Molecular Basis of Lipid Acyl Chain Selection by the Integral Outer Membrane Phospholipid:Lipid A Palmitoyltransferase PagP from *Escherichia coli.*

Mohammad Adil Khan Biochemistry, McMaster University, 2010

Molecular Basis of Lipid Acyl Chain Selection by the Integral Outer Membrane Phospholipid:Lipid A Palmitoyl Transferase PagP from *Escherichia coli*.

By

Mohammad Adil Khan, B.Sc., M.Sc., CCRA

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

Mohammad Adil Khan Biochemistry, McMaster University, 2010

McMaster University

© Copyright by Mohammad Adil Khan, January 2010

Doctor of Philosophy (2010) McMaster University, Hamilton, Ontario

(Biochemistry)

TITLE: Molecular Basis of Lipid Acyl Chain Selection by the Integral Outer Membrane Phospholipid:Lipid A Palmitoyl Transferase PagP from *Escherichia coli*.

AUTHOR: Mohammad Adil Khan, B.Sc, M.Sc. (University of Toronto), CCRA (McMaster University)

SUPERVISOR: Russell E. Bishop

NUMBER OF PAGES: xxiv, 235

Acknowledgements:

My interest in science began at a very young age. Obsessing over National Geographic documentaries or pretending to comprehend "A brief history of Time" by Stephen Hawking are a few highlights of my formative years. In school I was never a great student. I struggled with most of my courses, except for the sciences. Whether it was staring at stages of mitosis under a light microscope or reacting different compounds to generate funky colors in the ill equipped chemistry laboratory, science generated an unfamiliar passion. The years passed and I became an undergraduate at the University of Toronto (U of T). Here in my third year I ended up in the most unusual class. While the material was exceedingly difficult, the enthusiasm of the professor somehow made the experience enjoyable. In the following summer I found myself involved in an undergraduate research project in the same professor's lab. Eight years and nearly three degrees later I am still here, willingly. The main reason is the relationship I enjoy with my supervisor, Dr. Russell Bishop.

Russ and me have known each other nearly a decade and he has been a source of constant inspiration and support throughout this journey. I do not remember when the formalities of the supervisor-student relationship ended and a deep-rooted friendship began, but it has made all the difference in the world to me. I look back at the fond memories that I have shared with the many past and present members of the Bishop lab

and I sorely miss all of them. Under Russ's guidance we have all come to think of each other as family and share an unbreakable bond.

Many other people have helped me through this voyage. Namely my committee members Dr. Richard Epand and Dr. Paul Berti, whose timely guidance have helped me be successful in science. Thriving collaborations with Dr. Gilbert Privé and Dr. Régis Pomès at the U of T, Dr. Fraser Hof at University of British Columbia and Dr. Robert Woody at Colorado State University have been monumental in achieving our research goals. Furthermore, the constant moral support I have received from Dr. Lori Burrows has kept me going in those moments of weakness.

Lastly, but most importantly, the gratitude I owe to my family cannot be expressed in words. The love in my mother's eyes and the strength in my father's words have been the key ingredients for this accomplishment. The one person over the years who has made me realize my potential and has been the single driving force behind all my triumphs is undoubtedly my wife. I consider myself as one of the luckiest individual as life was generous in granting me my two true loves.

TABLE OF CONTENTS:

List of Figures:

List of Acronyms:

Acronym	Name
ACP	Acyl carrier protein
BMe	β-mercaptoethanol
CAMP	Cationic Anti Microbial Peptide
CD	Circular Dichroism
$Cu(OP)_3$	Copper-phenanthroline
bBBr	Dibromobimane
DDM	Dodecylmaltoside
DPC	Dodecylphosphocholine
DSC	Differential Scanning Calorimetry
EDTA	Ethylenediaminetetraacetic acid
Gdn-HCl	Guanidine Hydrochloride
GlcNAc	*N*-acetylglucosamine
IM	Inner Membrane
IPTG	Isopropyl-β-D-thiogalactopyranoside

L-Ara4-N	L-4-aminoarabinose
LDAO	Lauroyldimethylamine N-oxide
LPS	Lipopolysaccharide
MD	Molecular Dynamics
MNBS	Methyl-*p*-nitrobenzenesulfonate
MurNAc	*N*-acetylmuramic
NMR	Nuclear Magnetic Resonance
OM	Outer Membrane
OMP	Outer Membrane Protein
PtdEtn	Phosphatidylethanolamine
pEtN	Phosphoethanolamine
PtdCho	Phosphatidylcholine
PtdGro	Phosphatidylglycerol
TLR4	Toll like receptor 4
UDP-GlcNAc	UDP-*N*-acetylglucosamine
Und-PPi	Undecaprenylpyrophosphate

List Of Tables:

ABSTRACT:

The role of membrane-intrinsic enzymes of lipid metabolism in complex biological processes is being realized through comprehensive structure function studies. Detailed analysis of substrate-enzyme interactions occurring within the restrictive membrane environment has proved to be exceedingly challenging. Using detergent micelles, we describe a detailed model for substrate recognition and binding by the outer-membrane intrinsic enzyme PagP from *Escherichia coli*. PagP is an 8-stranded antiparallel β-barrel that transfers a palmitoyl group from a phospholipid molecule to lipid A, the endotoxin component of lipopolysaccharide. This simple modification provides bacterial resistance to host antimicrobial peptides and attenuates the inflammatory response signalled through the host toll-like receptor 4 pathway. We describe a molecular embrasure and a crenel, which display weakened transmembrane β-strand hydrogen bonding, to provide site-specific routes for lateral entry of substrates into the PagP active site. A Tyr147 localized to the L4 loop gates the entry of the phospholipid substrate through the crenel, while lipid A enters via the embrasure. The side chains of the catalytic residues that are located in the extracellular loops point towards the central axis of the enzyme, directly above the active site. An acyl-chain binding pocket known as the hydrocarbon ruler is buried within the transmembrane β barrel structure, and is optimized to accommodate a 16-carbon saturated palmitate chain. The hydrocarbon ruler, therefore, accounts for PagP's stringent selectivity for a palmitate chain. Substituting Gly88 lining the floor of the hydrocarbon ruler with residues

possessing linear, unbranched, aliphatic side chains changes the selectivity of PagP to utilize shorter acyl chains. The serendipitous discovery of an exciton interaction between Trp66 and Tyr26 at the floor of the hydrocarbon ruler provides an intrinsic spectroscopic probe to monitor the methylene unit acyl-chain resolution of PagP. A compromised acyl chain resolution of the Gly88Cys mutant is attributed to an unexpected decrease of the Cys sulfhydryl group pKa within the β-barrel interior, resulting in a burying of a charged thiolate within the PagP core. The structural perturbation associated with the Cys thiolate extinguishes the exciton and expands the acyl-chain selectivity. These molecular details of lateral lipid diffusion and acyl-chain selection provide the first such example for any membrane-intrinsic enzyme of lipid metabolism.

1.0 Introduction:

The molecular basis of enzyme-lipid interactions in a membrane-intrinsic environment is poorly understood because most membrane bound enzymes of lipid metabolism have proven refractory to structural determination. However, robust enzymes from the outer membranes (OM) of Gram-negative bacteria have proven to be an exception, and are now helping to reveal detailed molecular insights into long-standing unresolved structural questions of lipid metabolism. For example, how does a membrane-intrinsic enzyme measure the length of an acyl-chain in its lipid substrate? How do lipid substrates in the membrane plane access enzyme active sites? What are the conformational changes associated with the catalytic cycle of membrane-bound enzymes? Such questions are now being answered by recent structural studies of OM enzymes of Gram-negative bacteria. These structural studies have laid the groundwork necessary to predict, understand and define the principles of membrane protein enzymology.

My research interests have been to answer such questions, using an enzyme localized to the OM of Gram-negative pathogens. The enzyme is responsible for a specific modification of endotoxin and may prove to be of clinical significance. I will present the findings of my research in the following chapters by addressing substrate specificity and substrate entry during catalysis.

1.1: Gram-Negative Cell Envelope:

The cell envelope unique to Gram-negative bacteria constitutes the perimeter defence against cationic anti microbial peptides (CAMPs) and hydrophobic antibiotics. Structurally, it is composed of three parts: an inner membrane (IM) containing a phospholipid bilayer, the periplasmic space that contains a thin peptidoglycan layer, and an outer membrane (OM) (figure 1.1) (*1*).

1.1.1: Inner membrane:

Like the plasma membrane of eukaryotic cells, the bacterial IM is significant because it generates an electrochemical gradient responsible for generating the energy to drive all cellular functions. The phospholipid bilayer of the IM is permeable to lipophilic compounds. Therefore, CAMPs will cause membrane depolarization by puncturing holes in the phospholipid bilayer, effectively killing the cell. The proteins of the IM govern many necessary cellular functions ranging from lipid, protein and nutrients transport, cell signalling and several metabolic functions.

1.1.2: Peptidogylcan and Periplasmic Space:

The space between the OM and the IM contains a thin peptidoglycan layer that protects the IM from rupturing under the influence of high osmotic pressure. The peptidoglycan of *Escherichia coli* (*E. coli*) consists of alternating residues of β-(1, 4) linked *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues.

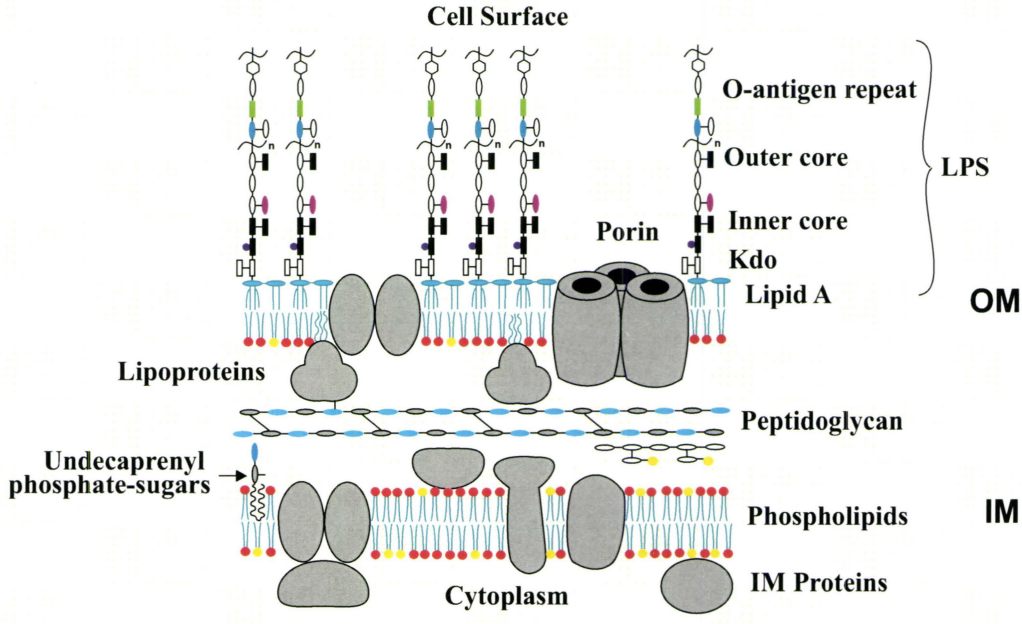

Figure 1.1 Gram-negative cell envelope: The IM is a phospholipid bilayer. This is followed by the periplasmic space, which includes a thin peptidoglycan layer. The OM is the external barrier for the bacterium and is a bilayer that has distinct lipid asymmetry between its leaflets. The inner leaflet is composed of phospholipids, while the outer leaflet is made up of LPS [Figure from Raetz *et al.*, 2007].

Each MurNAc is attached to a short (4- to 5-residue) amino acid chain, containing L and D-alanine, D-glutamic acid, and *meso*-diaminopimelic acid. The peptidogylcan is secured to the OM by Braun lipoprotein (*2*), and is surrounded in a gelatinous material called periplasm. The periplasm contains enzymes involved in nutrient and protein transport.

1.1.3: Outer Membrane:

The OM is a protective structure that is unique to Gram-negative bacteria. It possesses a distinct asymmetry of lipids between its inner and outer leaflets. While the inner leaflet is mainly composed of 80% phosphatidylethanolamine (PtdEtn) and 20% phosphatidylglycerol (PtdGro) and cardiolipin, the outer leaflet consists almost entirely of lipopolysaccharide (LPS). The low fluidity of LPS hydrocarbon domain and strong lateral interactions between neighbouring molecules (*3*) creates a barrier against spontaneous diffusion of lipophilic compounds (*4*). The OM houses a unique class of proteins with wide ranging functions, which are essential to bacterial survival.

1.1.3.1: LPS Structure:

A unique molecule of the OM resides in the outer leaflet of the OM and is known as LPS. LPS is composed of three parts: The lipid A hydrophobic anchor, the interconnecting core-oligosaccharide, and the O-antigen repeats (figure 1.2). The two Kdo sugars are connected to the lipid A molecule through a ketosidic linkage that is

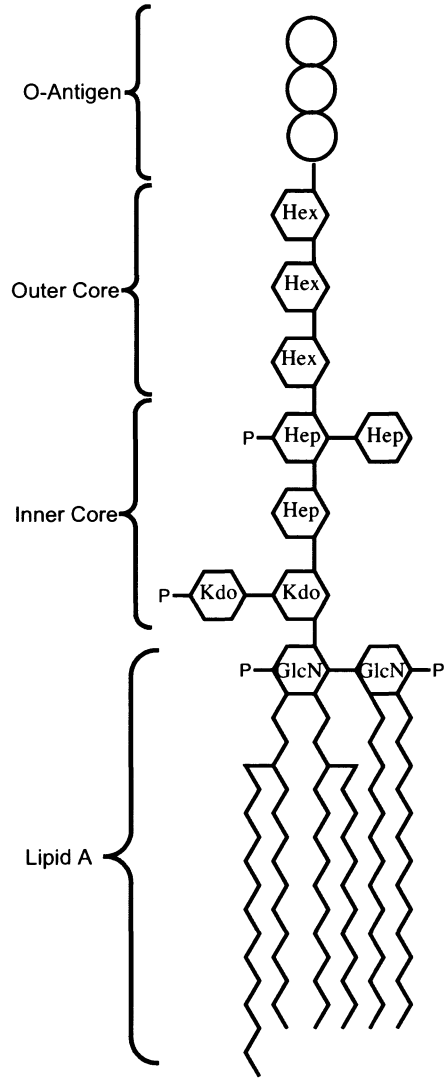

Figure 1.2 Schematic representation of LPS: Lipid A is the essential component of LPS and is required for growth. The core-sugars are conserved, but the O-antigen repeats are highly variable. In the figure *N*-acetylglucosamine, 3-deoxy-D-*manno*-2-octulosonic acid, Heptose and Hexose sugars are represented by GlcN, Kdo, Hep and Hex, respectively.

susceptible to mild-acid hydrolysis (*5*). Kdo_2-lipid A is an essential constituent of the LPS and is the simplest Re-chemotype of LPS that supports growth of *E. coli* (*1*) under laboratory conditions. Bacterial strains that exhibit the Re-chemotype are referred to as deep rough mutants and are typically not found in natural isolates. Attachment of core sugars and the O-antigen to lipid A generates smooth-LPS, which is necessary for bacterial survival in different ecological niches (*6*).

1.1.3.2: LPS: Clinical Significance and Inflammatory Response:

LPS is shed from the bacterial OM during an infection and induces an inflammatory response in the host through its lipid A component. It is detected via the Toll like receptor 4 (TLR4) located on mononuclear cells and neutrophils. The release of host chemokines, cytokines and antimicrobial agents results in an effective innate immune response against the invading bacteria. However, persistent infections can lead to the overwhelming production of tumor necrosis factor α (TNF-α) and interleukin 1β (IL-1β), resulting in septic shock that can be fatal (*7*). Similarly, activation of TLR4 by lipid A may result in an adaptive immune response by initiating the expression on antigen presenting cell surfaces of co stimulatory molecules (*8*). Recently, it was reported that the levels of TLR4 receptors are regulated at the surface of neutrophils and mononuclear cells during different stages of LPS induced sepsis (*9*). In many cases common urinary tract infections, if left unchecked, can easily result in occurrence of sepsis. It is a prominent problem amongst intensive care unit patients. Recent estimates put septic

shock related deaths in the United States at nearly one hundred thousand patients per year. More alarmingly, these numbers have steadily increased since the 1930s, when the first reported cases of septic shock were documented (*10*). For such reasons, LPS has always sparked a wide interest amongst immunologists and pathologists. Recently, it has been reported that LPS shed from *E. coli* can alter pharmacokinetic and pharmacodynamic properties of several drugs targeting hepatic and urinary tract diseases (*11*). In order to evade immune recognition and allow bacterial survival, evolutionary pressures have forced bacteria to develop mechanisms of lipid A modification.

1.1.3.3: Lipid A Synthesis:

The biosynthesis of lipid A is a cytoplasmic event and is best characterized in *E. coli (12)* (figure 1.3). Termed the Raetz pathway, the process begins with the key precursor molecule UDP-*N*-acetylglucosamine (UDP-GlcNAc). The first enzyme is a cytosolic acyltransferase (LpxA) that selectively transfers *R*-3-hydroxymyristate from acyl carrier protein (ACP) to the 3-OH of UDP-GlcNAc (*13*). Recently, inhibitor-bound and substrate-bound structures of LpxA have been resolved and provide a detailed picture of lipid interaction with the enzyme (*14, 15*). The enzyme exists as a homotrimer and forms three substrate-binding sites. The presence of conserved residues in substrate binding pockets allows for high selectivity of a hydroxymyristate chain, and provides a first example of a hydrocarbon ruler in proteins. *E. coli* LpxA is extraordinarily selective for hydroxymyristoyl-ACP (3-OH-14:0-ACP) as the acyl donor substrate, while the

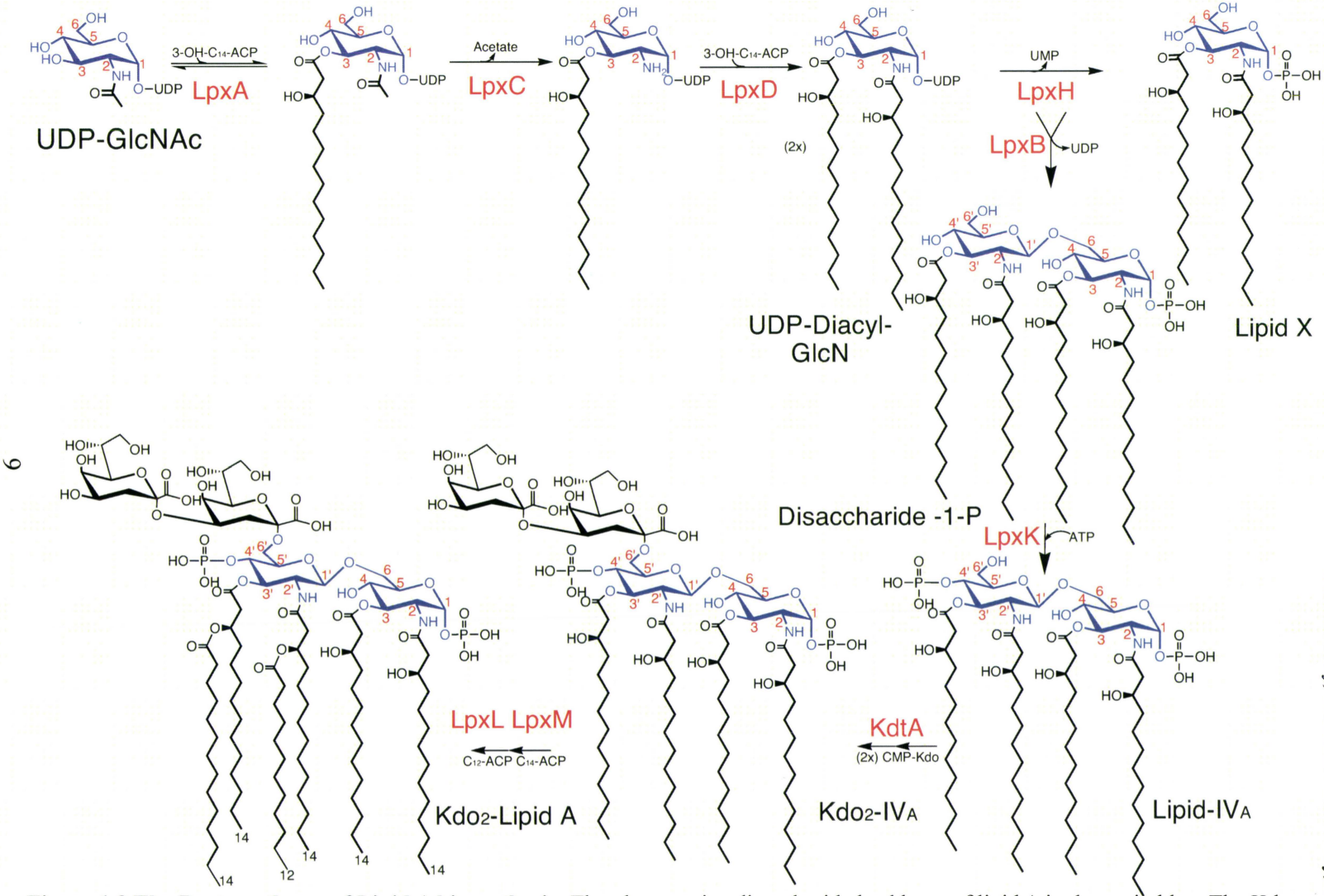

Figure 1.3 The Raetz pathway of Lipid A biosynthesis: The glucosamine disaccharide backbone of lipid A is shown in blue. The Kdo disaccharide is shown in black. The enzymes involved in each step are shown in red. The figure is adapted from Raetz *et al.*, 2007.

