (1998).

140. Reuter, K., Mofid, M. R., Marahiel, M. A. & Ficner, R. Crystal structure of the surfactin synthetase-activating enzyme Sfp: a prototype of the 4′-phosphopantetheinyl transferase superfamily. *The EMBO journal* **18,** 6823–6831 (1999).

141. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10,** 845–858 (2015).

142. Masoudi, A., Raetz, C. R. H., Zhou, P. & Pemble IV, C. W. Chasing acyl carrier protein through a catalytic cycle of lipid A production. *Nature* **505,** 422–426 (2013).

143. Agarwal, V., Lin, S., Lukk, T., Nair, S. K. & Cronan, J. E. Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. *Proceedings of the National Academy of Sciences* **109,** 17406–17411 (2012).

144. Rock, C. O., Cronan, J. E. & Armitage, I. M. Molecular properties of acyl carrier protein derivatives. *Journal of Biological Chemistry* **256,** 2669–2674 (1981).

145. Alberts, A. W. & Vagelos, P. R. Acyl Carrier Protein viii. Studies of acyl carrier protein and coenzyme a in *Escherichia coli* pantothenate or β-alanine auxotrophs. *Journal of Biological Chemistry* **241,** 5201–5204 (1966).

146. Kelly, W. L. *et al*. Characterization of the Aminocarboxycyclopropane-Forming Enzyme CmaC †. *Biochemistry* **46,** 359–368 (2007).

147. Keating, D. H., Carey, M. R. & Cronan, J. E. The unmodified (apo) form of *Escherichia coli* acyl carrier protein is a potent inhibitor of cell growth. *Journal of Biological Chemistry* **270,** 22229–22235 (1995).

148. De Lay, N. R. & Cronan, J. E. A genome rearrangement has orphaned the *Escherichia coli* K-12 AcpT phosphopantetheinyl transferase from its cognate *Escherichia coli* O157:H7 substrates: *E. coli* AcpT. *Molecular Microbiology* **61,** 232–242 (2006).

149. Mofid, M. R., Finking, R., Essen, L. O. & Marahiel, M. A. Structure-Based Mutational Analysis of the 4‘-Phosphopantetheinyl Transferases Sfp from *Bacillus*

subtilis : Carrier Protein Recognition and Reaction Mechanism [†,‡]. *Biochemistry* **43,** 4128–4136 (2004).

150. Naggert, J. *et al*. Cloning, Sequencing, and Characterization of *Escherichia coli* Thioesterase II. **266,** 11044–11050 (1991).

151. Cho, H. & Cronan, J. E. Escherichia coli Thioesterase I, molecular cloning and sequencing of the structural gene and identification as a periplasmic enzyme. *The Journal of Biological Chemistry* **268,** 9238–9245 (1993).

152. Nie, L., Ren, Y. & Schulz, H. Identification and Characterization of *Escherichia coli* Thioesterase III That Functions in Fatty Acid β-Oxidation [†]. *Biochemistry* **47,** 7744–7751 (2008).

153. Cantu, D. C., Chen, Y., Lemons, M. L. & Reilly, P. J. ThYme: a database for thioester-active enzymes. *Nucleic Acids Research* **39,** D342–D346 (2011).

154. Schwarzer, D., Mootz, H. D., Linne, U. & Marahiel, M. A. Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proceedings of the National Academy of Sciences* **99,** 14083–14088 (2002).

155. Kotowska, M. & Pawlik, K. Roles of type II thioesterases and their application for secondary metabolite yield improvement. *Applied Microbiology and Biotechnology* **98,** 7735–7746 (2014).

156. Conti, E., Stachelhaus, T., Marahiel, M. A. & Brick, P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *The EMBO journal* **16,** 4174–4183 (1997).

157. Jogl, G. & Tong, L. Crystal Structure of Yeast Acetyl-Coenzyme A Synthetase in Complex with AMP †. *Biochemistry* **43,** 1425–1431 (2004).

158. Gulick, A. M., Starai, V. J., Horswill, A. R., Homick, K. M. & Escalante-Semerena, J. C. The 1.75 Å Crystal Structure of Acetyl-CoA Synthetase Bound to Adenosine-5‘-propylphosphate and Coenzyme A †. *Biochemistry* **42,** 2866–2873 (2003).