Pseudomonas aeruginosa LpxA prefers hydroxydecanoyl-ACP (3-OH-10:0-ACP). However, the specificity could be modulated through mutation of certain residues lining the active site cleft. For example, the specificity for Gly173Met mutant of *E. coli* LpxA was shifted from 3-OH-14:0-ACP to 3-OH-10:0-ACP. In contrast, the specificity of *P. aeruginosa* LpxA could be extended to accommodate 3-OH-14:0-ACP in the corresponding Met169Gly mutant.

The acylation of UDP-GlcNAc is reversible and unfavourable (*16*). Therefore, the first committed step in lipid A biogenesis is the subsequent deacetylation by LpxC, a Zn^{+2} dependent enzyme (*17*). Sequence alignments of LpxC have indicated no homology with deacetylases, making the enzyme an important target for drug design (*18*). The first crystal structure of LpxC was published in 2003 at a resolution of 2 Å (*19*). The study identified the location of the enzyme active site, as well as a hydrophobic tunnel lined with many conserved residues (*19*). During catalysis the tunnel is occupied by a myristate chain, which coordinates the substrate with the enzyme active site. This work has allowed for testing of several inhibitors of LpxC that may be of pharmacological significance. Recently, an inhibitor complexed crystal structure of LpxC with compound BB 78485 (*20*) and a NMR structure with compound TU-514 (*21*) have been reported.

Following deacetylation, a second *N*-linked *R*-3 hydroxymyristate moiety derived from ACP is incorporated by LpxD to generate UDP-2,3-diacylglucosamine (*22*). Crystal structure of LpxD has recently been resolved (*23*), and it displays a homotrimeric enzyme

that shares many similarities with LpxA, specifically in its lipid binding domain regions. A highly selective pyrophosphatase LpxH then cleaves UDP-2,3-diacylglucosamine at its pyrophosphate bond to form lipid X and UMP (*24*). Next, a disaccharide synthase, LpxB, transfers the 2,3-diacylglucosamine portion of another UDP-2,3-diacylglucosamine molecule to position 6 of lipid X, generating the β 1'-6-linkage found in all lipid A molecules (*25*). LpxK, a membrane-bound kinase phosphorylates the 4' end of the disaccharide backbone generating lipid IV_A (*26*). This permits two 3-deoxy-D-*manno*-2-octulosonic acid (Kdo) sugars to be incorporated by a bifunctional Kdo-transferase encoded by the *kdtA* (*waaA*) gene (*27*). The labile nucleotide sugar CMP-Kdo serves as the Kdo donor. Attachment of the first sugar to lipid IV_A precedes the addition of the second Kdo sugar. In *Haemophilus*, a homolog of KdtA may add a phosphate group upon addition of the first Kdo sugar (*28*). The last steps of *E. coli* lipid A biosynthesis involve the transfer of lauroyl and myristoyl groups from ACP to the distal glucosamine unit, producing acyloxyacyl moieties generated by LpxL (HtrB) and LpxM (MsbB), respectively (*29*). Note that all the enzymes involved in the generation of Kdo_2-lipid A have a requirement for cytosolic substrates (figure 1.3).

Several variations of lipid A structures are scattered across the Gram-negative bacteria and many are known in detail. In *Francisella tularensis*, the causative agent of tularaemia (*30*), the lipid A moiety lacks Kdo sugars and therefore, has no core sugars or O antigen. Furthermore, a phosphatase, LpxE and a deacylase, LpxF, generates a lipid A lacking the 4' phosphate and a 3' myristate chain (*31, 32*).

1.1.3.4: Core sugar attachment and LPS transport:

Core oligosaccharides and O antigens are also synthesized in the cytoplasm. The core sugars are attached to the Kdo_2-lipid A moiety at the inner leaflet of the IM (*1*). After assembly, the core-Kdo_2-lipid A is flipped to the periplasmic side of the IM via an ABC transporter MsbA (*33*). The O-antigen is assembled on undecaprenylpyrophosphate (Und-PPi), which is flipped to the periplasmic side and then ligated to the core-Kdo_2-lipid A, releasing Und-PPi (*34, 35*). The completed LPS molecule is transported from the outer leaflet of the IM to the bacterial cell surface by the Lpt system (*36*). A complex localized to the inner membrane consists of LptB, LptC, LptF and LptG (YbrK) that form an ABC-transporter, and are proposed to be responsible for releasing the LPS molecule from the IM and delivering it to the periplasmic chaperone LptA (YhbN) (*37-40*). Once LPS has reached the inner leaflet of the OM, it is flipped to the outer leaflet by an OM protein LptD (Imp) (*41-43*), which is in complex with the lipoprotein LptE (RlpB) (*44*). Figure 1.4 outlines the major stages of LPS secretion.

1.4: Cationic antimicrobial peptides (CAMPs):

LPS contains phosphate groups and acidic sugars that confer a net negative charge on the molecule. In order to reduce the electrostatic repulsion between neighbouring LPS molecules at the cellular surface, the bacterial OM sequesters divalent cations consisting mainly of Mg^{+2} (*45*). The electrostatic attraction of the LPS to cationic molecules is what renders it susceptible to CAMP. CAMPs are amphipathic molecules (*46*), which are

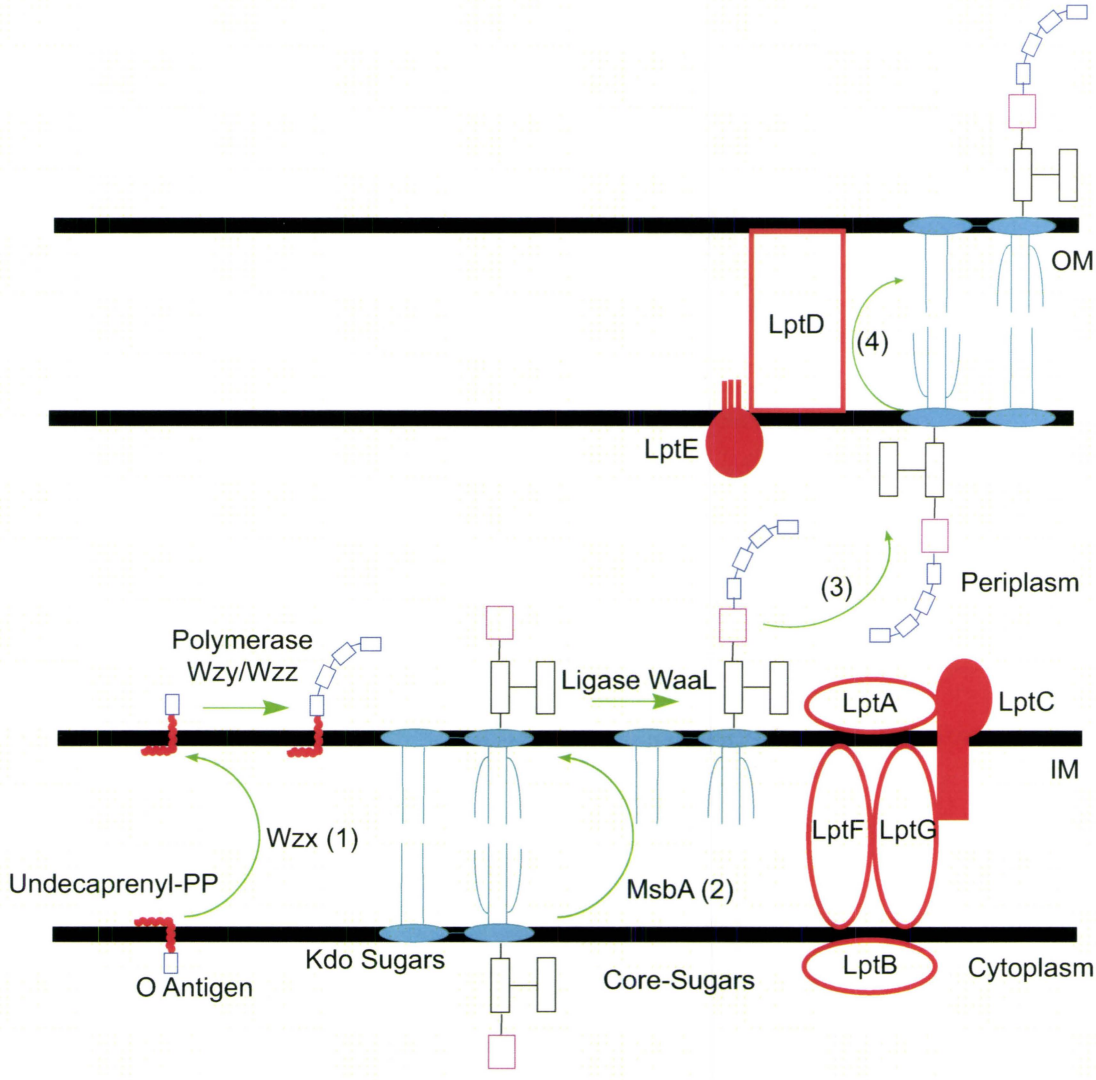

Figure 1.4: Schematic representation of the different stages of LPS secretion:

initially unstructured in the aqueous medium. Their initial electrostatic interaction with the bacterial surface serves to displace some Mg^{+2} ions. The reduced dielectric constant at the membrane interface induces dehydration of the peptide bonds, which become hydrogen-bonded in α or β-motifs. This reveals the amphipathic features that facilitate the translocation of CAMPs through the hydrocarbon layer by a non-porin pathway, termed the "self promoted uptake pathway" (*47*). Once inside the periplasm CAMPs are then believed to target the IM bilayer and to produce a detergent-like disruption of permeability. The consequences of IM permeation include the fatal depolarization of the transmembrane potential across the IM (*48*), leakage of cytoplasmic contents, cell lysis and cell death.

1.5 Lipid A modifications:

External stimuli such as Mg^{+2}-limitation and presence of CAMPs regulates the covalent modifications of lipid A. Figure 1.5 outlines such modifications in *E. coli* and *S. enterica (28, 49)*. Regulated modifications of lipid A can be divided into two major classes: reduction of lipid A negative charge and changes in acylation patterns.

1.5.1 Reduction of Lipid A Negative Charge:

CAMPs are able to gain access through bacterial perimeter defences by establishing electrostatic interactions with the OM. In order to minimize this attractive force, bacteria actively reduce the negative charge of the OM by adding phosphoethanolamine (pEtN) and L-4-aminoarabinose (L-Ara4-N) to lipid A. Under

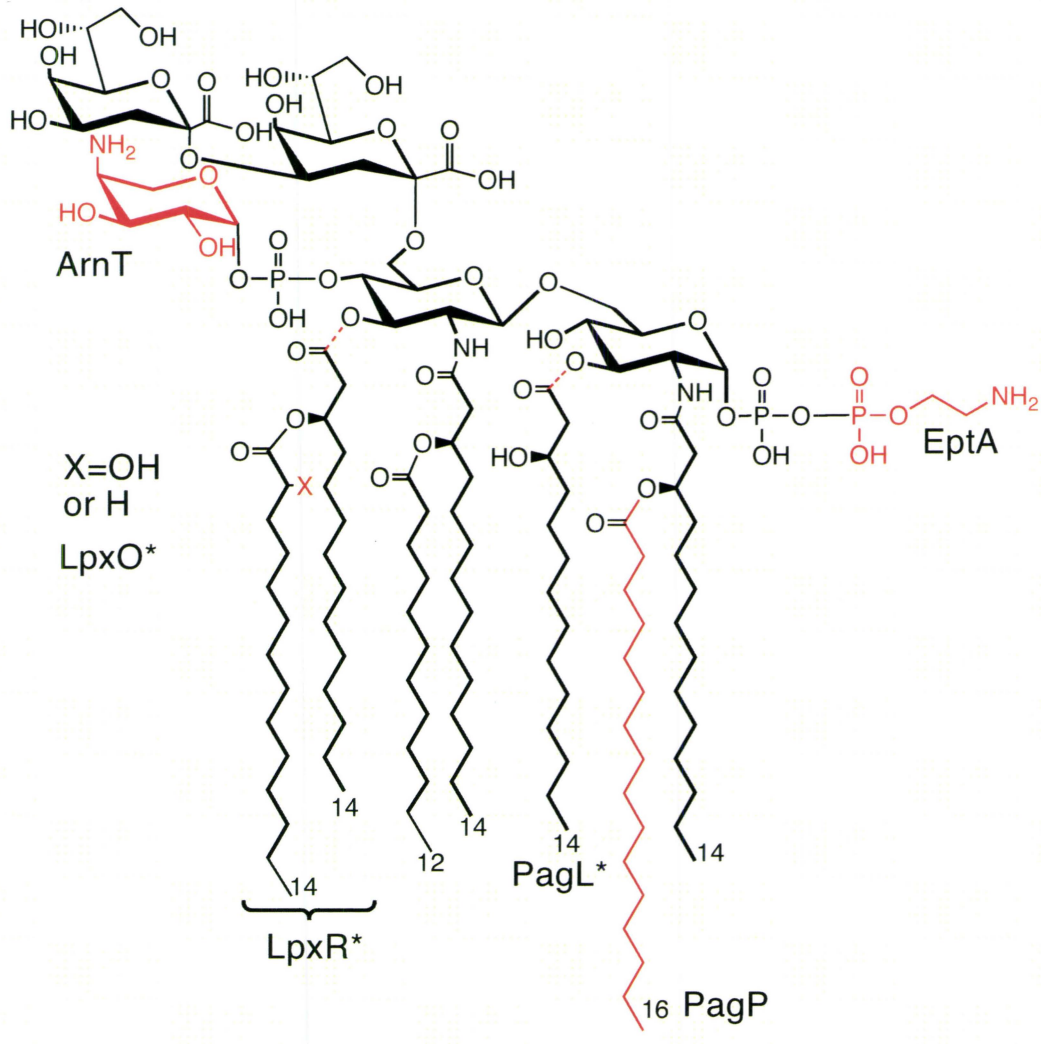

Figure 1.5 Lipid A modifications: PagP is responsible for adding a palmitate to lipid A, while PagL acts as a deacylase. LpxR is also a deacylase that cleaves 3'-acyloxyacyl moeity. LpxO, a hydroxylase, generates S-2-OH at the 3'-position. pEtN and L-Ara4-N groups are added to the 1 and 4' positions of lipid A, respectively. Enzymes with a star beside their names are not found in *E. coli*. [Adapted from Raetz *et al.*, 2007].

PhoP/PhoQ control, EptA adds pEtN to the phosphate at position 1 (*50*), while ArnT modifies the 4' phosphate with L-Ara4-N. The modification by EptA occurs at the outer surface of the inner-membrane (*28*). The cellular pool of phosphatidylethanolamine serves as the substrate source for the pEtN modification. Under certain conditions, where L-Ara4-N is not available, EptA may add a second pEtN moiety to the phosphate at 4' position (*51*). In *E. coli* K12, mild acidic conditions are required to trigger the addition of pEtN to the lipid A (*51*), while in case of *E. coli* O157:H7, such modifications are constitutively expressed (*52*).

The pathway for modification of lipid A by L-Ara4-N has recently been elucidated. The precursor molecule for the pathway is UDP-glucose. The oxidized species, UDP-glucuronic acid is acted upon by the dual purpose ArnA (PmrL) enzyme, to generate a UDP-4-Ketopentose (*53*). This serves as a substrate for the transaminase ArnB (PmrH), which forms UDP-β-L-Ara4N (*54*). ArnA further drives the pathway by formylating the amino group of UDP-β-L-Ara4N generating UDP-β-L-Ara4FN (*53*). ArnC (PmrF) transfers this molecule to the undecaprenyl phosphate moiety (*55*). ArnD (PmrJ) deformylates the Undecaprenyl Phosphate-α-L-Ara4FN molecule, to produce Undecaprenyl Phosphate-α-L-Ara4N (*55*). This is flipped by recently discovered proteins ArnE/ArnF (PmrL/PmrM) to the periplasmic side of the IM (*56*). Finally, ArnT (PmrK) catalyzes the addition of L-Ara4N to lipid A (*50*).

These substituents serve to reduce the overall negative charge of lipid A and inhibit the electrostatic interaction between the LPS and CAMPs (*57*). This neutralization of negatively charged bacterial surface is associated with resistance to polymyxin B, a lipid A-binding cationic cyclic peptide antibiotic, in *E. coli* and *S. enterica* (*58, 59*).

1.5.2 Changes in Acylation Patterns:

Under low Mg^{+2} conditions, PagP, a transacylase, incorporates a palmitate chain at the 2-position (*60*). PagL, a deacylase removes the acyl chain at the 3-position (*50*). These two enzymes are under the control of a 2-component regulatory system termed PhoP/PhoQ. A non-PhoP/PhoQ regulated protein, LpxO belongs to the Fe^{+2}/α-ketoglutarate/O_2 superfamily and catalyses the formation of the *S*-2-hydroxymyristate moiety. This presumably results in an increase in hydrogen bonding between lipid A moieties that may decrease penetration by organic molecules that could harm bacteria (*61, 62*). The active site of the enzyme faces the cytoplasm (*33*). Similar to PagL, a second OM deacylase, LpxR, that is not under PhoP/PhoQ control, catalyses the cleavage of 3'-acyloxyacyl moiety of lipid A. Such modifications of lipid A acylation may block the subsequent hydrophobic interaction between CAMPs and the membrane bilayer. Lipid A acylation is also critical in modulating its endotoxic activity through interaction with the TLR4 signal transduction pathway (*63, 64*). Recently, changes of lipid A acylation have been shown to attenuate lipid A mediated TLR4 activation (*65*).

1.5.3 Regulation of lipid A modifications:

Gram-negative bacteria have adapted the PhoP/PhoQ two-component signal transduction pathway to respond to environmental stimuli including divalent cation limitation and the presence of CAMPs that may be encountered during infections (*66*) (*67*). PhoQ was identified as the membrane bound sensor histidine kinase that is maintained in a repressed state in the presence of high concentrations of divalent cations (figure 1.6) (*67-69*). These cations associate with a patch of acidic residues located in the sensor domain of PhoQ. CAMPs will compete with divalent cations for these binding sites and trigger a signalling cascade resulting in phosphorylation of PhoP. Phosphorylated PhoP controls the expression of many genes that are involved in Mg^{+2} transport and in LPS modification. For example, transcription of *pagP* and *pagL*, which are involved in the modification of lipid A acyl chains, are under the direct influence of PhoP/PhoQ (*50, 61, 70*). The PmrA/PmrB, two-component regulatory system is a downstream regulator of the PhoP/PhoQ system. It is required for the modification of lipid A with pEtN and L-Ara4-N (*71*). PmrA/PmrB can be PhoP/PhoQ-activated via a mediating protein PmrD (*72*). A non-functional homolog of PmrD has been identified in *E. coli (73)*. In *Salmonella typhimurium*, the PmrD protein has high affinity for phosporylated PmrA and, upon binding PmrA it inhibits subsequent dephosphorylation steps (*74*). PmrA-induced genes can also be activated independently of PhoP/PhoQ by exposure to Fe^{+3} or mild acidic conditions (*75*). Similar to PhoQ, extracellular triggers cause autophosphorylation of the sensory transducer PmrB, which then transfers the

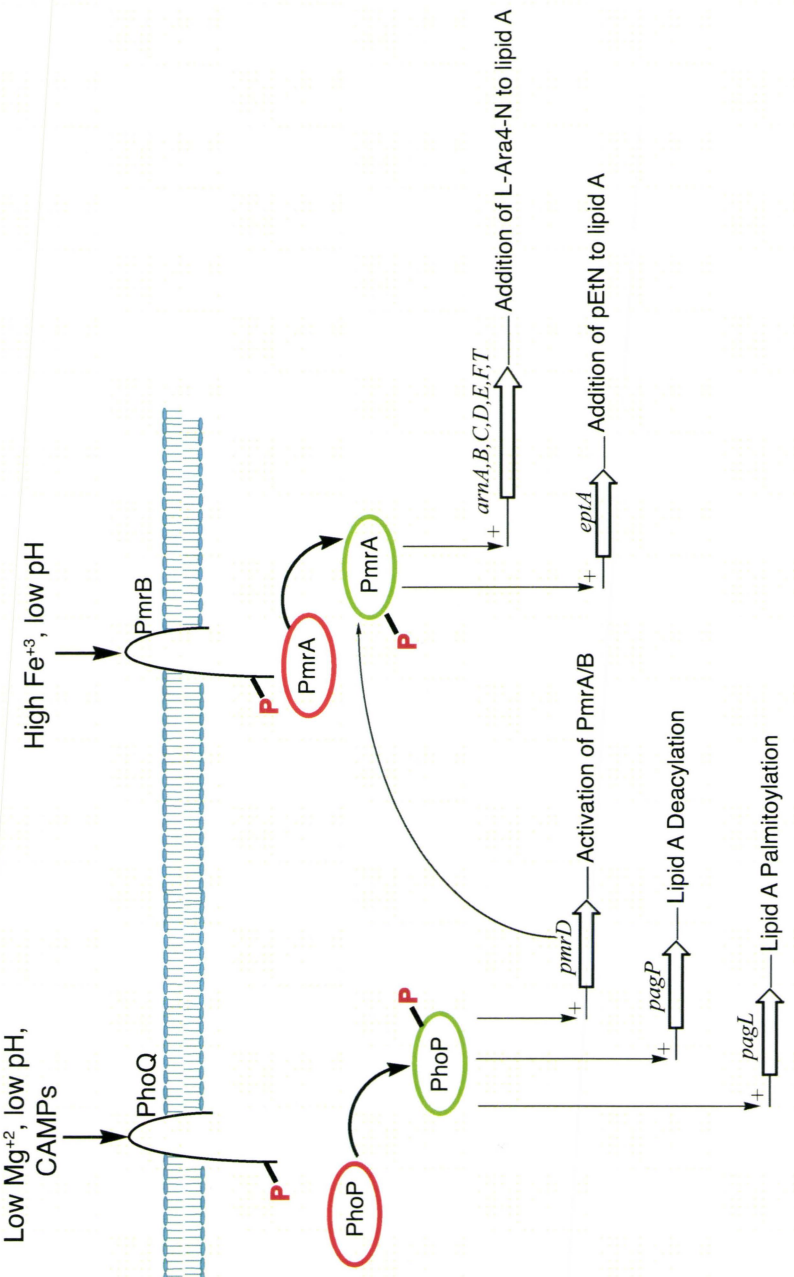

Figure 1.6: Genetic regulation of lipid A modifications in Salmonella typhimurium

phosphate group to PmrA. The activated PmrA can then initiate transcription of several downstream genes. PmrA/PmrB activation has also been shown to suppress PmrD expression (*76*), thereby creating a negative feedback loop.