159. Perego, M. *et al*. Incorporation of D-alanine into lipoteichoic acid and wall teichoic acid in *Bacillus subtilis*. *The Journal of biological chemistry* **270,** 15598–15606 (1995).

160. Conti, E., Franks, N. P. & Brick, P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* **4,** 287–298 (1996).

161. Kim, D. J. *et al*. Crystal Structure of Lipoate-Protein Ligase A Bound with the Activated Intermediate: insights into interaction with lipoyl domains. *Journal of Biological Chemistry* **280,** 38081–38089 (2005).

162. Fujiwara, K. *et al*. Crystal Structure of Lipoate-Protein Ligase A from Escherichia coli: determination of the lipoic acid-binding site. *Journal of Biological Chemistry* **280,** 33645–33651 (2005).

163. O’Donoghue, P. & Luthey-Schulten, Z. On the Evolution of Structure in Aminoacyl-tRNA Synthetases. *Microbiology and Molecular Biology Reviews* **67,** 550–573 (2003).

164. Stachelhaus, T., Mootz, H. D. & Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chemistry & Biology* **6,** 493–505 (1999).

165. Eppelmann, K., Stachelhaus, T. & Marahiel, M. A. Exploitation of the Selectivity-Conferring Code of Nonribosomal Peptide Synthetases for the Rational Design of Novel Peptide Antibiotics †. *Biochemistry* **41,** 9718–9726 (2002).

166. Agüero-Chapin, G., Pérez-Machado, G., Sánchez-Rodríguez, A., Santos, M. M. & Antunes, A. in *Nonribosomal Peptide and Polyketide Biosynthesis* (ed. Evans, B. S.) **1401,** 253–272 (Springer New York, 2016).

167. Rausch, C. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Research* **33,** 5799–5808 (2005).

168. Baranašić, D. *et al*. Predicting substrate specificity of adenylation domains of nonribosomal peptide synthetases and other protein properties by latent semantic indexing. *Journal of Industrial Microbiology & Biotechnology* **41,** 461–467 (2014).

169. Balibar, C. J., Vaillancourt, F. H. & Walsh, C. T. Generation of D Amino Acid Residues in Assembly of Arthrofactin by Dual Condensation/Epimerization Domains. *Chemistry & Biology* **12,** 1189–1200 (2005).

170. Osman, K. T., Du, L., He, Y. & Luo, Y. Crystal Structure of Bacillus cereus d-Alanyl Carrier Protein Ligase (DltA) in Complex with ATP. *Journal of Molecular Biology* **388,** 345–355 (2009).

171. Reichmann, N. T., Cassona, C. P. & Grundling, A. Revised mechanism of D-alanine incorporation into cell wall polymers in Gram-positive bacteria. *Microbiology* **159,** 1868–1877 (2013).

172. Percy, M. G. & Gründling, A. Lipoteichoic Acid Synthesis and Function in Gram-Positive Bacteria. *Annual Review of Microbiology* **68,** 81–100 (2014).

173. Yonus, H. *et al*. Crystal Structure of DltA: implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains. *Journal of Biological Chemistry* **283,** 32484–32491 (2008).

174. Du, L., He, Y. & Luo, Y. Crystal Structure and Enantiomer Selection by D -Alanyl Carrier Protein Ligase DltA from *Bacillus cereus* †. *Biochemistry* **47,** 11473–11480 (2008).

175. Trauger, J. W. & Walsh, C. T. Heterologous expression in *Escherichia coli* of the first module of the nonribosomal peptide synthetase for chloroeremomycin, a vancomycin-type glycopeptide antibiotic. *Proceedings of the National Academy of Sciences* **97,** 3112–3117 (2000).

176. McQuade, T. J. *et al*. A nonradioactive high-throughput assay for screening and characterization of adenylation domains for nonribosomal peptide combinatorial biosynthesis. *Analytical Biochemistry* **386,** 244–250 (2009).

177. Altschul, S. F. *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25,** 3389–3402 (1997).

178. Tanovic, A., Samel, S. A., Essen, L.-O. & Marahiel, M. A. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* **321,** 659–663 (2008).