1.6: Membrane Proteins:

Membrane proteins are critical to bacterial survival. There are several classes of membrane proteins based on the wide range of functions they perform. The associated biochemistry of membrane proteins localized in the IM, in comparison to the OM, shows vast differences. Structurally, IM proteins are anchored by transmembrane α helical segments, while OM proteins (OMPs) are β barrels with even numbers of β strands (*77*). Our discussion will focus on OMPs.

1.6.1: Structure:

The β-barrel structure of OMPs have short turns facing the periplasm and larger loops protruding beyond the OM. β-barrels satisfy the hydrogen-bonding requirements of the polypeptide backbone, which must be dehydrated in the membrane environment. The exterior of the barrel that is exposed to the lipids is generally composed of hydrophobic residues, while residues exposed to the barrel interior are mostly hydrophilic (*78*). The latter point is especially true in the case of porins and specific channels, but is also observed in smaller β-barrel proteins that function as enzymes such as OmpT and OMPLA. Therefore, the interiors of β-barrel proteins are generally shielded from membrane lipids by hydrogen-bonding interactions between β-strands in the

transmembrane domains. One other characteristic feature of these OM proteins is their "heat modifiability"; that is, when the OM proteins are heated in SDS solution, their electrophoretic mobilities are altered in SDS-PAGE. This is thought to result from an intrinsic resistance to denaturation by detergents in the absence of heat treatment (*79*).

1.6.2: OMP Transport:

OMP transport to their final destination posses a significant problem. The machinery for protein production is present in the cytoplasm. These proteins need to transverse the IM and the periplasmic space before they reach the OM, where they have to be correctly folded and inserted. OMPs are synthesized with an *N*-terminal signal sequence, which directs the protein to the SecYEG translocon (*80, 81*). The signal peptide is cleaved by a peptidase at the periplasmic face of the membrane (*82*), where the protein can be captured by chaperone proteins. Several putative periplasmic chaperones involved in protein transport have been suggested including Skp and SurA (*83, 84*). Deletion mutations of Skp have been shown to lead to reduction in properly folded OMPs (*85*). Double deletion mutations of Skp and DegP are not viable, since in the absence of the DegP periplasmic protease, the accumulation of misfolded OMPs is lethal to the bacteria (*86*). Similarly, SurA was shown to directly interact with immature forms of PhoE, LamB and OmpF, through recognition of a aromatic-random-aromatic (Ar-X-Ar) motifs (*87*). Some chaperones like SurA and FkpA serve not only as molecular chaperones but also as folding catalysts (*88*). Chaperone and folding activities are

typically independent of each other. Once the protein is delivered to the OM, the BAM complex help to fold the protein in its correct conformation (*89*). The complex consists of BamA (YaeT), BamB (YfgL), BamC (NlpB), BamD (YfiO) and BamE (SmpA). BamB, BamC, BamD and BamE are lipoproteins, while BamA is an essential β barrel protein (*89, 90*). It should be noted that studies have confirmed that LPS is required for proper protein assembly of certain OMPs in the OM (*91*).

1.7: Examples of OM enzymes:

1.7.1: Outer-Membrane Phospholipase A (OMPLA):

Phospholipases belong to a family of lipolytic enzymes that catalyse the hydrolysis of phospholipids. OMPLA was identified in 1971 as an integral membrane protein in *E. coli* (*92*). The *pldA* gene encoded a 32 kDa protein that exhibits both phospholipase A1 and A2 activity (*93*). Several triggers have been reported for OMPLA activation including phage induced lysis (*94*), temperature shift (*95*), EDTA treatment (*96*) and colicin release (*97*). All of these triggers have in common that they likely perturb the integrity of the OM.

The crystal structure of OMPLA was described a decade ago (*98*). OMPLA is a twelve-stranded anti-parallel β barrel composed of a convex and a flat side. Monomeric OMPLA is not active and only upon dimerization does the protein exhibit any enzymatic activity (figure 1.7). The dimer is stabilized via hydrogen bond formation along the flat barrel surface. The crystal structure revealed two calcium-binding sites for the OMPLA

dimer. The monomeric calcium binding site, located close to the active site of the protein is a low affinity binding site and has a K_d of 358μM. Upon dimerization the calcium binding site adopts a high affinity conformation and has a K_d of 36μM (*93*). Since under conditions of membrane perturbation divalent cations are stripped from the OM, a high affinity Ca^{+2} binding site in the active state might ensure catalytic activity of the protein. The active site is composed of His142, Ser142 and Asn156 whose role in catalysis has been confirmed through mutagenesis. These residues are organized in a catalytic triad, a characteristic of classical serine hydrolases. The bound Ca^{+2} plays a significant role in catalysis by orientating water molecules in the active site to form the oxyanion hole (*99*). His142 abstracts a proton from Ser142, allowing it to perform a nucleophilic attack on the phospholipid ester linkage, leading to formation and hydrolysis of an acyl-enzyme intermediate.

The phospholipid substrate and OMPLA are physically separated in the asymmetric OM. Membrane perturbation would cause phospholipids from the inner leaflet to migrate to the outer leaflet. Therefore signals that activate OMPLA also eliminate the physical separation of substrate and enzyme. As a result of dimerization, substrate binding pockets are formed that run down the two active sites along the dimer interface. It was initially thought that the architecture of the active site allows for a wide array of substrates to be accommodated in the binding site. Current research shows that while OMPLA has no preference for the phospholipid head group it exhibits a strong selectivity for acyl chains of 14 and 16 methylene units (*100*).

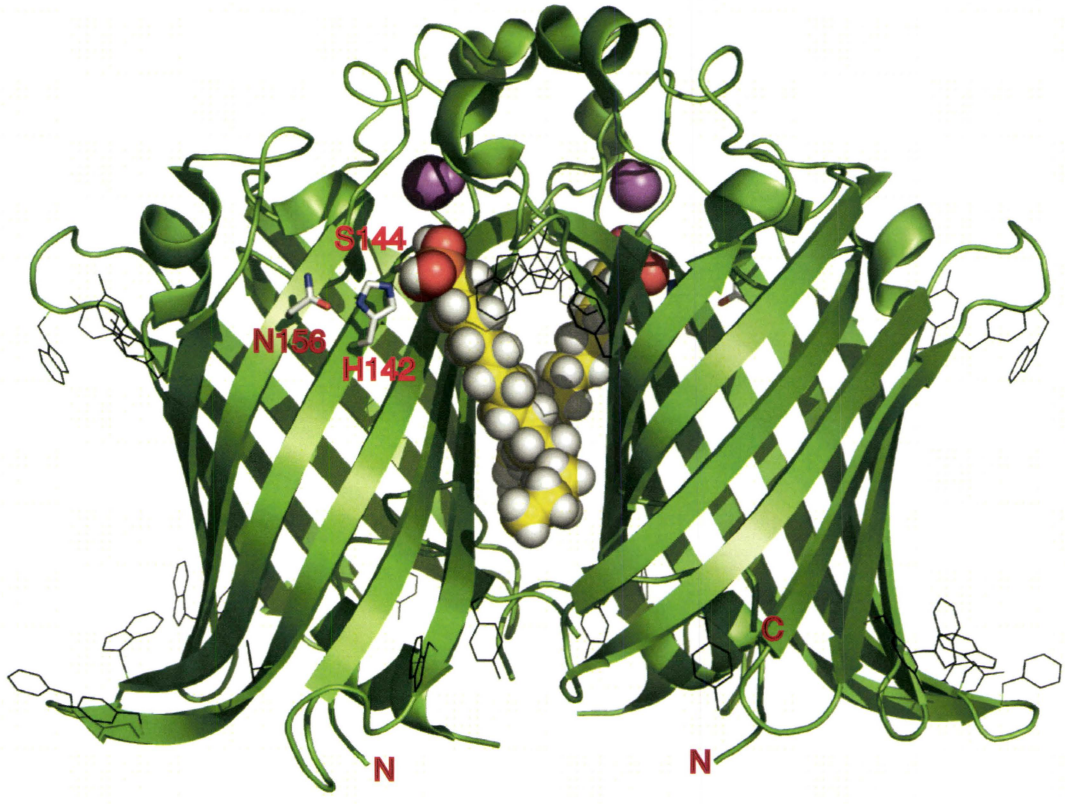

Figure 1.7: Crystal structure of OMPLA [pdb: 1QD6]. The aromatic belt residues are depicted in black sticks. The enzyme is bound to a cofactor Ca^{+2} (purple) and hexadecanesulfonyl group. Residues implicated in catalysis, as well as the N and C termini are labeled in red.

1.7.2 Outer-Membrane Protein T (OmpT):

OmpT is an OM protease that cleaves between two basic amino acids (*101*) and requires LPS for activity (*102*). The protein belongs to a highly homologous family of OM proteases known as omptins, and is implicated in virulence. The crystal structure of OmpT has been resolved to 2.6 Å (*103*). The protein is a 10-stranded anti-parallel β barrel that exists as a monomer in the OM and adopts a vase shaped structure (figure 1.8). The active site of the protease is located in a negatively charged groove at the exterior face of the protein and is exposed to the extracellular medium. The existence of a negatively charged active site was anticipated because the protease recognizes two consecutive basic residues as the target sequence. Mutagenesis studies have associated a ten thousand fold decrease in enzyme activity with Asp83, Asp85, Asp210 and His212 mutants (*102*). Within the catalytic site, four putative binding sites were identified with different specificities for substrate residues. Amongst these, position P1 is highly specific for Lys and Arg, while position P2' exhibits specificity for small hydrophobic residues like Ile, Val and Ala. Positions P1' and P2 display a broad specificity for residues (*101*). Associated with these highly specific sites are grooves S1 and S2' that accommodate the specific residues. The S1 groove is lined with Gly27 and Asp208 and therefore can accommodate Lys and Arg at position P1. Similarly, Met81 and Ile170 line the bottom of the S2' groove, explaining the preference of small hydrophobic residues at position P2'. Interaction of the residues bound at these sites is significant, since it would position the scissile peptide bond between Asp83-85 and Asp210-H212, which are

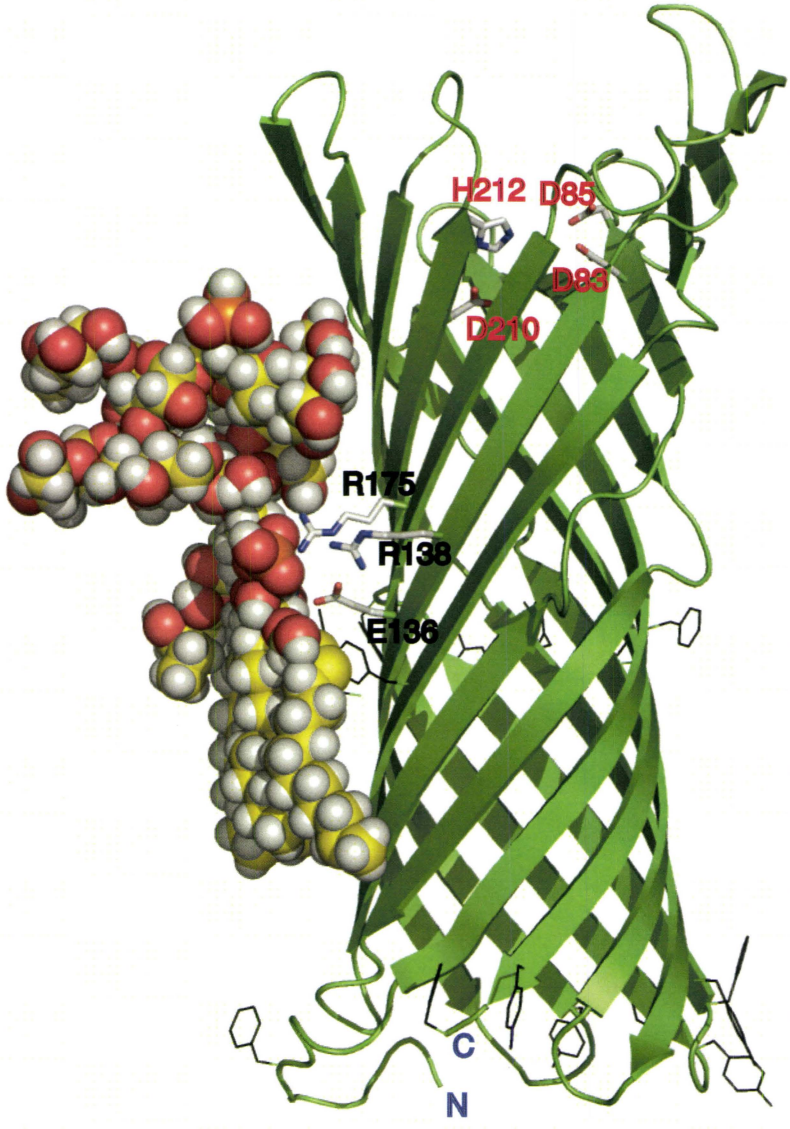

Figure 1.8 Crystal structure of OmpT (pdb: 1I78): The coordinates for LPS bound to OmpT are provided by Lucy Rutten and Piet Gros at the University of Utrecht (unpublished data). The aromatic belt residues are depicted in black sticks. Catalytically significant residues are labelled in red, while the *N* and the *C* termini of the enzyme are labelled in blue. Residues in black are proposed to be significant in LPS binding.

located on opposite sides of the active site. The proposed mechanism of OmpT suggests that His212-Asp210 pair resembles His-Asp pairs seen in catalytic triads. Although in this case, an identifiable nucleophile is not present in close proximity to the His. It is hypothesized based on the structure of the active site that a water molecule present in between the Asp83 and His212 is activated by the His212-Asp210 pair and it can hence perform a nucleophilic attack on the scissile bond (*103*).

1.7.3 LpxR:

LpxR found in *S. enterica,* but absent in *E. coli,* is an OM lipase. The enzyme catalyzes the cleavage of the 3'-acyloxyacyl moiety of lipid A. Recently, a high resolution crystal structure of the enzyme has been published (figure 1.9) (*104*). The 12-stranded β barrel enzyme requires Ca^{+2} as a cofactor, but was crystallized in the presence of Zn^{+2} ions. Although Zn^{+2} does not support enzyme activity, Ca^{+2} is expected to occupy the same position. Modelling of Kdo_2-lipid A, the substrate of LpxR, within the proposed active site of the protein has led the authors to suggest a catalytic mechanism. His121 acts as a base and is stabilized by the carbonyl oxygen of Asn128. The Ca^{+2} ion forms the required oxyanion hole, and is coordinated by Asp10 and Thr34.

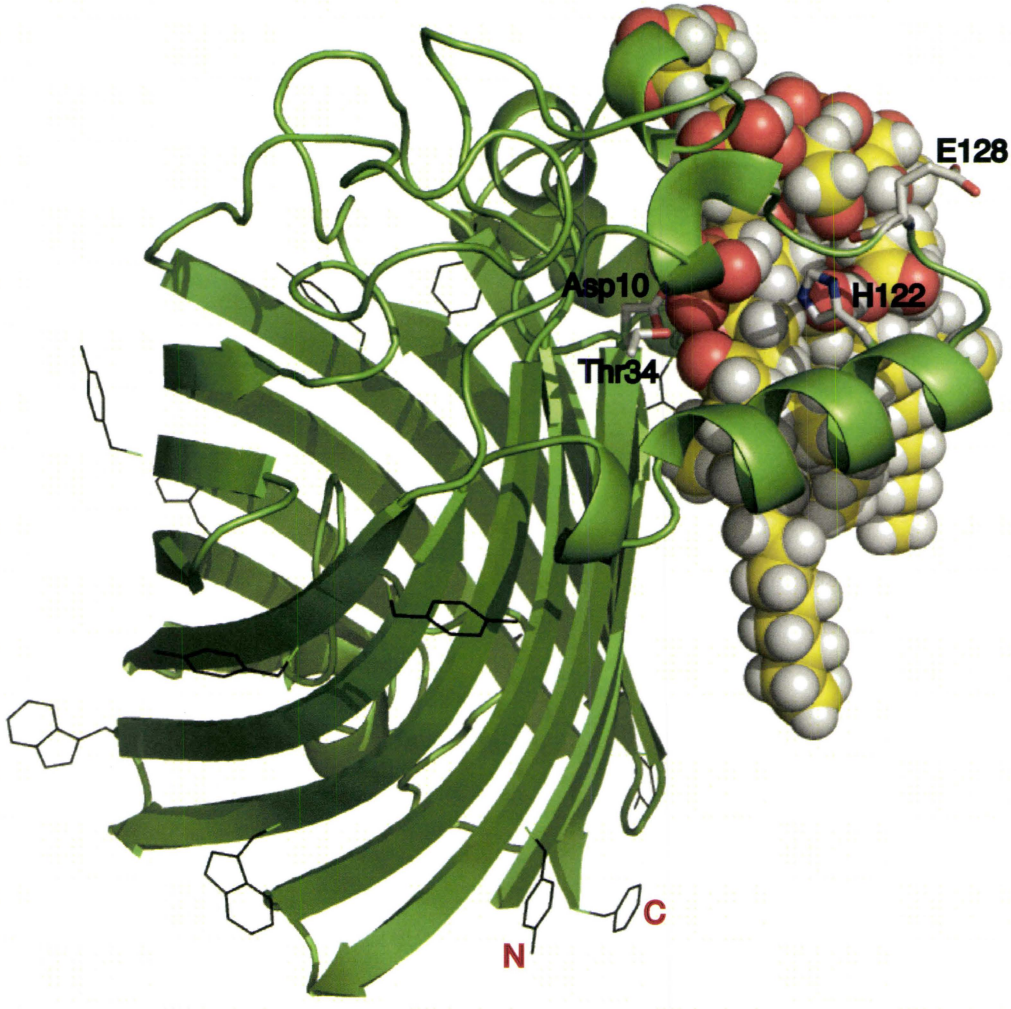

Figure 1.9 Crystal structure of LpxR (pdb: 3FID): The coordinates for Lipid A bound to LpxR are provided by Lucy Rutten and Piet Gros at the University of Utrecht (Unpublished data). The aromatic belt residues are depicted in black sticks. Residues implicated in catalysis are shown in black, while the *N* and the *C* termini of the protein are labelled in red.

1.7.4 PagL:

PagL is a 3-*O*-deacylase that is found in many Gram-negatives and is localized to the OM. The *pagL* gene is under the control of the PhoP/PhoQ regulatory system. PagL deacylates lipid A at position 3 and may decrease its endotoxic properties. PagL has recently been reported to be responsible for resistance to polymyxin B in L-Ara4-N and pEtN mutants of *Salmonella enterica*. PagL is maintained in a latent state in wild-type *S. enterica*, since no lipid A deacylation can be observed *in vivo*. It has been recently shown that the extracellular loops of PagL can detect the presence of L-Ara4-N and maintain inactivity (*105*). In L-Ara4-N and pEtN mutants, PagL is no longer latent and the deacylation of lipid A can be observed (*65, 106*). Furthermore, it has been reported that vaccines against whooping cough generated from *B. pertussis* expressing PagL exhibit greater efficacy without altering vaccine reactogenicity (*107*). Recent research has uncovered the importance of PagL in pathogenesis. In the case of *Pseudomonas aeruginosa*, clinical isolates from patients with Cystic Fibrosis (CF) had PagL constitutively expressed, while isolates from acute infections of blood and the urinary tract, as well as environmental isolates, had a regulated expression of PagL. The constitutively expressed PagL is required for the formation of specific lipid A species that are seen in CF patients (*108*). Therefore, the lipid A deacylation may allow the bacteria to be in a stealth mode and remain undetected by the host innate immune system.

The crystal structure of PagL has been elucidated in *P. aeruginosa*. It is an 8-stranded anti-parallel β barrel with short periplasmic turns and large extracellular loops (figure 1.10). The protein's position in the OM is at a 30° tilt with respect to the membrane normal. Two identifiable regions of aromatic residues are present along the hydrophilic/hydrophobic interface of the protein (*109*). The catalytic site of the protein is located at the extracellular face of the protein. Mutagenesis identified Ser128 and His126 as critical for catalytic activity. The structure of the active site allows the protein to have a catalytic triad. PagL Ser128 is the nucelophile that attacks the scissile bond of the substrate and His126 is a general base that abstracts a proton from the serine residue making it available to perform the nucleophilic attack. In order for His126 to abstract a proton from Ser128, the Nε2 of His needs to be deprotanated. This Nε2 deprotonation of His is stabilized via hydrogen binding of Nδ1 proton to Gln140. The oxyanion hole is formed by Ala130 and Gly131.

In an intriguing experiment, PagL from *P. aeruginosa* was able to deacylate lipid A from *E. coli*. Since *P. aeruginosa* lipid A carries 3-OH C10 rather than 3-OH C14 found in *E. coli*, PagL was shown to lack fatty acid chain length specificity (*65*). The deacylase does exhibit specificity for substrates that have the *R*-3-hydroxyl group in the correct position. It is postulated that Asp106 might be the residue responsible for the specificity that PagL exhibits, although this yet remains to be established.

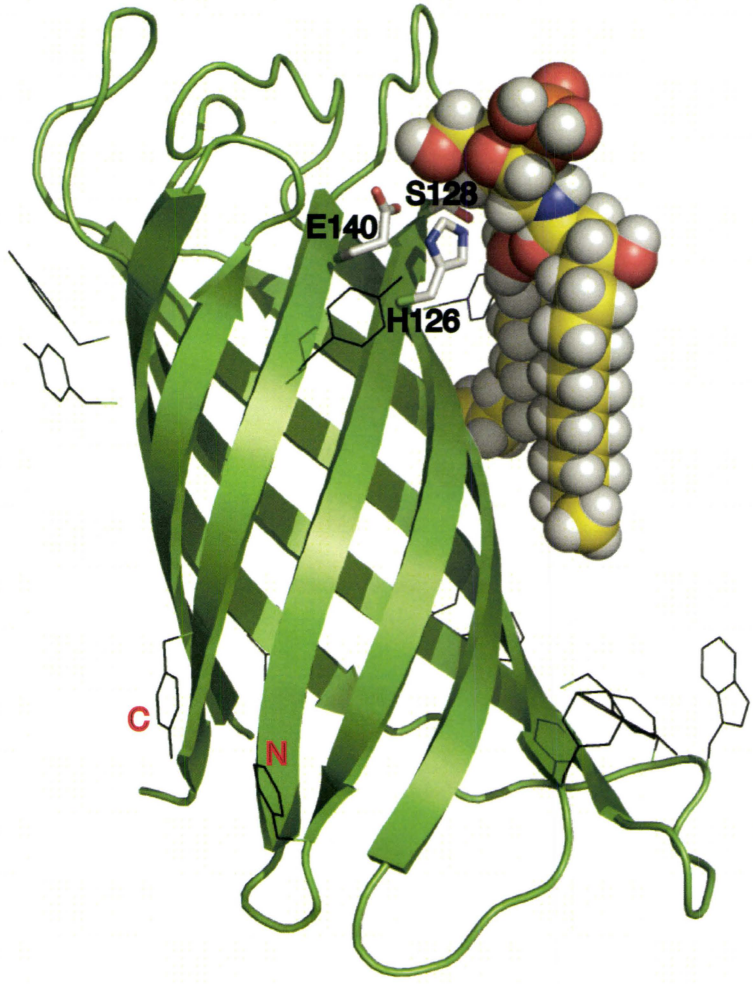

Figure 1.10 Crystal structure of PagL (pdb 2ERV): The coordinates for Lipid X bound to PagL are provided by Lucy Rutten and Piet Gros at the University of Utrecht (unpublished data). The aromatic belt residues are depicted in black sticks. The residues implicated in catalysis are shown in black. The *N* and the *C* termini of the protein are labelled in red.

1.7.5 PagP:

The molecular structure and dynamics of PagP have been elucidated by both nuclear magnetic resonance (NMR) spectroscopy (*110*), and by x-ray crystallography (*111*). PagP, like other OM proteins, has a β-barrel structure, but due to its small size (19kDa), PagP was expected to adopt the smallest possible β-barrel topology in membranes of 8 strands (*112*). Through NMR spectroscopy, PagP was first confirmed to be an 8-stranded antiparallel β-barrel preceded by an amino-terminal amphipathic α-helix. The crystal structure of PagP has been resolved to 1.9Å resolution (figure 1.11). The crystal structure confirmed the findings of NMR spectroscopy but it also indicated a 25° tilt of the barrel axis with respect to the membrane normal (*111*). A tilted position in PagP is supported by the alignment of aromatic belt residues (*78*) with the membrane interface (*111*). This orientation creates a defined hydrophobic zone bordered by hydrophilic regions, which is not seen in the untilted PagP, and was the first example of a tilted β-barrel membrane protein.

1.8 PagP:

1.8.1 PagP in pathogenic Gram-negative bacteria:

Figure 1.12 depicts the sequence alignment of PagP across its homologues. PagP has been identified in Gram-negative pathogens, including *Yersinia, Bordetella, Legionella, Erwinia* and *Photorhabdus* (*113-117*).

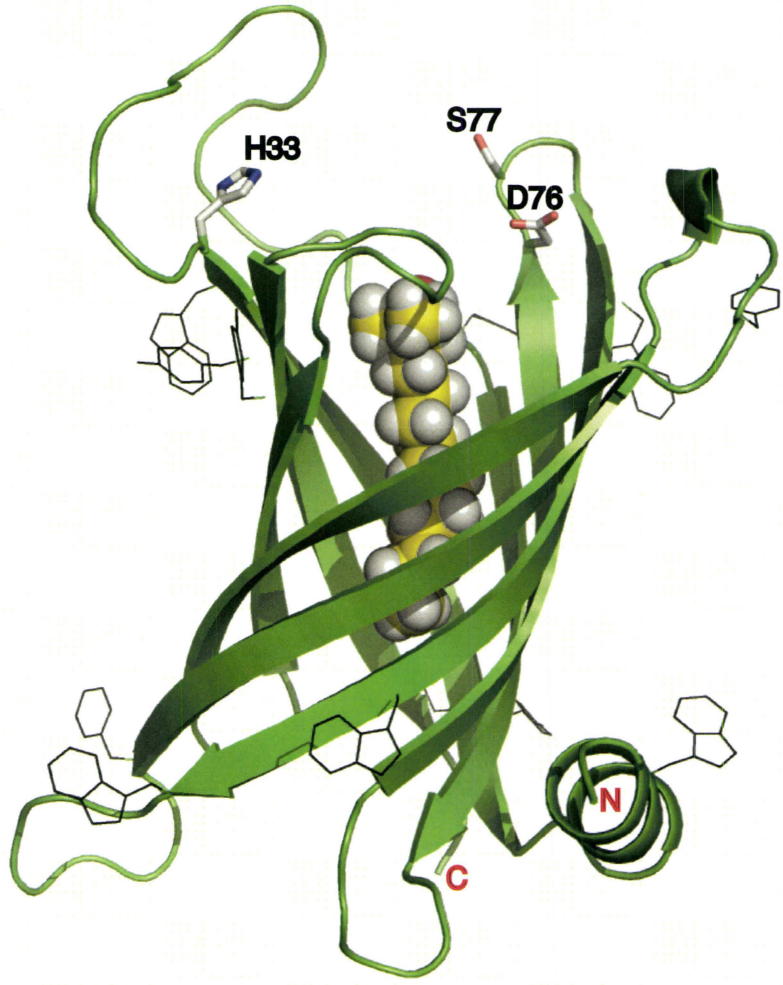

Figure 1.11 Crystal structure of PagP (pdb: 1THQ): The coordinates with L1 loop introduced between β-strands A and B are provided by Chris Neale and Régis Pomès at the University of Toronto. The aromatic belt residues are depicted in black sticks. A molecule of LDAO is bound in the hydrocarbon ruler. Catalytically significant residues are shown in black, while the *N* and the *C* termini of the enzyme are labelled in red.

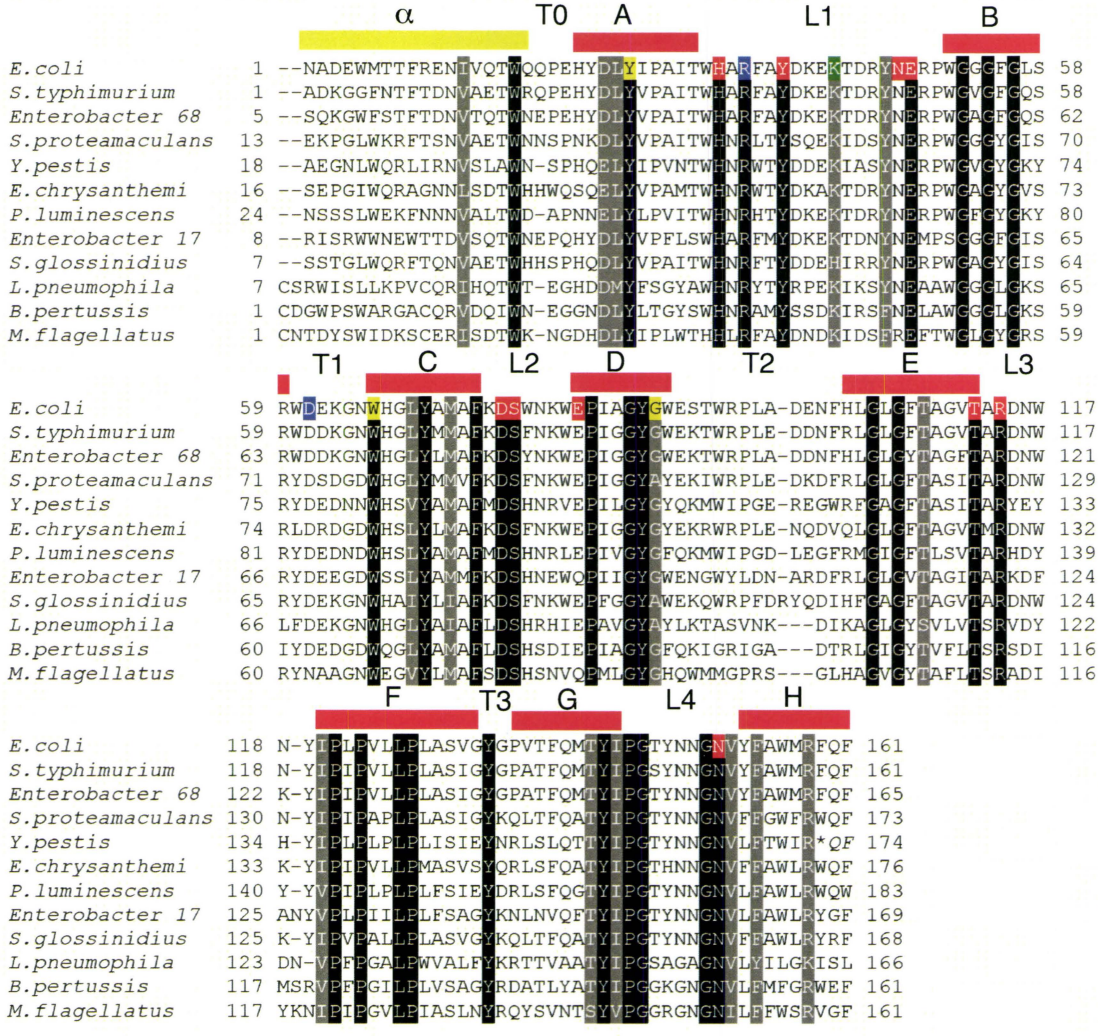

Figure 1.12 Multiple sequence alignment of PagP: The N-terminal signal peptides have been removed. The mature N-terminal 7 residues of *E. coli* PagP are disordered in the crystal structure, and several homologues have poorly conserved extensions of this region, so some of these extensions were removed as well. Black boxes identify identities and gray boxes similarities. Red boxes identify residues that we mutated to Ala and found the enzyme was expressed in membranes, but without catalytic activity. Residues in blue were mutated to Ala and retained significant residual activity, but were expressed poorly. Residues in yellow mark the floor of the hydrocarbon ruler. The alignment was constructed using Clustal W (Adapted from Bishop *et al.*, 2005).

1.8.1.1 PagP in *Legionella*:

A *pagP* homologue, *rcp* (resistance to cationic antimicrobial peptides) has been discovered in *Legionella pneumophila* (*117*). The *rcp* gene is required for the intracellular survival in amoeba and macrophage models of Legionnaire's disease. Also, *rcp* mutant bacteria displayed a defect in colonization of the mouse lung, indicating PagP as a requirement for intracellular infection and virulence.

1.8.1.2 PagP in *Bordetella*:

PagP incorporates a palmitate chain into the lipid A of *Bordetella bronchiseptica*, a respiratory pathogen of mammals (*114*). Recently it has been reported that *B. bronchiseptica pagP*, which is not required for initial colonization of the mouse respiratory tract, is required 7 days after colonization for persistence of the infection (*114*). These results were later confirmed by the observation that *B. bronchiseptica pagP* provides resistance to antibody-mediated complement cell lysis (*118*). As in the case of *E. coli*, the *pagP* promoter is under the control of a two-component virulence signal transduction pathway, termed BvgA/BvgS (*114*). Although a Bvg-regulated *pagP* homologue is reported to be present in *B. parapertussis*, it appears that the *pagP* promoter in *B. pertusis*, the causative agent of whooping cough in children, has been inactivated by the insertion of a transposable genetic element. Comparative genomics indicate that the acquisition of transposable genetic elements was important during the

evolution of *B. parapertussis* and *B. pertussis* from a *B. bronchiseptica*-like ancestor (*119*).

1.8.1.3 PagP in *Yersinia*:

The genus *Yersinia* includes the enteropathogenic species *Y. enterocolitica* and *Y. pseudotuberculosis*, in addition to the highly virulent causative agent of the bubonic and pneumonic plague, *Y. pestis*. Virulence mechanisms in pathogenic Yersiniae are under temperature-controlled promoters, that is, upon infection of a human host, the external temperature increases from 21°C to 37°C. In the case of *Y. pseudotuberculosis*, such a shift in temperature results in induced palmitoylation of lipid A (*120*). Palmitoylation is also induced in lipid A isolated from *Y. enterocolitica* grown under Mg^{+2}-limited conditions (*113*). In case of *Y. pestis* no such palmitoylation occurs likely because of a nonsense mutation at its 3' end. This results in a truncated PagP, which may not be able to fold correctly in the OM (*60*).

1.8.1.4 PagP in *Erwinia*:

Erwinia carotovora is a plant pathogen that degrades plant cell walls by generating specific enzymes. Interestingly, the palmitoylated lipid A species can only be isolated from the bacteria under specific pH, temperature and divalent cation concentrations (*116*).

1.8.1.5 PagP in *Photorhabdus*:

Photorhabdus luminescens is a symbiont of entomopathogenic nematodes and an insect pathogen. A homologue of the PhoP/PhoQ system is associated with virulence in these bacteria. However, *P. luminescens* PagP is induced by Mg^{+2}-limitation through a PhoP/PhoQ-independent mechanism (*117*).

1.8.2 Function:

PagP is predominantly an acyltransferase, but also exhibits slow phospholipase activity in the absence of lipid A *in vitro*. For the acyltransferase activity, a phospholipid donor and the lipid A component of the LPS are the required substrates. As an acyltransferase PagP catalyses the transfer of a palmitate chain from the *sn*-1 position of a phospholipid to the *R*-3 hydroxymyristate chain at position 2 of lipid A (figure 1.13). Although PagP is not specific for the head group of the phospholipid donor, it shows high specificity for the length of the acyl chain transferred (*70*). In *in vivo* systems, observable activity of PagP is limited due to the distinct asymmetry of the OM. The phospholipid donor is present exclusively in the inner leaflet while LPS is restricted to the outer leaflet. A large body of evidence has suggested that in the events of membrane perturbation, for example, after ethylenediaminetetraacetic acid (EDTA) treatment, 30-50% of LPS molecules are shed from the cell surface (*121*), promoting the migration of phospholipids from the inner to the outer leaflet to replace the lost LPS (*122*). The resultant patches of phospholipid bilayer in the OM causes cells to become permeable to lipophilic solutes

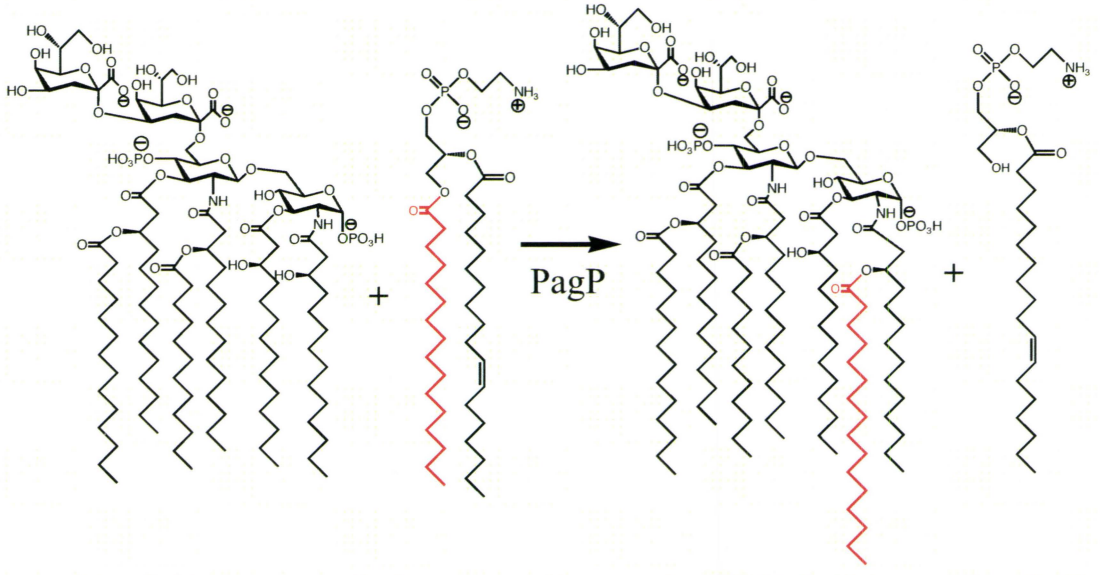

Figure 1.13 PagP enzymatic reaction: PagP catalyzes the transfer of palmitate from the *sn*-1 position of phosphatidylethanolamine (PtdEtn) to the lipid A *R*-3-hydroxymyristate group at position 2, generating hepta-acylated lipid A and *sn*-1-lyso PtdEtn.

that normally cannot penetrate the OM. In this scenario, PagP can access both lipid A and phospholipids that have migrated to the outer leaflet following membrane perturbations. By catalyzing the transesterification reaction, PagP may contribute to the restoration of OM lipid asymmetry (*123*). EDTA-induced lipid A palmitoylation occurs at a faster rate than *pagP* induction through PhoP/PhoQ and is independent of *de novo* protein synthesis (*123*). This suggests that PagP remains dormant in the OM until both required substrates are present in the OM outer leaflet to be accessed by PagP.

1.9 Thesis Problem:

The overall aim of my research project is to establish a detailed structural and functional model of PagP. In order to construct such a model, I approached the problem by examining two distinct questions:

1. Why does PagP exhibit high selectivity for a palmitate chain?
2. What is the mechanism of substrate entry into PagP's active site?

Why does PagP exhibit high selectivity for a palmitate chain?

As indicated previously, PagP shows a high degree of specificity for a palmitate chain in its acyltransferase reaction. An explanation for this behaviour surfaced upon solving of the PagP crystal structure. Since PagP is a membrane protein, its crystal structure was obtained in the presence of a zwitterionic detergent, lauroyldimethylamine N-oxide (LDAO). An LDAO molecule bound in the upper part of the barrel identified the

presence of an internal hydrophobic pocket in PagP (figure 1.11). The detergent molecule was aligned with the membrane normal and had its polar head group facing the extracellular face, while the acyl tail extended into the protein.

The presence of the hydrophobic pocket suggested a role as a hydrocarbon ruler, conferring substrate specificity upon the enzyme. It was hypothesized that by manipulating the depth of the hydrophobic pocket in PagP, the ruler's preference may shift to acyl chains of different length. A Gly88 lining the floor of the pocket was replaced with a series of non-polar amino acids possessing unbranched side chains, specifically, Ala, Met and Cys. These mutants decreased the depth of the pocket and shifted PagP's preference to C15 and C12 for the Ala and Met mutants, respectively. The Cys mutant failed to exhibit single methylene unit resolution and selected both C14 and C15 (*111*). Furthermore, the role of this hydrophobic pocket in PagP catalysis was also signified by the fact that PagP activity is inhibited in the presence of detergents like LDAO and dodecylphosphocholine (DPC) that have linear acyl tails that can occupy the hydrophobic interior of PagP. On the other hand, detergents like dodecylmaltoside (DDM) and CYFOS 7, which have bulkier groups and are unable to occupy the hydrophobic pocket, can support PagP activity (*111*).

As a consequence of the study published by our group in Ahn *et al.*, 2004, we revisited the hydrocarbon ruler hypothesis in order to establish a detailed understanding of lipid-enzyme interactions occurring in PagP's core. We wished to construct a PagP

mutant that would have preference for a C13 acyl chain. Since no naturally occurring amino acid that has a hydrophobic, linear 3-atom side chain exists, we decided to use chemical methylation to generate such a PagP protein. The rationale behind this route was that we had available to us the G88C mutant and techniques had been reported where the sulfhydryl (SH) group of a Cys residue could be chemically methylated. Furthermore, the fact that native PagP lacked Cys residues would allow us to conduct site-specific chemical methylation. There were 2 main challenges for this strategy. Firstly, in order to ensure methylation, the Gly88Cys residue had to be made available to the methylating reagent. This residue sits at the bottom of the hydrophobic pocket and is blocked by LDAO when PagP is folded. Therefore, we had to purify the protein in an unfolded state in guanidine-HCl in order to expose the SH group of the Cys residue. The second challenge was to establish a quantitative technique to determine the degree of methylation. Electrospray ionization mass spectrometry (ESI-MS) was chosen, as it was sensitive enough to detect a small molecular mass difference corresponding to a methylene unit of 14.03 Dalton (Da). Since detergents and salts are not compatible with ESI techniques, a proper solvent system that did not interact with ESI had to be developed. In this regard, we found that unfolded PagP precipitated by dialysis in water could be dissolved in dilute formic acid for ESI (*124*). For the purposes of NMR and crystallography, the protein in guanidine-HCl was refolded by dilution into an appropriate detergent (*110, 111*). Correct refolding of PagP can then be confirmed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Upon purification of all hydrocarbon ruler mutants, we proceeded with evaluation of structural changes associated with mutations and/or chemical modification. By packing the cavity of PagP with hydrophobic side chains, we expected a stabilizing effect on the protein. We decided to measure PagP thermal stability using far-ultraviolet circular dichroism (CD) spectroscopy. We discovered an intrinsic spectroscopic feature known as an aromatic exciton couplet, which provided us with a probe to resolve methylene-units in PagP lipid acyl-chain discrimination. We believe these findings provide a novel approach to the study of lipid acyl-chain selection in enzymology.