179. Sundlov, J. A., Shi, C., Wilson, D. J., Aldrich, C. C. & Gulick, A. M. Structural and Functional Investigation of the Intermolecular Interaction between NRPS

Adenylation and Carrier Protein Domains. *Chemistry & Biology* **19,** 188–198 (2012).

180. Mitchell, C. A., Shi, C., Aldrich, C. C. & Gulick, A. M. Structure of PA1221, a Nonribosomal Peptide Synthetase Containing Adenylation and Peptidyl Carrier Protein Domains. *Biochemistry* **51,** 3252–3263 (2012).
181. Hughes, A. J. & Keatinge-Clay, A. Enzymatic Extender Unit Generation for In Vitro Polyketide Synthase Reactions: Structural and Func-tional Showcasing of *Streptomyces coelicolor* MatB. *Chemistry & Biology* **18,** 165–176 (2011).
182. Lai, J. R., Koglin, A. & Walsh, C. T. Carrier Protein Structure and Recognition in Polyketide and Nonribosomal Peptide Biosynthesis †. *Biochemistry* **45,** 14869–14879 (2006).
183. Drake, E. J., Duckworth, B. P., Neres, J., Aldrich, C. C. & Gulick, A. M. Biochemical and Structural Characterization of Bisubstrate Inhibitors of BasE, the Self-Standing Nonribosomal Peptide Synthetase Adenylate-Forming Enzyme of Acinetobactin Synthesis,. *Biochemistry* **49,** 9292–9305 (2010).
184. Zhou, Z., Lai, J. R. & Walsh, C. T. Directed evolution of aryl carrier proteins in the enterobactin synthetase. *Proceedings of the National Academy of Sciences* **104,** 11621–11626 (2007).
185. Manning, J. M., Merrifield, N. E., Jones, W. M. & Gotschlich, E. C. Inhibition of bacterial growth by β-chloro-D-alanine. *Proceedings of the National Academy of Sciences* **71,** 417–421 (1974).

186. Hishikawa, D. *et al*. Discovery of a lysophospholipid acyltransferase family essential for membrane asymmetry and diversity. *Proceedings of the National Academy of Sciences* **105,** 2830–2835 (2008).

187. Lands, W. E. & others. Metabolism of glycerolipides; a comparison of lecithin and triglyceride synthesis. *J Biol Chem* **231,** 883–888 (1958).

188. Oger, P. M. in *High Pressure Bioscience* (eds. Akasaka, K. & Matsuki, H.) **72,** 383–403 (Springer Netherlands, 2015).

189. Zhang, Y.-M. & Rock, C. O. Membrane lipid homeostasis in bacteria. *Nat. Rev. Microbiol*. **6,** 222–233 (2008).

190. Ernst, R., Ejsing, C. S. & Antonny, B. Homeoviscous Adaptation and the Regulation of Membrane Lipids. *Journal of Molecular Biology* (2016). doi:10.1016/j.jmb.2016.08.013

191. Clementz, T., Bednarski, J. & Raetz, C. R. H. Function of the htrB high temperature gene of *Escherichia coli* in the acylation of lipid A. *Journal of Biological Chemistry* **271,** 12095–12102 (1996).

192. Vorachek-Warren, M. K. An *Escherichia coli* Mutant Lacking the Cold Shock-induced Palmitoleoyltransferase of Lipid A Biosynthesis. absence of unsaturated acyl chains and antibiotic hypersensitivity at 12 degreesC. *Journal of Biological Chemistry* **277,** 14186–14193 (2002).

193. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide Endotoxins. *Annual Review of Biochemistry* **71,** 635–700 (2002).

194. Boll, J. M. *et al*. Reinforcing Lipid A Acylation on the Cell Surface of *Acinetobacter baumannii* Promotes Cationic Antimicrobial Peptide Resistance and Desiccation Survival. *MBio* **6,** e00478-415 (2015).

195. Somerville, J. E., Cassiano, L. & Darveau, R. P. *Escherichia coli* msbB gene as a virulence factor and a therapeutic target. *Infection and immunity* **67,** 6583–6590 (1999).