What is the mechanism of substrate entry into PagP's active site?

The crystal and NMR structures of PagP showed the presence of Pro28 and Pro50, which limit the number of hydrogen bonds that can form between strands A and B. This results in the formation of a β-bulge (*110, 125*), which may prove to be a site through which substrates can gain access to the catalytic residues of PagP. Also, the presence of proline rich strands F and G might also provide an access site to the β barrel interior without the formation of a β-bulge. We reasoned that the most likely scenario in PagP is lateral entry of lipid substrates from the membrane plane to the enzyme's active site. Our strategy for approaching the problem is to systematically block each potential entry site and perform detailed *in vitro* analysis.

Previously, intimate understanding of Cys chemistry allowed for successful chemical methylation of the Gly88Cys mutant. A similar approach can be adopted for

barricading substrate entry sites by introducing Cys residues on β strands A and B and on β strands F and G. Cys residues under oxidizing conditions not only lend themselves to disulfide bond formation, but also provide reactive thiols that may be modified with chemical reagents.

The mutants were designed at sites where their side chains are oriented towards each other. By examining PagP structure, we selected Pro28/Pro50 and Pro127/Pro144 as the mutant pairs, since their γ carbons are oriented sufficiently close to each other that disulfide bonds may form. There are several chemical and spectroscopic techniques available that report the formation of disulfide bonds. For example, disulfide bridges give rise to signals in near-UV CD spectrum. Sometimes these signals may be hard to detect due to the high level of noise that is present in the near-UV spectrum. Alternatively, specific chemicals i.e. Ellman's reagent react with free sulfhydryl groups to detect the absence of disulfide bridges. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) can also prove to be a useful tool to establish refolding and the presence of disulfide bonds. We have previously used this technique to study refolding of PagP (*126*). The four Cys mutations are localized on the surface of PagP and are exposed. Therefore, intermolecular disulfide bonds can form instead of intramolecular disulfide bonds and retard the mobility of PagP in an SDS-PAGE gel. One must also consider that spontaneous intramolecular disulfide bonding may not occur, either due to the reduced state of the sulfhydryl groups in a membrane environment, or perhaps because of large distances over which disulfide bridges may have to form. The problem

of the reduced state of –SH group can be solved through use of copper phenanthroline $Cu(OP)_3$, which promotes oxidation of the sulfhydryl groups and promotes disulfide bridge formation (*127*). The large distance over which a disulfide bridge may have to form can be assisted through use of Dibromobimane (bBBr). The compound reacts with two thiol groups that are within 3 to 5 Å of each other forming an alkyl bridge (*128*). It has been used to link domains in investigations of P-glycoproteins, lactose permease and ArsA ATPase (*129-131*). Its fluorescence properties make it a valuable tool in the detection of correctly formed disulfide bridges.

Upon protein purification and refolding, site-specific entry of substrates will be investigated by measuring the phospholipase and the acyltransferase activities of PagP. The mutant that blocks lipid A access should convert PagP into a dedicated phospholipase, while the prevention of phospholipid entry will abolish all enzymatic activity.

2.0 Gauging a Hydrocarbon Ruler by an Intrinsic Exciton Probe:

Work presented in this chapter was published in the peer-reviewed journal *Biochemistry* in 2007: Khan, M.A., Neale, C., Michaux, C., Pomès, R., Privé, G.G., Woody, R.W., Bishop, R.E. "Gauging a hydrocarbon ruler by an intrinsic exciton probe". *Biochemistry*. 2007;46(15):4565-79. This project was initiated as part of my Masters program at the University of Toronto (U of T), and was completed during my PhD studies at McMaster University. The construction and analysis of Gly88Ala, Gly88Cys, Gly88Cys-*S*-methyl and Gly88Met mutants was conducted at the U of T. Investigation of Tyr26Phe, Trp66His and Trp66Phe mutants and experiments involving high-temperature CD and the effects of SDS upon PagP's folding states were performed at McMaster University. Theoretical CD calculations were conducted in the laboratories of Dr. Robert Woody (Colorado State University). Chris Neale in the laboratory of Dr. Régis Pomès (U of T) and Dr. Catherine Michaux in the laboratory of Dr. Gilbert Privé (U of T) performed MD simulations and DSC experiments, respectively.

2.1 Introduction:

The diversity of acyl chains found in membrane lipids reflects the ability of cells to modulate membrane biophysical states and to employ specific lipids in signal transduction pathways (*132, 133*). Enzymes of lipid metabolism encounter an effective combinatorial library in their substrates and are necessarily endowed with mechanisms for selecting specific acyl-chain types. Most lipid-metabolizing enzymes of known structure represent soluble globular domains that exist either free in solution or as

monotopic membrane proteins (*134, 135*). Some of these structures indicate how enzymes interact with soluble lipid substrates and have revealed acyl-chain measuring devices known as hydrocarbon rulers (*136, 137*). The monotopic membrane enzyme structures further indicate how globular domains interact with their substrates at the periphery of lipid bilayers. However, many key lipid metabolic steps are catalyzed by integral membrane enzymes, which are embedded in their substrates and encounter a low dielectric milieu not normally accessible to soluble globular protein domains. The biophysical basis of integral membrane lipid-protein recognition is advancing with increasing numbers of determined structures, but integral membrane enzymes of lipid metabolism have mostly proven refractory to structure determination (*138*). Exceptions are found in the OMs of Gram-negative bacteria, which have recently revealed integral membrane protein structures for a phospholipase, two lipid deacylases and a lipid acyltransferase (*98, 109-111, 139*). These intrinsic membrane β-barrel enzymes are more robust than the α-helical integral membrane enzymes of lipid metabolism, and are now providing insights into the molecular mechanisms of lipid acyl-chain selection within a lipid bilayer environment.

Many integral membrane β-barrel proteins can be purified in an unfolded state and refolded at high concentrations in detergent micelles, which are transparent to certain spectroscopic analyses of protein structure and stability. Extrinsic spectroscopic probes can be introduced into proteins to study structure-function relationships, but these studies must be carefully scrutinized to avoid artefacts arising from structural perturbations

introduced by the probe itself. If Trp and Tyr residues are localized in a functionally interesting protein region, their aromatic side chains can be exploited as intrinsic probes to prevent such artefacts. However, it can be difficult to deconvulate the signals that arise from a single aromatic side chain if multiple copies are present in a given protein. More rarely, two or more Trp and/or Tyr residues that interact within specific geometrical and distance constraints can afford a so-called exciton interaction. The exciton arises from the delocalization of the excited states of two interacting chromophores (*140*). The strong $\pi \rightarrow \pi^*$ transitions arising from Trp and Tyr side chains do not on their own generate Cotton effects that can be detected by Circular Dichroism (CD) spectroscopy. Cotton effects arise when parallel components of electric and magnetic dipole transition moments are combined, and strict group theoretical rules show that this cannot happen when a chromophore possesses centers or planes of symmetry (*141*). When they are placed in a chiral environment, such as that provided by the polypeptide backbone of a protein molecule, an exciton effect arising from a pair of interacting chromophores can generate two Cotton effects of equal magnitude and opposite sign that are slightly separated by the so-called Davydov energy (*142*). The resultant of these two overlapping Cotton effects is a bisignate curve known as an exciton couplet, which can be detected in the far-UV region of a protein CD spectrum (*143*).

Exciton theory lies at the heart of protein CD spectroscopy because it accounts for key spectroscopic signatures that arise between $\pi \rightarrow \pi^*$ transitions in interacting peptide groups located within the secondary structural elements. However, exciton couplets

arising from interacting aromatic side chains have largely been ignored in protein secondary structure analysis, despite their clear overlap in the far-UV range with the signals arising from secondary structure (*144*). Due to the strong rotational strengths associated with the α-helix, the weaker aromatic exciton couplets in the far-UV range are most apparent in proteins with low α-helical content, but their presence can significantly influence the far-UV CD spectrum of any protein (*145, 146*). Theoretical analyses show the exciton couplets arising from aromatic side-chain interactions are among the more reliably predicted CD spectroscopic signatures that can occur within a protein molecule (*147*). Algorithms that consider aromatic exciton interactions for predicting a protein CD spectrum from a given set of crystal structure coordinates are available. One can usually pinpoint interacting aromatic side chains associated with an exciton couplet by systematically replacing each amino acid with Ala *in silico* and subtracting the calculated CD spectrum from that of the wild type (wt)-protein (*147*). Given the strict geometrical requirements of the exciton effect, aromatic exciton couplets are anticipated to be extraordinarily sensitive to local structural perturbations that might well be introduced during the experimental modification of any given protein structure. A considerable degree of untapped potential lies in the identification and application of aromatic exciton interactions to study protein function.

We have been investigating the structure and function of PagP, a 161-amino acid membrane protein that resides within the OMs of pathogenic Gram-negative bacteria such as *E. coli* (*60*). PagP is an enzyme of lipid metabolism that transfers a palmitate

chain from a phospholipid molecule to the lipid A (endotoxin) component of LPS (*148*). The Gram-negative OM is an asymmetric bilayer that normally displays lipopolysacchride (LPS) in the outer leaflet and restricts phospholipids to the inner leaflet (*149*). This asymmetric lipid organization creates a permeability barrier to hydrophoboic antibiotics and detergents that are normally encountered in the host and natural environments (*3*). PagP provides bacteria with a degree of resistance to host-derived antimicrobial agents and attenuates the ability of endotoxin to activate the TLR-4 host defense pathway (*65, 113*). Additionally, PagP can function as an apical sensory transducer that reports perturbations of lipid asymmetry by a novel signal transduction mechanism in bacteria (*150*).

PagP is an eight-stranded antiparallel β-barrel with a short α-helix at its N-terminus, and it sits in the membrane with the barrel axis tilted by roughly 25° (Figure 2.1A) (*111*). The PagP palmitate recognition pocket, known as the hydrocarbon ruler, resides within the interior of the β-barrel and is localized in the outer LPS-exposed region of the protein (*111*). A single molecule of the detergent lauroyldimethylamine *N*-oxide (LDAO) serves to identify the position of the hydrocarbon ruler within the β-barrel interior (Figure 2.1A). Two discontinuities in β-strand hydrogen bonding within this LPS-exposed region provide obvious routes for lateral access of lipids to the hydrocarbon ruler (*110, 125*). The implication that PagP depends on the aberrant migration of phospholipids into the outer leaflet is supported by observations that the enzyme

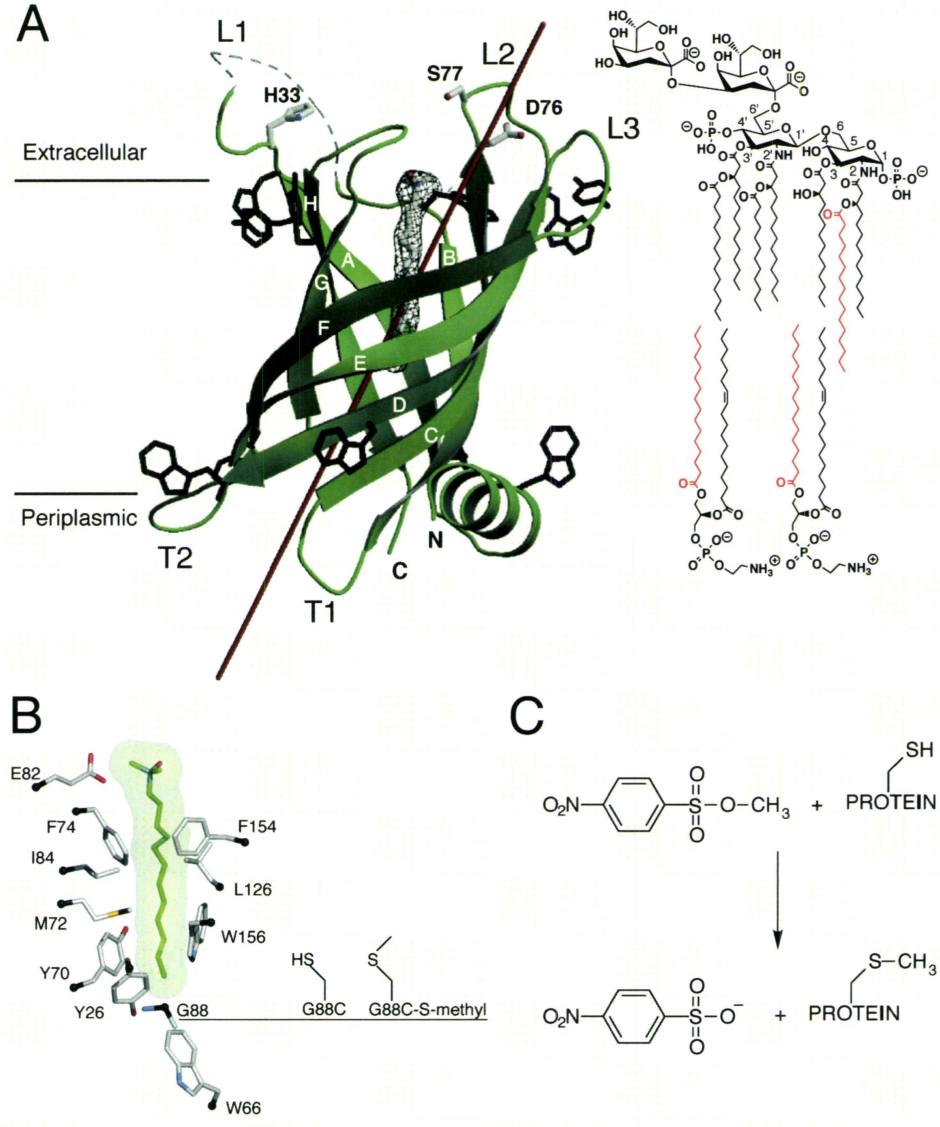

FIGURE 2.1: Site-specific chemical modification of the PagP hydrocarbon ruler: (A) PagP resides in the *E. coli* outer membrane, where LPS (shown as Kdo_2-lipid A) lines the outer leaflet and phospholipids line the inner leaflet. A bound LDAO detergent molecule resides within the PagP hydrocarbon ruler, which has lateral access to lipid substrates in the outer leaflet. PagP transfers a palmitate chain (red) from a phospholipid to lipid A and depends on the aberent migration of phospholipids into the outer leaflet.(B) The hydrocarbon ruler dimensions determine PagP's ability to strictly select a 16-carbon saturated palmitate chain. Substitution of Gly88 at the base of the hydrocarbon ruler can cause PagP to select shorter acyl chains. (C) Strategy for site-specific *S*-methylation of Gly88Cys PagP using MNBS.

normally remains dormant in outer membranes until its activity is directly triggered by perturbations to lipid asymmetry (*123*).

A striking enzymological feature of PagP lies in its ability to distinguish a 16-carbon saturated palmitate chain in phospholipids from all other chains, even those that differ by a single methylene unit (*70, 111*). This ability to distinguish acyl chains of different length is due to the presence of a hydrocarbon ruler, which was identified upon solving the PagP crystal structure. The depth of the PagP hydrocarbon ruler, which is lined at its base by Gly88, determines this single-methylene unit resolution (Figure 2.1B). PagP lacks any Cys residues, but is particularly rich in aromatic amino acids and includes 10 Phe, 11 Tyr and 12 Trp residues, not including the N-terminal signal peptide. During our efforts to modulate acyl-chain selection by site-specific chemical methylation of a Gly88Cys PagP mutant, we have discovered an exciton couplet that could be pinpointed to an interaction between Tyr26 and Trp66 at the floor of the hydrocarbon ruler. We demonstrate that this intrinsic exciton probe can provide a gauge for reporting methylene unit resolution in the PagP hydrocarbon ruler.

2.2 Materials and Methods:

All materials used in these experiments are listed in table 1.

2.2.1 Expression and Purification of PagP:

To purify PagP in a denatured state, we expressed it without its native signal peptide in *E. coli* BL21 (DE3) transformed with the plasmid pETCrcAHΔS or its mutant derivatives that were constructed using the QuickChange protocol (Stratagene) as described previously (*110*). Table 2 outlines all oligonucleotide primers used in our mutagenesis, and table 3 identifies all plasmids and bacterial cell lines used. Bacteria were cultured in 1 L of Luria-Bertani (LB) medium supplemented with ampicillin (amp) at 100 μg/ml and grown to an optical density at 600 nm of 0.50 and then induced with 1 mM isopropyl β-D-thiogalactopyranoside for 3 hr. The induced cells were harvested, suspended in 20 mL of 50 mM Tris-HCl (pH8.0) and 5.0 mM EDTA, and broken by being passed twice through a French pressure cell at 8,000 psi, and the insoluble material was recovered by centrifugation in a Beckman MLA-80 rotor at 27,000 rpm and 4°C for 20 min using an Optima MAX-E ultracentrifuge. The pellets were washed in 20 mL of 50 mM Tris-HCl (pH8.0) and 2% Triton X-100 followed by 20 mL of 50 mM Tris-HCl (pH8.0) and then solubilised in 10 mL of 50 mM Tris-HCl (pH8.0) and 6 M Gdn-HCl. Supernatants were collected after centrifugation and loaded onto a 5 mL bed of His-bind resin (Novagen) that was charged with 50 mM $NiSO_4$ and equilibrated with 3 column volumes of 10 mM Tris-HCl (pH8.0), 250 mM NaCl, 6 M Gdn-HCl and 5 mM imidazole. The sample was applied to the column and washed with 10 column volumes of the equilibration buffer, followed by 5 column volumes of 10 mM Tris-HCl (pH8.0), 250 mM NaCl, 6 M Gdn-HCl, and 20 mM imidazole. The samples were eluted in a step

gradient using 5 mL each of 35, 50, 75, 100 and 125 mM imidazole. SDS-PAGE confirmed the presence of purified PagP in the 100 and 125 mM imidazole fractions, which were pooled and dialyzed against distilled water, and the precipitated protein was collected. The protein was resuspended in 8.8 mL of 8 M Gdn-HCl and 0.34 M Tris-HCl (pH8.6), and the concentration of the purified wt protein was determined, using a calculated extinction coefficient (ε_{280}) 82630 M^{-1} cm^{-1}, to be 6.67 mg/ml for a total yield of 58.7 mg. The remaining mutant proteins were obtained in similar yields.

PagP Trp66 mutant that failed to refold *in vitro* could be folded *in vivo* when it was expressed with the intact signal peptide in plasmid pETCrcAH and then purified directly from LDAO-solubilized membranes in 1 mg quantities from 1 L cultures as described previously (*70*).

2.2.2 Methylation of PagP:

WtPagP and Gly88Cys mutant PagP in Gdn-HCl were subjected to an *S*-methylation procedure using methyl-*p*-nitrobenzenesulfonate (MNBS). Control reaction mixtures excluding only MNBS were also prepared, and the four reactions were carried out in capped glass tubes. Protein in 8 M Gdn-HCl and 0.34 M Tris-HCl (pH8.6) was adjusted to 5 mg/ml in a final volume of 5 mL containing 6 M Gdn-HCl, 0.25 M Tris-HCl (pH8.6), 3.3 mM EDTA, and 25% (v/v) acetonitrile. The solutions were flushed with N_2(g) for 1 min to create an anoxic barrier, and 50 µl of 260 mM β-mercaptoethanol (βMe) was added (10-50 fold molar excess of the protein). The tubes were tightly sealed,

placed in a 50°C bath for 1 hr, and then gradually cooled to 37°C. Under the N_2 barrier, 0.5 mL of 52.0 mM MNBS (2 fold molar excess of the β-Me) was added, and the tubes were tightly sealed and placed in a 37°C bath for 2 hr. The reactions were quenched via addition of 5 µL of 14 M β-Me. The reaction mix was dialyzed exhaustively against distilled water, and the precipitated proteins were collected and dissolved in 10 mM of 10 mM Tris-HCl (pH8.0) and 6 M Gdn-HCl. The protein yield was nearly quantitative (25 mg) for all four samples. An aliquot of each sample (0.5 mL) was dialyzed against water and the precipitated protein used for mass spectrometry.

2.2.3 Electrospray Ionization Mass Spectrometry (ESI-MS):

MS was performed at the Advance Proteomics Facility at the Hospital for Sick Children. Our procedure for dissolving precipitated PagP was adapted from a prior study of lactose permease (*124*). Precipitated samples to be analyzed were dissolved in 5 mL of a 1:1 (v/v) acetonitrile/1% formic acid mixture at a concentration of ~1 ng/µl just prior to ESI-MS and were injected directly into a triple-quadrupole mass spectrometer. The positive ion mode was used and the cone potential maintained at 48 eV and the collision energy at 4 V. The spectra were reconstructed using Mass Lynx 3.5. In one instance, a Sciex API III^+ triple-quadrupole mass spectrometer was used, but all other analysis were performed with a Micromass Quattro Ultima LC-ESI/APCI triple-quadrupole mass spectrometer.

2.2.4 Refolding of PagP:

Samples in 10 mM Tris-HCl (pH 8.0) 6 M Gdn-HCl, and 20 mM βMe were diluted drop wise (~1 drop per 2s) into a 10-fold excess of 10 mM Tris-HCl (pH8.0) and 20 mM βMe containing 0.5% LDAO at room temperature with vigorous stirring and left to stir overnight at 4°C. The βMe was excluded in subsequent refolding experiments using PagP mutants that lack any Cys residues. The refolding sample was then applied to a 4 mL bed of His-bind resin (Novagen) charged with 50 mM $NiSO_4$ and equilibrated with 10 mM Tris-HCl (pH 8.0), 0.1% LDAO, and 5 mM imidazole. The column was washed with 10 column volumes of the equilibration buffer and 10 column volumes of 10 mM Tris-HCl (pH 8.0), 0.1% LDAO, and 20 mM imidazole and then eluted with 2 mL of 10 mM Tris-HCl (pH 8.0), 0.1% LDAO, and 250 mM imidazole. The sample was dialyzed against 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO, and the refolded samples (6 μg) were resolved via SDS-PAGE on a 1 mm Novex 16% Tris glycine precast gel under reducing or non-reducing conditions (Invitrogen). The heated samples were boiled at 100°C for 10 min before loading, and protein was stained using Coomassie blue dye. Disulfide bonds were encouraged to form under nonreducing conditions by uncapping the heated and unheated samples on the bench for several hours prior to loading on the gel. Refolded PagP protein concentrations were determined using either the bicinchoninic acid assay (*151*) or by absorbance using an extinction coefficient (ε_{280}) determined experimentally by the Edelhoch method (*152, 153*).

2.2.5 CD Spectroscopy:

Samples to be analyzed for CD were maintained at a concentration of 0.3 mg/ml in 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO and were analyzed using a cuvette with a path length of 1 mm. The samples were analyzed using either a Jasco J-812 or an Aviv 215 CD spectrometer, each of which were linked to Peltier devices for temperature control. For each sample, three accumulations were averaged at a data pitch of 1 nm and a scanning speed of 10 nm/min. The temperature was maintained at 25°C, and data sets were obtained from 200 to 280 nm. For thermal denaturation profiles, samples at a concentration of 0.80 mg/ml in a 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO were analyzed in a cuvette with a path length of 1 mm. The samples were heated from 20 to 100°C at a rate of 2°C/min, with a response time of 16 s.

2.2.6 Differential Scanning Calorimetry:

The specific heat capacity (C_p) as a function of temperature was obtained in a N-DSC (cell volume of 0.3 mL) at a scan rate of 1.0 °C/min and a pressure of ~37 psi. The protein sample was first dialyzed overnight at 4°C against 100 mM sodium phosphate buffer (pH 8.0) with 0.05% LDAO. The solution was then degassed for 5 min before it was loaded into the DSC cells. A blank scan with buffer in both calorimeter cells was subtracted to correct for the difference between the cells. The DSC instrument that was employed was a CSC model 6100 nano II. The software used to collect the data was DSCRUN, N-DSC control program (version 2.5.0.29s), Calorimetry Science Corp. The

software used to visualize the data was cpcalc (version 2.1, Applied Thermodynamics), while the software used to process the data was Microcal TM (version 5.0, Microcal Software Inc.).

2.2.7 Theoretical Calculations of CD:

The CD spectrum for wtPagP was calculated using the X-ray structure (*111*) (PDB entry 1THQ). The methods and parameters have been described previously (*146*) (*154*). Calculations were also performed for mutants generated *in silico* by deletion, one by one, of each of the Tyr and Trp side chains in the vicinity of the G88 in the crystal structure: Tyr26, Trp60, Trp66, Tyr70, Tyr87 and Trp156. (Phe side chains were not mutated because they have only the weak 260 nm band above 220 nm). In each case, a difference spectrum was calculated by subtracting the predicted spectrum of the mutant from that of the wt, thus providing the contribution of the mutated side chain to the wt CD spectrum.

2.2.8 Hydrocarbon Ruler Assays:

Kdo_2-lipid A was prepared from heptose-deficient *E. coli* WBB06 as described previously (*155*) and quantified as described (*156*). Synthetic diacylphosphatidylcholines were obtained from Avanti Polar Lipids (Alabaster, AL). The hydrocarbon ruler assays were performed at 30°C by a TLC-based radiolabeling procedure as described previously (*111*), except that 25 μCi of [^{32}P]-orthophosphate was employed in place of sodium [^{14}C]-acetate.

2.2.9 Molecular Dynamics:

We have used high-temperature simulated annealing to probe the local conformational changes induced by a mutation of Gly88. The same procedure has been applied to two starting conformations. The initial conformation was the X-ray structure (*111*) (PDB entry 1THQ) in the absence of LDAO and crystal waters. Another starting conformation was obtained by equilibration of the X-ray structure containing a rebuilt L1 loop and bound LDAO, and embedded in an explicit 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (*157*) bilayer and TIP3P water (*158*). All nonprotein atoms were then removed. Residues His22 and His102 were protonated in $N^{\delta 1}$, while residues His33 and His67 were protonated on $N^{\varepsilon 2}$. All simulations were carried out using the CHARMM software package (*159*) using the CHARMM22 all-hydrogen topology and parameter files. Parameters for the methylated cysteine residue were developed on the basis of methionine via removal of C^{β}. All nonbonded interactions (Lennard-Jones and Coulombic) were switched off from 10 to 11Å. The neglect of the long-range electrostatic interactions is justified because conformational sampling is restricted to the local side-chain rearrangements of neutral residues within a static protein framework.

Initial conformations of the mutants were obtained as follows. Hydrogen atoms to be replaced by heavy atoms were chosen such that the nascent side chain was in the LDAO binding pocket and pointed toward the extracellular region. Mutants were created according to the following succession: G88A ($H^{\alpha 1} \rightarrow C^{\beta}$, $H^{\alpha 2} \rightarrow H^{\alpha}$), A88C ($H^{\beta 3} \rightarrow S^{\gamma}$), C88-

S-methyl ($H^{\gamma} \rightarrow C^{\delta}$), and C-*S*-methyl88M ($S^{\gamma} \rightarrow C^{\gamma}$. $C^{\delta} \rightarrow S^{\delta}$, $H^{\delta 1} \rightarrow C^{\varepsilon}$). A harmonic restraining potential was applied for C′, N′, C^{α}, and O′ atoms of the residue 88 with a force constant of 10 kcal mol^{-1} A^{-2}, and all other residues were held fixed. Each mutant was energy minimized.

In the remaining calculations, a harmonic restraining potential was applied to the backbone heavy atoms of residues Trp66, Gly68, Tyr70, X88, Thr108, and Leu128 with a force constant of 10 kcal mol^{-1} A^{-2} (greater by 1 order of magnitude for calculations starting from the X-ray structure). All other residues were held fixed. Each of the conformations resulting from the procedure outlined above was simulated at 3000 K for 75 ps. The temperature was controlled by Langevin dynamics with a friction coefficient of 2.0 ps^{-1}. High temperature seed conformations were taken 15, 25, 35, 45, 55, 65 and 75 ps. Each seed was then cooled to 260 K over a minimum of 1.3 ns by reducing the temperature to 0.8 of its previous value after each segment. Dynamics at and below 3000 K were produced with time steps pf 0.5 and 1.0 fs, respectively. The final cooling step was followed by energy minimization.

Figures were generated with VMD version 1.8.3 (*160*) and rendered with POV-Ray version 3.6 (Persistence of Vision Pty. Ltd.). W66 and X88 are overlaid for all seven replicas. Contouring of the β-barrel interior was performed by HOLE with a cutoff radius of 1.4 Å (*161*). For the wt and each mutant, contours were generated on the basis of the initial wt structure plus the coordinates of the residue 88 overlaid from all seven replicas.

All other figures derived from the PagP crystal structure coordinates were rendered using PyMOL.

2.3 Results:

In our prior investigation, we analyzed membrane-derived wtPagP and its site-specific mutants Gly88Ala, Gly88Cys and Gly88Met (*111*). To improve the quality of these initial studies, we removed the signal peptide from each of the mutant proteins for expression and purification in a denatured state to be followed by refolding. Since PagP lacks any Cys residues, the single Cys in Gly88Cys PagP was available for a chemical methylation procedure that occurs in Gdn-HCl needed to unfold PagP, which is necessary to expose the Cys residue and to normalize the reactivity of other functional groups (*162*). By synthesizing Gly88Cys-*S*-methyl PagP, we predicted that we could create a suite of hydrocarbon ruler mutants that select 15-carbon (Gly88Ala), 14-carbon (Gly88Cys), 13-carbon (Gly88Cys-*S*-methyl), and 12-carbon (Gly88Met) saturated acyl chains. The hydrocarbon ruler hypothesis predicts that the substituted amino acid side chain at position 88 will plug the floor of the hydrocarbon ruler and make it shallower by the same length as the introduced side chain. The change in hydrocarbon ruler dimensions should afford a corresponding shortening of the acyl chain that is selected in enzymatic assays.

2.3.1 Site-Specific Chemical Modification of Gly88Cys:

We subjected the unfolded wtPagP and Gly88Cys proteins to methylation with MNBS (Figure 2.1 C) as described previously (*111*). Control reactions were also performed without MNBS. To detect the incorporation of a single methylene group (14.03 Da) into an ~20 kDa protein, we chose to determine the mass by ESI-MS. The method is incompatible with the Gdn-HCl or LDAO used to keep PagP in solution, but we were able to develop a mild procedure, adapted from a much harsher one using formic acid to perform ESI-MS on lactose permease (*124*); our adaptation involves removing denaturants by dialysis against water and dissolving the precipitated protein in a 1:1 acetonitrile/1% formic acid solution immediately prior to ESI-MS analysis. We could demonstrate that MNBS leads to a 14 Da increase in the mass of the Gly88Cys mutant, but not in the wtPagP (Table 4). The mass spectra revealed that the Gly88Cys mutant in the absence of MNBS treatment forms two species, the major one being a disulfide-linked dimer (Figure 2.2). MNBS treatment of the Gly88Cys mutant reveals a single molecular species corresponding to a monomeric *S*-methylated protein. These findings validate the claim that the methylation procedure is both quantitative and highly selective for Cys under the specified conditions (*162*). We were also able to verify the masses of the remaining PagP mutants by ESI-MS.

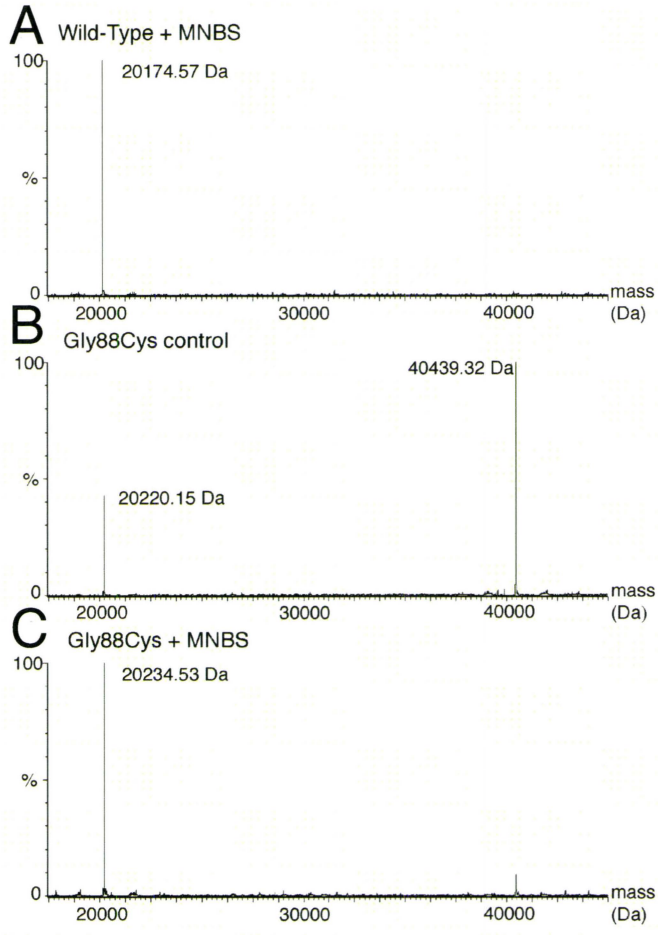

FIGURE 2.2 ESI-MS confirmation of Gly88Cys-*S*-methyl PagP: (A) Wild-type PagP treated with MNBS exhibits a single mass that corresponds to an unmodified monomeric species. (B) In the absence of MNBS, Gly88Cys PagP exhibits two masses that correspond to an unmodified monomer and a disulfide-linked dimer. (C) MNBS treatment of Gly88Cys PagP results in a 14 Da mass increase for the monomeric species and a loss of the disulfide linked dimer.

2.3.2 Evaluation of PagP Refolding:

The unfolded PagP variants were refolded by dilution into LDAO and evaluated by a qualitative SDS-PAGE assay (Figure 2.3). Many OMPs like PagP exhibit the phenomenon of heat modifiability on SDS-PAGE. In these gels, the unheated form of PagP appears folded and more compact as it migrates slightly ahead of its expected molecular mass of 20 kDa (*70*). However, after irreversible heat denaturation, PagP migrates as its expected molecular mass. This explanation for the observed heat modifiability is reinforced by the observation that only the heat-denatured Gly88Cys is capable of migrating as a disulfide-linked dimer in the nonreducing gel. The single Cys in this mutant appears to be unexposed in the unheated sample, consistent with the faster migrating species reflecting a folded state of the protein. WtPagP and its Gly88 mutant derivatives all appeared to be adequately folded by this SDS-PAGE criterion.

2.3.3 Spectroscopic Examination of Refolded PagP Mutants:

We wanted to determine whether subtler structural details could be ascertained by a more refined CD spectroscopic examination of the refolded PagP proteins. The LDAO micelles are transparent down to 200 nm in CD analysis of the refolded PagP proteins. The far-UV CD spectrum of wtPagP (Figure 2.4 A, B) displays a negative CD band at 218 nm that is characteristic of the $n \rightarrow \pi^*$ transition derived from peptide bonds in a largely β sheet conformation (*163*). This observation is consistent with the β barrel structure of PagP. However, the positive ellipticity at 232 nm is likely the first Cotton

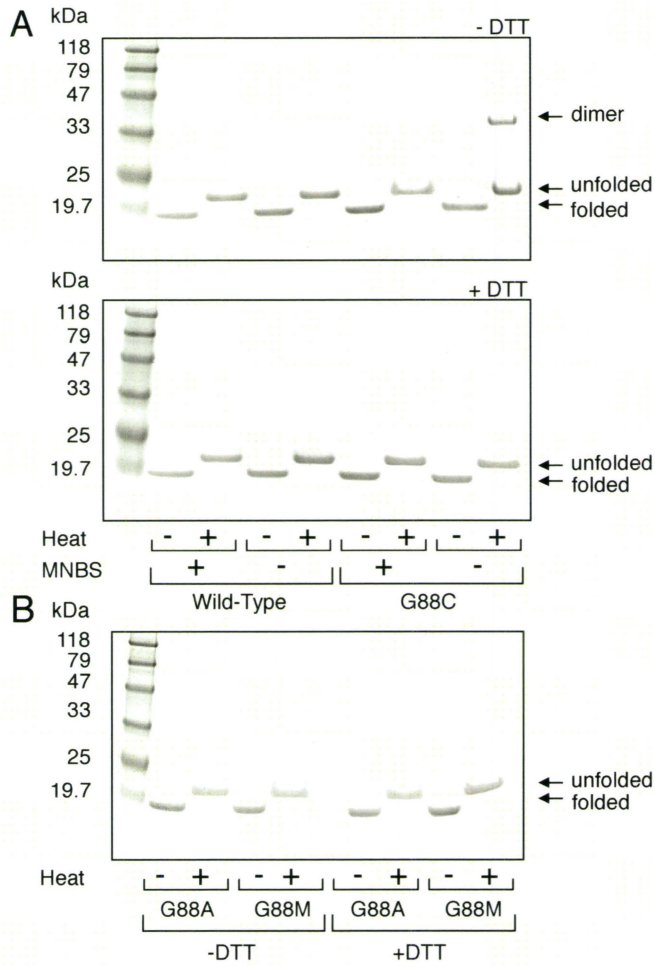

FIGURE 2.3 SDS-PAGE analysis of PagP refolding: PagP requires heat treatment to become denatured via SDS-PAGE and migrates anomalously fast in the absence of heat treatment. (A) Analysis of wild-type and Gly88Cys PagP under reducing and nonreducing conditions, in the presence and absence of heat treatment, and with or without prior exposure to MNBS. (B) Analysis of Gly88Ala and Gly88Met PagP as described above, but without exposure to MNBS.

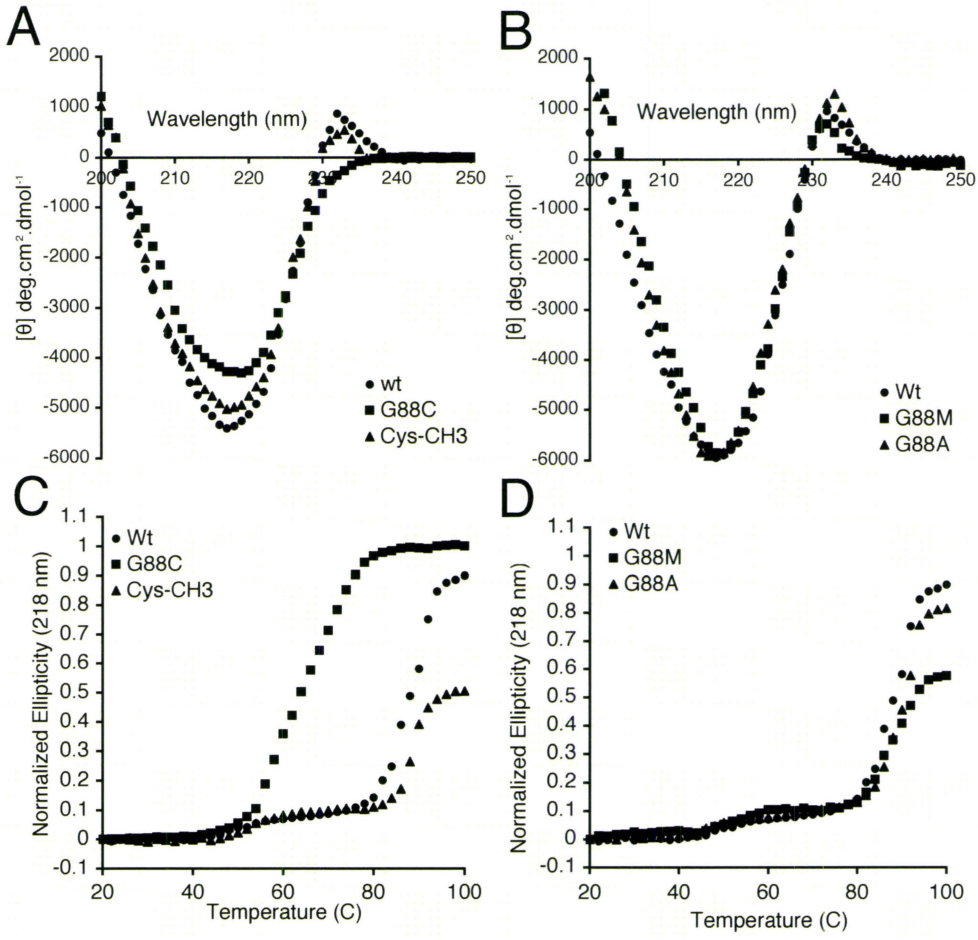

FIGURE 2.4 Spectroscopic analysis of refolded PagP variants and evaluations of thermal stability: (A and B) Far UV-CD spectra of refolded PagP variants. (C and D) Thermal unfolding profiles of the refolded PagP variants assessed by following the loss of negative ellipticity maximum at 218 nm observed in panels A and B.

effect of an exciton couplet, which becomes fully apparent only by its absence in Gly88Cys PagP. The second Cotton effect of this exciton couplet is largely superimposed on the β sheet n→π* transition. The Gly88Cys PagP sulfhydryl group appears to be responsible for extinguishing the exciton couplet, because Gly88Cys-*S*-methyl PagP restores the exciton couplet to its full strength. The far-UV CD spectra of Gly88Ala and Gly88Met similarly exhibit a robust exciton couplet.

We reasoned that the absence of the exciton couplet in Gly88Cys might result from a local structural perturbation arising from the sulfhydryl group. We followed loss of the negative ellipticity maximum at 218 nm in the thermal unfolding experiments using wtPagP and its hydrocarbon ruler mutants (2.4 C, D). PagP was unfolded irreversibly in these experiments because it precipitated after incubation at 100°C. However, the proteins that exhibit the exciton couplet all share a thermal denaturation temperature near 88°C, while Gly88Cys denatures near 65°C. The smaller perturbation apparent near 50°C in all but Gly88Cys likely represents loss of the excitonic Cotton effect that contributes to the ellipticity at 218 nm because a transition of a similar magnitude is also observed near 50°C by following the loss of the associated positive ellipticity at 232 nm (Figure 2.5 A). This finding suggests that the exciton is much more sensitive to heat, disappearing at a temperature that is easily tolerated by the β barrel structure. The transition at 88°C most likely represents protein unfolding because wtPagP was shown to exhibit a thermal denaturation at the same temperature in a DSC experiment (Figure 2.5 B).