196. Low, K. *et al*. Lipid A mutant *Salmonella* with suppressed virulence and TNFα induction retain tumor-targeting in vivo. *Nature Biotechnology* **17,** 37–41 (1999).

197. Khan, S. *et al*. A lethal role for lipid A in *Salmonella* infections. *Molecular Microbiology* **29,** 571–579 (1999).

198. Somerville, J. E., Cassiano, L., Bainbridge, B., Cunningham, M. D. & Darveau, R. P. A novel *Escherichia coli* lipid A mutant that produces an antiinflammatory lipopolysaccharide. *Journal of Clinical Investigation* **97,** 359–365 (1996).

199. Watson, P. R. *et al*. Mutation of waaN Reduces *Salmonella enterica* Serovar Typhimurium-Induced Enteritis and Net Secretion of Type III Secretion System 1-Dependent Proteins. *Infection and immunity* **68,** 3768–3771 (2000).

200. d'Hauteville, H. *et al*. Two msbB Genes Encoding Maximal Acylation of Lipid A Are Required for Invasive *Shigella flexneri* to Mediate Inflammatory Rupture and Destruction of the Intestinal Epithelium. *The Journal of Immunology* **168,** 5240–5251 (2002).

201. Ranallo, R. T. *et al*. Virulence, Inflammatory Potential, and Adaptive Immunity Induced by *Shigella flexneri* msbB Mutants. *Infection and Immunity* **78,** 400–412 (2010).

202. Tran, A. X. *et al*. Resistance to the Antimicrobial Peptide Polymyxin Requires Myristoylation of *Escherichia coli* and *Salmonella* typhimurium Lipid A. *Journal of Biological Chemistry* **280,** 28186–28194 (2005).

203. Clements, A. *et al*. Secondary Acylation of *Klebsiella pneumoniae* Lipopolysaccharide Contributes to Sensitivity to Antibacterial Peptides. *Journal of Biological Chemistry* **282,** 15569–15577 (2007).

204. Chatterjee, S. N. & Chaudhuri, K. Lipopolysaccharides of *Vibrio cholerae*. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1639,** 65–79 (2003).

205. Baldridge, J. R. & Crane, R. T. Monophosphoryl Lipid A (MPL) Formulations for the Next Generation of Vaccines. *Methods* **19,** 103–107 (1999).

206. Needham, B. D. *et al*. Modulating the innate immune response by combinatorial engineering of endotoxin. *Proceedings of the National Academy of Sciences* **110,** 1464–1469 (2013).

207. Brabetz, W., Muller-Loennies, S., Holst, O. & Brade, H. Deletion of the heptosyltransferase genes rfaC and rfaF in *Escherichia coli* K-12 results in an Re-type lipopolysaccharide with a high degree of 2-aminoethanol phosphate substitution. *European Journal of Biochemistry* **247,** 716–724 (1997).

208. Wang, J., Ma, W., Wang, Z., Li, Y. & Wang, X. Construction and Characterization of an *Escherichia coli* Mutant Producing Kdo2-Lipid A. *Marine Drugs* **12,** 1495–1511 (2014).

209. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97,** 6640–6645 (2000).

210. Henderson, J. C., O'Brien, J. P., Brodbelt, J. S. & Trent, M. S. Isolation and Chemical Characterization of Lipid A from Gram-negative Bacteria. *Journal of Visualized Experiments* (2013). doi:10.3791/50623

211. Bligh, E. G. & Dyer, W. J. A rapid method of total lipid extraction and purification. *Canadian journal of biochemistry and physiology* **37,** 911–7 (1959).

212. Liu, D. & Reeves, P. R. *Escherichia coli* K12 regains its O antigen. *Microbiology* **140,** 49–57 (1994).

213. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. . Predicting transmembrane protein topology with a hidden markov model: application to complete genomes Edited by F. Cohen. *Journal of Molecular Biology* **305,** 567–580 (2001).

214. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* **8,** 785–786 (2011).

215. Smith, P. K. *et al*. Measurement of protein using bicinchoninic acid. *Analytical Biochemistry* **150,** 76–85 (1985).