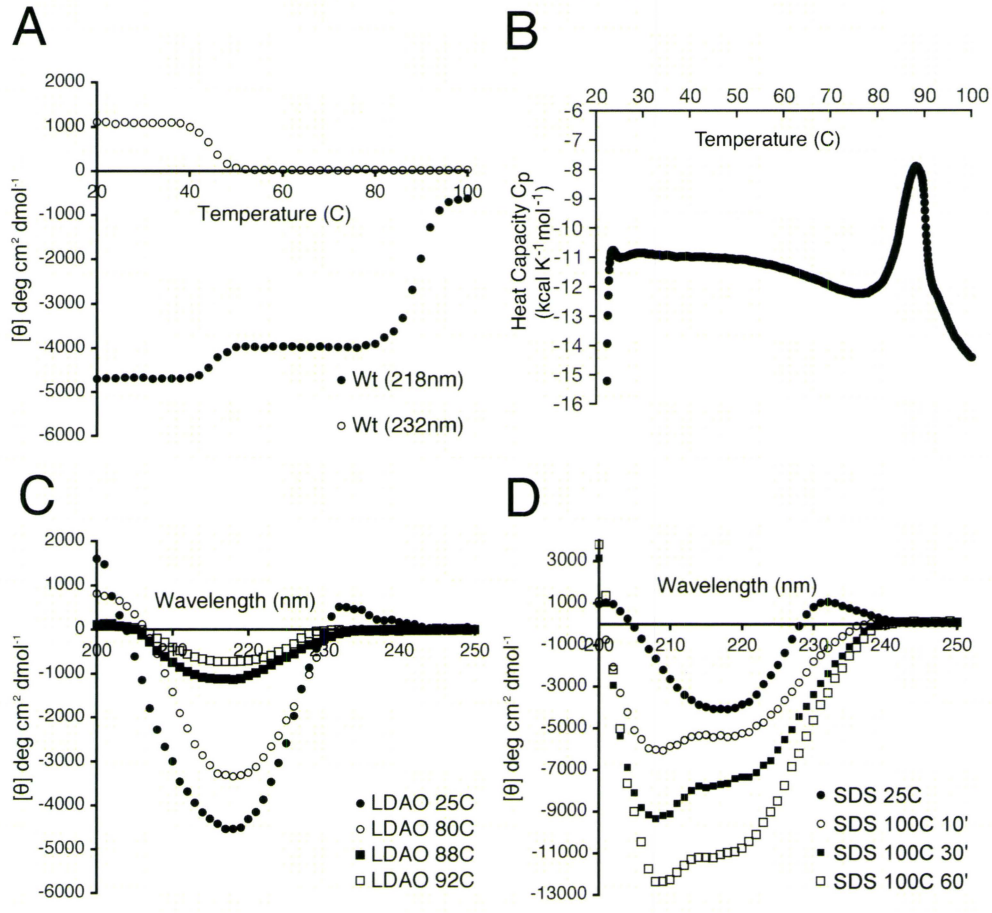

FIGURE 2.5 Secondary structure associated with thermal stability of wtPagP. (A) Thermal unfolding profile of refolded wtPagP assessed by following the loss of ellipticities at 218 nm and 232 nm. (B) Thermal unfolding profile of wtPagP assessed by DSC. (C and D) Far-UV CD spectra of wtPagP refolded in LDAO and recorded at various points along the thermal unfolding profile after a 15 min equilibration (C) and in the presence of 1% SDS as a function of the time of incubation at 100 C.

Far-UV CD spectra recorded at high temperatures validate the disappearance of the excitonic signature prior to protein unfolding and reveal that residual β structure remains at 92°C (Figure 2.5 C). Precipitation ensuing at higher temperatures is consistent with the formation of intermolecular β structure that would compensate for intramolecular β structure lost during β barrel unfolding, because both structures exhibit negative ellipticity at 218 nm. Consequently, the different plateaus reached at high temperatures by the various mutants (Figure 2.4 C, D) likely reflect different tendencies to form intermolecular β structure in the denatured state. Compared to wtPagP, more extensive residual ellipticity is apparent at 92°C in Gly88Cys-*S*-methyl PagP (Figure 2.S1 A), which is consistent with its lower plateau as observed in figure 4 C. In contrast to the aggregation observed in LDAO at high temperatures, a distinct conformation is indicated after heat denaturation in the presence of 1% SDS because we previously observed that heat-denatured PagP migrates in a monodisperse state via SDS-PAGE (Figure 2.3). We find that PagP retains the apparent excitonic signature associated with its β barrel structure after dilution into 1% SDS at room temperature, but boiling the sample over the course of 1 hr followed by cooling to room temperature reveals characteristics of an α helical protein in far-UV CD spectra (Figure 2.5 D). Helix formation in SDS is also indicated as a function of increasing temperature (Figure 2.S1 B) by a gradual increase in rotational strength, an associated red shift of the amide $n \rightarrow \pi^*$ transition from 218 to 222 nm, and the acquisition of a new negative ellipticity maximum at 208 nm, which arises from exciton splitting of the amide $\pi \rightarrow \pi^*$ transition.

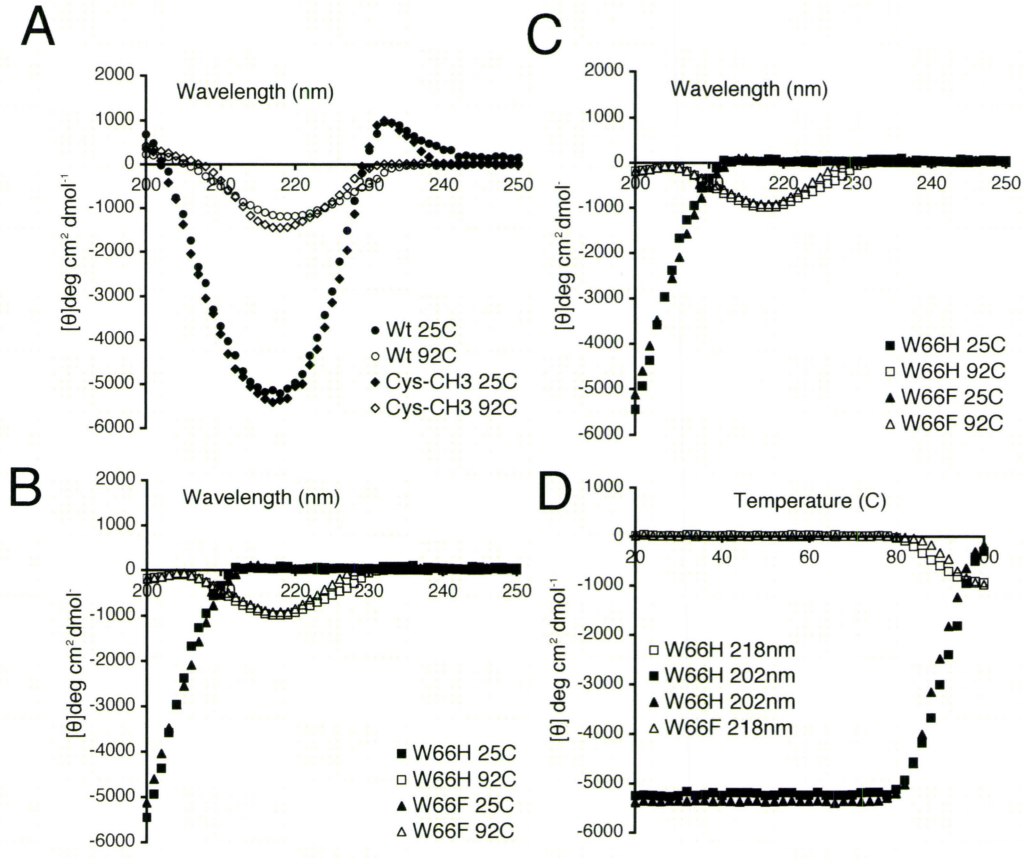

FIGURE 2.S1 Secondary structural analysis of PagP variants at different temperatures: Spectra were taken at the indicated temperature after a 15 minute equillibration using PagP derivatives that were refolded in LDAO as indicated in Experimental Procedures. Panel A shows spectra for wild-type (Wt) and Gly88Cys-*S*-methyl PagP. Panel B shows spectra for Wt PagP refolded in LDAO and diluted into 1% SDS. Panel C shows Trp66His and Trp66Phe PagP, whereas Panel D shows their thermal precipitation curves.

2.3.4 Theoretical Exciton Analysis:

To pinpoint aromatic amino acid residues that contribute to the observed exciton couplet, we used the high-resolution PagP crystal structure coordinates (*111*) (PDB: 1THQ) to predict the CD spectrum. We were encouraged to find that calculations using standard methods and parameters reproduced the experimental positive CD band at 232 nm (predicted $[\lambda]_{max}$ = 1500 at 237 nm). Analysis of the band responsible for the positive feature shows that it is located at 232 nm and is largely due to the 230 nm band of Tyr26 and 225 nm band of Trp66, with contributions of 1-3% from Trp60, Tyr70, Tyr87 and Trp156. This analysis was supported by calculations on *in silico* mutants, in which Tyr26, Trp66 and other participants in the long-wavelength positive band were deleted. Deletion of Tyr26 gave negative CD throughout the 220-240 nm region, and the difference spectrum of Tyr26Ala with the wtPagP exhibited a positive exciton couplet centered at 227 nm. Deletion of the Trp66 gave a spectrum with a weak positive maximum at 238 nm, with only 20% of the wtPagP intensity, and the difference spectrum for Trp66Ala a positive couplet comparable in magnitude to that for Tyr26Ala. For the other mutants, the long-wavelength positive band persisted, with nearly constant contributions from Tyr26 and Trp66. The magnitude of the band was changed only slightly upon deletion of Trp60 or Tyr70. However, deletion of Tyr87 led to a 70% increase in amplitude, whereas deletion of Trp156 led to a 50% decrease. These two side chains make significant but opposing contributions to the long-wavelength exciton band, despite their low fractional participation. Nevertheless, Tyr26 and Trp66 dominate the 232 nm exciton band, and any

perturbation that eliminates this band must affect one or both of these side chains. Therefore, exciton theory applied to the PagP crystal structure has conclusively identified Tyr26 and Trp66 as the main interacting aromatic side chains responsible for the observed exciton couplet.

The structural relationship between the hydrocarbon ruler and the nondegenerate Tyr26-Trp66 exciton pair is shown to indicate how the exciton effect arises from an interaction between $\pi \rightarrow \pi^*$ transitions with similar but distinct energies (Figure 2.6 A). The ^1L_a and ^1B_b electric dipole transition moments in Tyr26 and Trp66, respectively, couple to give two exciton states separated by the splitting energy, $\Delta_{ij} = (\delta_{ij}^2 + 4V_{ij}^2)^{1/2}$, where δ_{ij} is the energy difference between the two transitions, δ_i and δ_j, and V_{ij} is the interaction (Davydov) energy (Figure 2.6 B). Considering only the coupling between the Tyr26 ^1L_a band and the Trp66 ^1B_b band, and using the monopole-monopole interaction energy for calculating the wavelengths and the rotational strengths, the resulting exciton couplet is predicted to be positive and to exhibit a peak-to-trough strength of ~2500 deg cm^2 $dmol^{-1}$ (Figure 2.6 C). The sign of an exciton couplet is defined in accordance with the long wave-length (first) Cotton effect and will be reversed in the enantiomeric structure. In contrast to the degenerate exciton case (with two identical chromophores), where Δ_{ij} reduces to $2V_{ij}$, the mixing of Tyr26 ^1L_a and Trp66 ^1B_b bands is incomplete. The Tyr26 ^1L_a band dominates the long-wavelength Cotton effect, and the Trp66 ^1B_b band dominates the short-wavelength Cotton effect. The exciton couplet will disappear in CD when either V_{ij} or the rotational strength (R_o), representing the area under the two

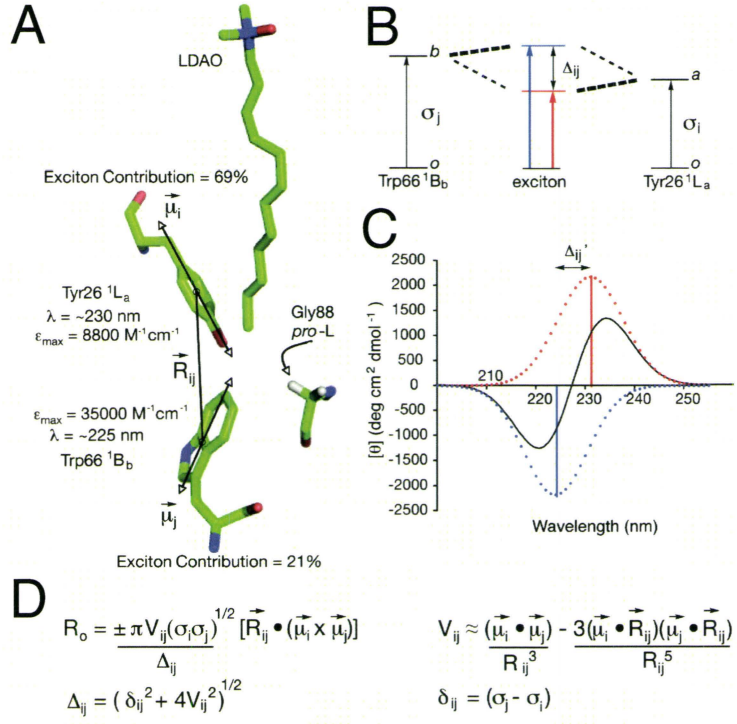

FIGURE 2.6 Theoretical analysis of a nondegenerte exciton interaction in the PagP: (A) Structural relationships between the PagP hydrocarbon ruler, emphasizing the LDAO molecule and the pro-L hydrogen of Gly88, and the exciton interaction between Tyr26 and Trp66. The polarization axes for the interacting Tyr 26 ^1L_a (μ_i) and Trp66 ^1B_b (μ_j) electric dipole moment vectors are shown to emphasize their relationships with the interchromophore distance vector R_{ij}. Key absorption parameters and the calculated exciton contribution to the wild-type PagP far UV-CD spectrum are also shown. (B) The π to π^* transitions from ground states o to excited states a (for Tyr 26 ^1L_a) or b (for Trp66 ^1B_b) are split be nondegenerate exciton interaction into two new exciton states, which generate CD Cotton effects equal in magnitude, opposite in sign, and separated by the splitting energy Δij. This splitting energy is given by $(\delta_{ij}^2 + 4V_{ij}^2)^{1/2}$, where δ_{ij} $(=\sigma_j - \sigma_i)$ is the difference in the transition energies of the two chromophores and V_{ij} is the energy of interaction between the transition moments. Tyr26 and Trp66 each dominate the low-energy (red) and high-energy (blue) exciton states, respectively. (C) The exciton couplet (solid line) calculated from the Tyr26-Trp66 exciton interaction is the resultant of the two exciton-split CD Cotton effects (red and blue dotted curves), which are separated on the wavelength scale by Δ_{ij}', the wavelength equivalent of Δ_{ij}. (D) Theoretical rotational strengths (R_o) of nondegenerate exciton split CD Cotton effects are proportional to V_{ij}, Δ_{ij}, and a scalar triple product involving the vector connecting the centers of the two transitions and the dipole transition moments. V_{ij} can be approximated by the point-dipole approximation, but a more exact method is used in calculating the spectra shown here.

Gaussian distributions, is zero (*140*). Although the point-dipole approximation for calculating V_{ij} is only accurate when the interchromophore distance vector R_{ij} is large compared to the dimensions of the interacting chromophores and is replaced in practice by a monopole (distributed dipole) approximation, it usefully illustrates how V_{ij} is related to the electric dipole transition moment vectors for μ_i and μ_j (Figure 2.6 D). The dot products indicate that V_{ij} is maximized by parallel and minimized by orthogonal vector relationships. In contrast, the scalar triple product in R_o for a nondegenerate exciton interaction dictates that R_o is maximized by orthogonal and vanishes for parallel relationships. Theoretically, the exciton can be extinguished by any structural perturbation arising from a Gly88 substitution that induces parallel or orthogonal relationships from the oblique vector relationships normally found to exist between the Tyr26 ^1L_a and Trp66 ^1B_b transitions.

2.3.5 Experimental Exciton Analysis:

To experimentally validate the predicted Tyr26-Trp66 nondegenerate exciton in PagP, we constructed, purified, and analyzed the Tyr26Phr, Trp66Phe and Trp66His mutants. The Trp66 mutants failed to refold as assessed by SDS-PAGE (Figure 2.7 A), which indicates that Trp66 is a critical determinant of PagP folding. The Trp66His and Trp66Phe mutants remained in solution after failing to refold in LDAO at 25°C, but precipitated after heating to 100°C; their far-UV CD spectra revealed negative ellipticity below 210 nm at 25°C, but 218 nm negative ellipticity was induced at 92°C to a similar

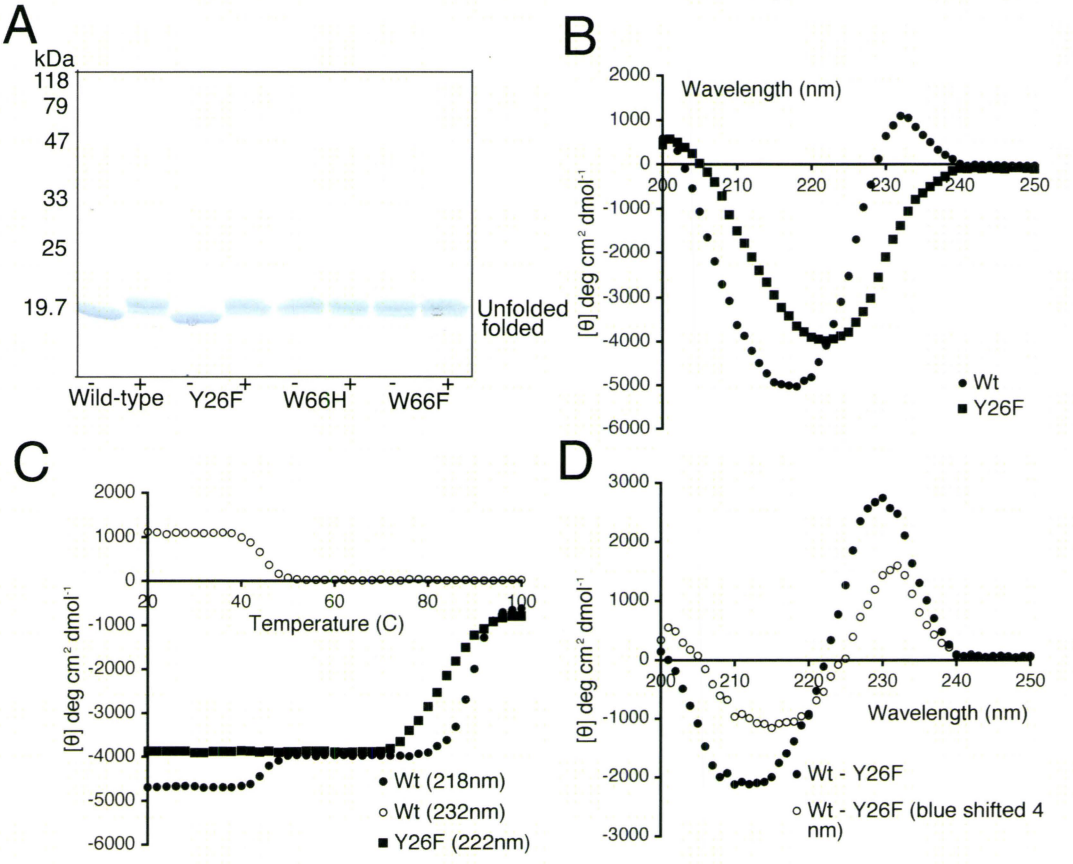

FIGURE 2.7 Refolding and spectroscopic analysis of PagP exciton mutants: (A) Wild-type PagP and Tyr26Phe, Trp66Phe and Trp66His mutants were subjected to refolding in LDAO and evaluated by SDS-PAGE. (B and C) Wild-type PagP and the Tyr26Phe mutant were further analyzed by far-UV CD spectroscopy (B), and the thermal unfolding profiles were obtained by following the loss of the peak ellipticity bands (C). (D) Difference far-UV CD spectra determined by subtracting the Tyr26Phe PagP spectrum, with and without introduced 4 nm blue shift, from the wild-type spectrum.

extent as observed previously for wtPagP at the same temperature (Figure 2.S1 C). These observations supports our suggestion that the residual ellipticity observed at high temperatures reflects intermolecular β structure formed prior to precipitation. By following the loss of ellipticity at 202 nm and the formation of ellipticity at 218 nm, we found the Trp66 mutants revealed that unfolded PagP in LDAO begins the transition to an aggregated state above 80°C (Figure 2.S1 D).

The Tyr26Phe mutant refolded successfully, and its far-UV CD spectrum did not exhibit any positive CD in the 230 nm region, thus confirming the predicted contribution of Tyr26 to the long-wavelength exciton band (Figure 2.7 B). However, the negative band was red-shifted from 218 to 222 nm, which indicates that the peptide backbone conformation is modulated by contributions from Tyr26 because loss of the exciton prior to β barrel unfolding at 80°C (Figure 2.5 C), or in the Gly88Cys mutant (Figure 2.4 A), was not accompanied by a red-shift. By following loss of 222 nm negative CD band in Tyr26Phe, we observed an 8°C reduction in the thermal transition temperature when compared to the 218 nm negative CD band lost from wtPagP; only the latter displays loss of the exciton in the 40-50°C range, and an apparent contribution of the Tyr26 to folding cooperativety is revealed by comparing the slopes of the two precipitation curves. Subtracting Tyr26Phe far-UV CD spectrum from that for wtPagP reveals the exciton couplet (Figure 2.7 D), but the peak-to-trough strength better matches the theoretical prediction (Figure 2.6 C) when the Tyr26Phe spectrum is first blue-shifted manually back to 218 nm to 222 nm.

Although the Trp66 mutants failed to refold in vitro, we could successfully purify folded Trp66F and Trp66H when they were expressed *in vivo* with an intact signal peptide (Figure 2.S2 A) (*70*). The *in vivo*-folded wt and Trp66 mutant PagP proteins appeared folded by both SDS-PAGE and far-UV CD criteria (Figure 2.S2 A, B). The far-UV CD spectra validate the predicted contribution of Trp66 to the nondegenerate exciton interaction by revealing the loss of positive ellipticity above the 230 nm and the formation of a weaker negative ellipticity maximum at 218 nm. In contrast to Tyr26Phe, no red shift was associated with the loss of the exciton in the Trp66 mutants. The predicted dominance of the Tyr26 and Trp66 contributions to the long and short wavelength exciton bands, respectively (Figure 2.S3 A, B), appears to be validated by the experimental observations (Figure 2.7 B, 10 B). The failure of the Trp66 mutants to fold *in vitro* is reflected in the instability associated with their reduced thermal unfolding temperature of 75°C compared with a temperature of 88°C observed for the *in vivo* folded wtPagP protein (Figure 2.S2 C). However, the Gly88Cys mutant was successfully folded *in vitro* despite that fact it displays a thermal unfolding temperature of only 65°C (Figure 2.4 C). Periplasmic chaperones are known to assist the folding of OMPs in *E. coli* and likely to help to overcome the folding defect associated with Trp66 mutations. The observation of reduced thermal unfolding temperatures for both the Tyr26 (Figure 2.7 C) and Trp66 (Figure 2.S2 C) mutants suggests that destabilizing effects will likely be associated with proximal groups that perturb these aromatic side chains.