216. Neron, B. *et al*. Mobyle: a new full web bioinformatics framework. *Bioinformatics* **25,** 3005–3011 (2009).

217. Heath, R. J. & Rock, C. O. A conserved histidine is essential for glycerolipid acyltransferase catalysis. *Journal of Bacteriology* **180,** 1425–1430 (1998).

218. Rottig, A. & Steinbuchel, A. Acyltransferases in Bacteria. *Microbiology and Molecular Biology Reviews* **77,** 277–321 (2013).

219. Six, D. A., Carty, S. M., Guan, Z. & Raetz, C. R. H. Purification and Mutagenesis of LpxL, the Lauroyltransferase of *Escherichia coli* Lipid A Biosynthesis †. *Biochemistry* **47,** 8623–8637 (2008).

220. Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiology Reviews* **40,** 133–159 (2016).

221. Roy, H. & Ibba, M. Broad Range Amino Acid Specificity of RNA-dependent Lipid Remodeling by Multiple Peptide Resistance Factors. *Journal of Biological Chemistry* **284,** 29677–29683 (2009).

222. Dare, K., Shepherd, J., Roy, H., Seveau, S. & Ibba, M. LysPGS formation in *Listeria monocytogenes* has broad roles in maintaining membrane integrity beyond antimicrobial peptide resistance. *Virulence* **5,** 534–546 (2014).

223. Arendt, W., Groenewold, M. K., Hebecker, S., Dickschat, J. S. & Moser, J. Identification and Characterization of a Periplasmic Aminoacyl-phosphatidylglycerol Hydrolase Responsible for *Pseudomonas aeruginosa* Lipid Homeostasis. *Journal of Biological Chemistry* **288,** 24717–24730 (2013).

224. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research* **44,** D457–D462 (2016).

225. Goarant, C. *et al*. Molecular epidemiology of *Vibrio nigripulchritudo*, a pathogen of cultured penaeid shrimp (*Litopenaeus stylirostris*) in New Caledonia. *Systematic and Applied Microbiology* **29,** 570–580 (2006).

226. Goudenège, D. *et al*. Comparative genomics of pathogenic lineages of *Vibrio nigripulchritudo* identifies virulence-associated traits. *The ISME journal* **7,** 1985–1996 (2013).

227. Le Roux, F. *et al*. Virulence of an emerging pathogenic lineage of *Vibrio nigripulchritudo* is dependent on two plasmids: Plasmids and virulence in *Vibrio nigripulchritudo*. *Environmental Microbiology* **13,** 296–306 (2011).

228. Walling, E., Vourey, E., Ansquer, D., Beliaeff, B. & Goarant, C. *Vibrio nigripulchritudo* monitoring and strain dynamics in shrimp pond sediments. *Journal of Applied Microbiology* (2009). doi:10.1111/j.1365-2672.2009.04601.x

229. Yildiz, F., Fong, J., Sadovskaya, I., Grard, T. & Vinogradov, E. Structural Characterization of the Extracellular Polysaccharide from *Vibrio cholerae* O1 El-Tor. *PLoS ONE* **9,** e86751 (2014).

230. Dzieciatkowska, M. *et al*. Mass Spectrometric Analysis of Intact Lipooligosaccharide: Direct Evidence for *O* -Acetylated Sialic Acids and Discovery of *O* -Linked Glycine Expressed by *Campylobacter jejuni* †. *Biochemistry* **46,** 14704–14714 (2007).

231. Molinaro, A. *et al*. Full structural characterization of *Shigella flexneri* M90T serotype 5 wild-type R-LPS and its galU mutant: glycine residue location in the inner core of the lipopolysaccharide. *Glycobiology* **18,** 260–269 (2007).

232. Engskog, M. K. R., Deadman, M., Li, J., Hood, D. W. & Schweda, E. K. H. Detailed structural features of lipopolysaccharide glycoforms in nontypeable *Haemophilus influenzae* strain 2019. *Carbohydrate Research* **346,** 1241–1249 (2011).

233. Henderson, J. C. *et al*. Antimicrobial Peptide Resistance of *Vibrio cholerae* Results From an LPS Modification Pathway Related to Non-ribosomal Peptide Synthetases. *ACS Chemical Biology* 140728100653003 (2014). doi:10.1021/cb500438x

234. Pride, A. C., Herrera, C. M., Guan, Z., Giles, D. K. & Trent, M. S. The Outer Surface Lipoprotein VolA Mediates Utilization of Exogenous Lipids by *Vibrio cholerae*. *mBio* **4,** e00305-13-e00305-13 (2013).

235. Pride, A. C., Guan, Z. & Trent, M. S. Characterization of the *Vibrio cholerae* VolA Surface-Exposed Lipoprotein Lysophospholipase. *Journal of Bacteriology* **196,** 1619–1626 (2014).

236. Giles, D. K., Hankins, J. V., Guan, Z. & Trent, M. S. Remodelling of the *Vibrio cholerae* membrane by incorporation of exogenous fatty acids from host and aquatic environments: Remodelling of the *V. cholerae* membrane. *Molecular Microbiology* **79,** 716–728 (2011).

237. Samanta, P., Ghosh, P., Chowdhury, G., Ramamurthy, T. & Mukhopadhyay, A. K. Sensitivity to Polymyxin B in El Tor *Vibrio cholerae* O1 Strain, Kolkata, India. *Emerging Infectious Diseases* **21,** 2100–2102 (2015).

238. Bina, X. R., Howard, M. F., Ante, V. M. & Bina, J. E. *Vibrio cholerae* LeuO links the ToxR regulon to the expression of lipid A remodeling genes. *Infection and Immunity* IAI.00445-16 (2016). doi:10.1128/IAI.00445-16

239. Doss, M., White, M. R., Tecle, T. & Hartshorn, K. L. Human defensins and LL-37 in mucosal immunity. *Journal of Leukocyte Biology* **87,** 79–92 (2009).

240. Huq, A. *et al*. Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. *Applied and Environmental Microbiology* **45,** 275–283 (1983).

241. Rawlings, T. K., Ruiz, G. M. & Colwell, R. R. Association of *Vibrio cholerae* O1 El Tor and O139 Bengal with the Copepods Acartia tonsa and Eurytemora affinis. *Applied and Environmental Microbiology* **73,** 7926–7933 (2007).

242. Scheich, C., Kummel, D., Soumailakakis, D., Heinemann, U. & Bussow, K. Vectors for co-expression of an unrestricted number of proteins. *Nucleic Acids Research* **35,** e43–e43 (2007).

243. Shaw, J. B. *et al*. Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation. *Journal of the American Chemical Society* **135,** 12646–12651 (2013).

244. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276,** 307–326 (1997).

245. Brunger, A. T. *et al*. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica Section D: Biological Crystallography* **54,** 905–921 (1998).

246. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallographica Section D Biological Crystallography* **61,** 458–464 (2005).

247. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* **3,** 1171–1179 (2008).

248. Adams, P. D. *et al*. *PHENIX* : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography* **66,** 213–221 (2010).

249. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D: Biological Crystallography* **53,** 240–255 (1997).

250. Emsley, P. & Cowtan, K. *Coot* : model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography* **60,** 2126–2132 (2004).

251. Touze, T., Tran, A. X., Hankins, J. V., Mengin-Lecreulx, D. & Trent, M. S. Periplasmic phosphorylation of lipid A is linked to the synthesis of undecaprenyl phosphate. *Molecular microbiology* **67,** 264–77 (2008).

252. Liu, J., Chen, C.-Y., Shiomi, D., Niki, H. & Margolin, W. Visualization of bacteriophage P1 infection by cryo-electron tomography of tiny *Escherichia coli*. *Virology* **417,** 304–311 (2011).

253. Vanderheiden, G. J., Fairchild, A. C. & Jago, G. R. Construction of a laboratory press for use with the French pressure cell. *Applied microbiology* **19,** 875 (1970).

254. Dounce, A. L., Witter, R. F., Monty, K. J., Pate, S. & Cottone, M. A. A method for isolating intact mitochondria and nuclei from the same homogenate, and the influence of mitochondrial destruction on the properties of cell nuclei. *The Journal of biophysical and biochemical cytology* **1,** 139–153 (1955).