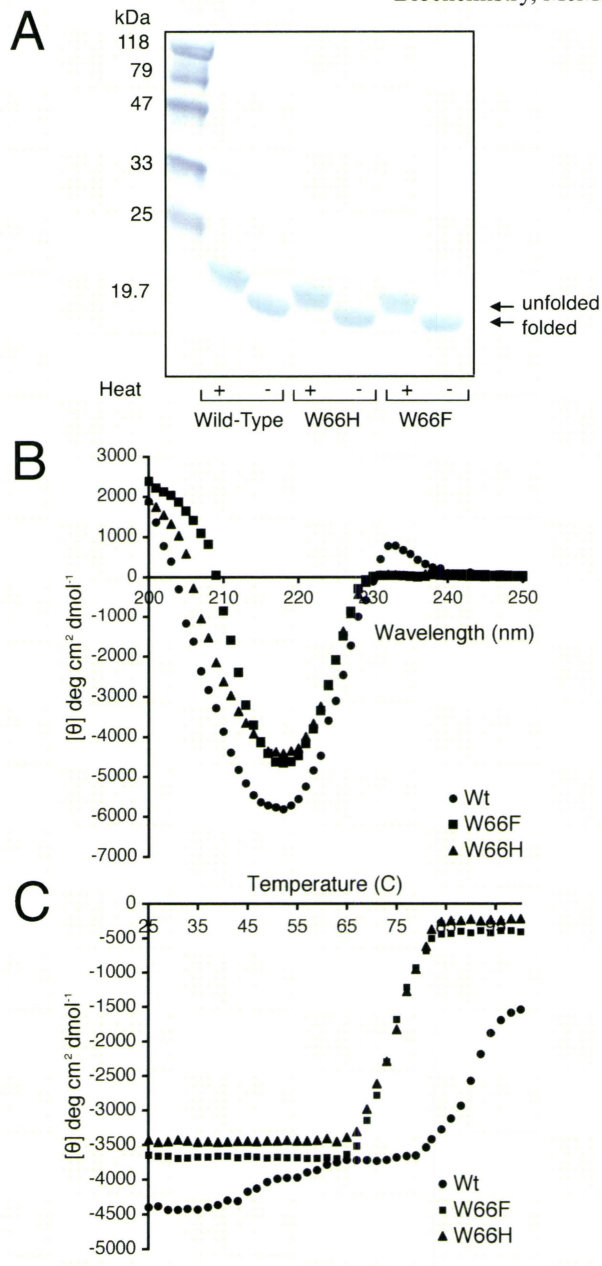

FIGURE 2.S2: Secondary structural analysis of PagP and its variants purified from membranes: Wild-type (Wt) is compared with Trp66His and Trp66Phe PagP mutants. Panel A shows SDS-PAGE, Panel B shows far-UV CD, and Panel C shows thermal precipitation curves following the loss of minimal ellipticity maximum at 218 nm.

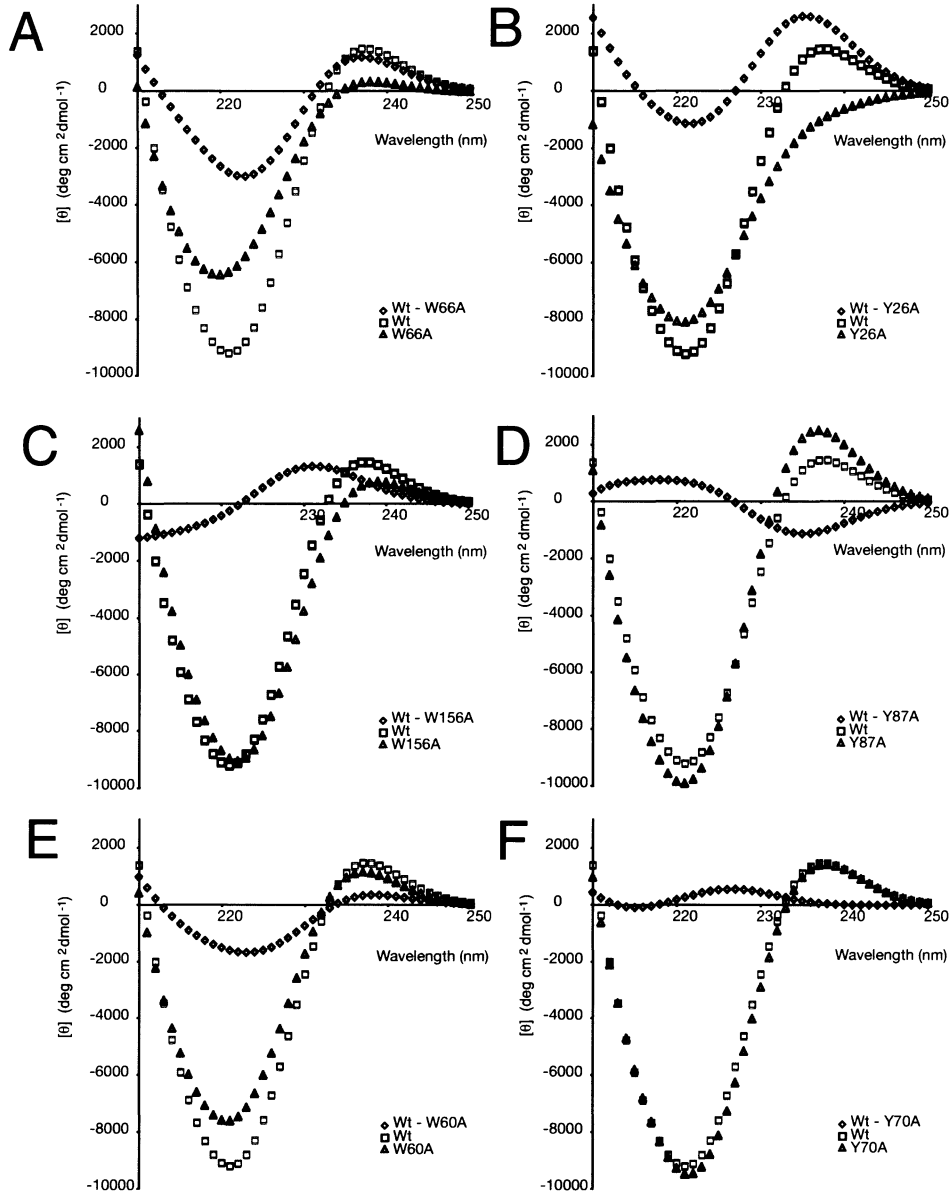

FIGURE 2.S3: Theoretical CD spectra of PagP, its *in silico* aromatic amino acid substitutions, and their difference spectra: Wild-type (Wt) is compared with W66A, Y26A, W156A, Y87A, W60A and Y70A PagP in panels A-F, respectively.

2.3.6 Hydrocarbon Ruler Mutants:

The remarkable thermal stability of PagP in detergent micelles provides the destabilized Gly88Cys mutant with sufficient residual stability to easily tolerate the temperature of 30°C encountered in the PagP hydrocarbon ruler assay. The assays are performed using a suite of synthetic diacylphosphatidylcholine donors that possess saturated acyl chains varying in length from 10 to 18 carbons in single-methylene unit increments (*111*). [^{32}P] Kdo_2-lipid A (Figure 2.1 A) is used as the acyl acceptor, and acylation is assessed by a TLC-based assay (Figure 2.8). All PagP derivatives that exhibit exciton couplets are able to select acyl chains of the length predicted by the hydrocarbon ruler hypothesis, and they do it with single-methylene unit resolution. Only Gly88Cys, which lacks the exciton because of a local structural perturbation arising from its single sulfhydryl group, could not accommodate a single acyl chain of the predicted length. Rather, Gly88Cys accommodated two acyl chains: one with the predicted length and one extended by an extra methylene unit.

2.3.7 *In Silico* Mutagenesis:

Efforts to crystallize the PagP mutant proteins have so far not met with any success. Therefore, we opted to perform *in silico* mutagenesis and assess the conformational preferences of the substituted side chains by constrained simulated annealing molecular dynamics. The main features surrounding the floor of the hydrocarbon ruler in the wtPagP crystal structure are outlined in figure 2.6 A and given a

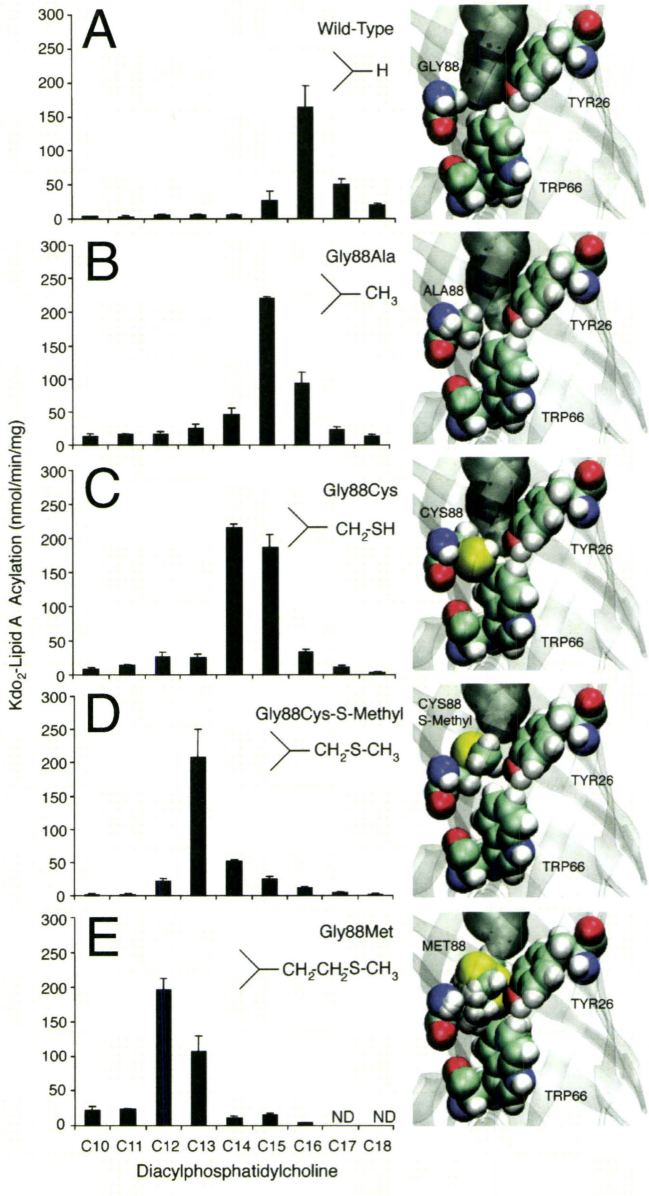

FIGURE 2.8 Hydrocarbon ruler analysis of refolded PagP variants: Synthetic diacylphosphatidylcholines with saturated acyl chains varying in single methylene increments from 10 to 18 carbon atoms were used as donors for the enzymatic acylation of Kdo_2-lipid A. Panels A-E present acyl-chain selection for the refolded PagP variants. The structure of PagP (PDB entry 1THQ) emphasizing the positions of Gly88 and the exciton interaction between Tyr26 and Trp66 is shown at the right of panel A. Contouring reveals the β-barrel interior region after removal of the LDAO detergent molecule. In silico mutagenesis followed by simulated annealing was performed to reveal a predicted orientation of the PagP Gly88 substitutions, which are shown at the right of panels B-E. ND means not detected

different perspective in the right portion of figure 2.8 A. Here the LDAO molecule that occupies the hydrocarbon ruler has been removed, and the vacated space is outlined with a contoured shell. The position of Gly88 reveals the *pro*-L hydrogen is directed toward the base of the hydrocarbon ruler and should direct and substituted amino acid side chain into the base of the cavity. Flanking this site are Tyr26 and Trp66, which approach each other in a geometric relationship that affords an exciton delocalization of the two chromophore-excited states. The amino acid substitutions associated with expression of both the exciton couplet and single-methylene unit resolution were found to advance into the hydrocarbon ruler floor (Figure 2.8 B, D, E), but the Cys sulfhydryl group was found to recede downward and induce a 1 Å shift in the position of the indole ring of Trp66, which alters its geometric relationship with respect to Tyr26 (Figure 2.8 C). Another possible source of perturbation of the local structure of the enzyme is the presence of a weak electrostatic interaction between the sulfhydryl group and the hydroxyl group of Tyr26, which is 4 Å away. These predictions suggest that the loss of the exciton in Gly88Cys results from a local perturbation of Trp66 and /or Tyr26 arising from the downward orientation of the Cys sulfhydryl group, which also serves to expand the hydrocarbon ruler and allow the enzyme to accommodate an additional acyl chain with one extra methylene unit. We propose that the Gly88Cys hydrocarbon ruler is tailored to optimally accommodate an acyl chain that lies somewhere between C14 and C15 and is thus capable of using both.

The thioether (sulphide) bond introduced into the Gly88Cys-*S*-methyl PagP is a decidedly hydrophobic entity (*164*) that is expected to occupy the hydrophobic pocket as predicted by the hydrocarbon ruler hypothesis. *In vitro*, this mutant selects the expected 13-carbon saturated acyl chain with single-methylene unit resolution (Figure 2.8 D), and it exhibits a full exciton couplet (Figure 2.4 A). Although the simulated annealing calculations fall short of exhaustively sampling conformational space, the annealed conformations of the Gly88Met suggest that this side chain spans a wider spatial distribution that the other mutants of Gly88 (Figure 2.8 A-E). This observation is consistent with the fact that Gly88Met, while preferentially selective for the expected 12-carbon saturated acyl chain, also exhibits residual ability to bind an acyl chain with one extra methylene unit. These observations might reflect a positional preference within the hydrocarbon ruler environment for the inherent conformational flexibility around the thioether CH_3S-CH_2 torsional unit and/or dispersion forces associated with the thioether sulphur atom (*164*).

Despite our experimental observations that the exciton is retained in all mutants except Gly88Cys, exciton calculations using structures from simulated annealing did not predict its presence in any of them. Even for wtPagP, the simulated annealing results are not predicted to have an exciton. This finding suggests that the X-ray conformation is a better representation of the conformational ensemble of the protein than the locally relaxed conformations obtained from our simulated annealing calculations. Obtaining more realistic conformations of the mutant would require letting the enzyme relax and

reach thermal equilibrium in explicit membrane and solvent. Nevertheless, the high-temperature simulated annealing calculations are a first step in modeling the local conformational preferences of the mutant side chains. These preferences are qualitatively different for the Cys mutant, which is consistent with the measured extinction of the exciton.

2.4 Discussion:

Excitons in proteins can arise from through-space interactions between the excited states of neighbouring Trp and/or Tyr aromatic chromophores. Although such excitons are by no means present in every protein structure, they are more common than is widely appreciated (*147*). Exciton couplets observed in some proteins have been previously employed to analyze protein structure (*165, 166*) but efforts to identify the residues actually responsible for the exciton interaction, and to employ them as instrinsic probes for monitoring structure-function relationships, have been rare. An exciton couplet formed by Trp172 and Trp215 in chymotrypsinogen and chymotrypsin was used to demonstrate conversion of the active enzyme conformation to the zymogen conformation at high pH and to show that the diisopropyl fluorophosphate-inactivated enzyme does not undergo this transition (*167*). However, the excitonic origin of the far-UV features in the CD spectra of chymotrypsinogen and chymotrypsin was not demonstrated until much later (*147*). Exciton coupling between Trp47 and Trp74 in *E. coli* dihdrofolate reductase was identified as the source of the first transient that can be detected by stopped flow in

the refolding of the urea-denatured enzyme (*168*). More recently, a Gly95Ala mutation has been shown to abolish the Trp-Trp exciton, whereas binding of the methotrexate restores it (*169*). Given the forecasted output of current structural proteomics consortia, it is likely that the pool of proteins exhibiting clearly discernible exciton couplets will continue to grow rapidly into the future. Such knowledge will provide investigators with highly sensitive probes for monitoring the functional effects of locally introduced structural alterations.

A case in point is revealed by our current study of PagP, which has a number of favourable features that made the study possible. (1) The compact β barrel structure of PagP produces a relatively weak backbone CD that makes obvious the overlapping exciton couplet. (2) PagP can resolve methylene units in the acyl chains of its lipid substrate by using a molecular measuring device known as a hydrocarbon ruler, the floor of which is located in the proximity of the interacting exciton partners. (3) PagP is an extremely stable small protein that can be unfolded and refolded *in vitro*, enabling the transient exposure of buried functional groups for chemical modification. The fact that PagP is an integral membrane protein makes these favourable features all the more remarkable. Consequently, PagP beautifully demonstrates both the sensitivity of the exciton and how it can be employed as a spectroscopic probe to monitor a local structural perturbation.

The PagP hydrocarbon ruler was identified in a prior X-ray structure determination (*111*) and validated in that study by using enzyme that had been purified directly from membranes (*70*). However, the LDAO detergent used to purify the enzyme is a competitive inhibitor (*111*) and must be exchanged by dilution of the enzyme into a reaction cocktail that includes DDM or another detergent that supports PagP activity (*70*). While PagP purified directly from membranes is obtained at ~1 mg/ml (*70*), PagP can also be expressed without its N-terminal signal peptide, purified in an unfolded state, and reliably refolded in LDAO to reach a concentration of >10 mg/ml (*110*) (*111*). Consequently, LDAO can be removed more effectively by dilution when using refolded PagP, which gives more reliable results in the enzyme assays. By using refolded PagP derivatives in this investigation, we observed for the first time that Gly88 substitutions, with the exception of the Cys mutation, modulate hydrocarbon ruler acyl-chain selection with methylene-unit precision.

The predicted interactions of the Cys sulfhydryl group with the Trp66 and Tyr26 aromatic side chains can explain our main experimental observations that Gly88Cys is both devoid of the exciton and compromised in terms of acyl-chain selection. Our simulated annealing experiments demonstrate that the amino acid substitutions observed experimentally to retain the exciton are advancing upward into the hydrocarbon ruler floor, while the one side chain that is able to recede downward (Cys) is observed to extinguish the exciton. The question of why the Cys sulfhydryl group would prefer the position between Tyr26 and Trp66 instead of inside the hydrophobic hydrocarbon ruler

pocket remains. The hydrocarbon ruler floor in PagP marks the boundary between the upper hydrophobic and lower polar interior regions of the β barrel (*111*), which might indicate that the Cys sulfhydryl group has an inherent tendency to seek out a more polar microenvironment. Hydrophobicity scales derived primarily from α-helical transmemebrane domains typically identify Cys as being hydrophobic (*170*), but Cys in an α-helix normally donates a hydrogen bond to the carbonyl oxygen atom of the preceding peptide bond (*171*). In contrast, the sulfhydryl group of the Gly88Cys is forced into a β structured conformation where it is faced above with nonpolar and below with polar environments. In the absence of a suitable hydrogen bond acceptor, the Cys sulfhydryl group appears to have rotated about its χ1 dihedral angle to establish a weak electrostatic interaction with the Tyr26 hydroxyl group and/or a so-called π (aromatic)-lone pair (sulphur) attraction with Trp66 (*172*). The polarizability of the sulphur atom and the acidity of the sulfhydryl group are reported to promote its interaction with aromatic partners. Although our force field contains only partial point charges on each nucleus, which is but a rough approximation to the electronic distribution associated with (pure and hybrid) π orbitals, the force field may be capable of capturing these effects qualitatively.

The lactose permease provides an example of an α-helical transmembrane protein in which substitutions of Cys for Gly have had unpredictable effects on protein stability. The replacement of Cys154 with Gly within a transmembrane domain of the lactose permease has been shown to have stabilizing effects that are incompatible with active

transport but necessary to obtain crystals for determining the structure by X-ray diffraction (*173, 174*). Only Cys at position 154 was capable of introducing a local instability that was necessary for normal transport (*175*). The lactose permease in this region includes local distortions of helical conformation and close packing interactions with several aromatic partners. Perhaps the unusual physical properties associated with Cys154 in lactose permease have in common with Gly88Cys the need to establish interactions with aromatic side chains. Clearly, the PagP hydrocarbon ruler affords unique opportunities to evaluate the physical forces at work in determining membrane protein stability. Our ability to enzymatically synthesize endotoxin analogues with defined acyl-chain substitutions might also reveal pharmacologically interesting derivatives, because acyl-chain length is known to be an important factor in determining endotoxin structure-activity relationships

3.0 A Thiolate Anion Buried Within The Hydrocarbon Ruler Perturbs PagP Lipid Acyl Chain Selection

Work presented in this chapter was accepted in the peer-reviewed journal *Biochemistry*, on Feb 22, 2010: Khan, M.A., Moktar, J., Mott, P. J., Bishop, R.E. "A thiolate anion buried within the hydrocarbon ruler perturbs PagP lipid acyl chain selection". *Biochemistry* (DOI: 10.1021/bi 9016699). Joel Moktar, a fourth year undergraduate research student, assisted in the construction, purification and folding of the mutants described in this manuscript. Patrick Mott, also a fourth year undergraduate research student, assisted in the *in vivo* hydrocarbon ruler experiments

3.1 Introduction:

The outer membrane acyltransferase PagP catalyzes the transfer of a palmitate chain from the *sn*-1 position of a glycerophospholipid to the free hydroxyl group of the *R*-3-hydroxymyristate chain at position 2 of lipid A (endotoxin) (Figure 3.1) (*28, 70, 176*). In *Escherichia coli* and related pathogenic Gram-negative bacteria, PagP combats host immune defences by restoring the bacterial outer membrane permeability barrier (*114, 115, 118, 150*), and attenuating endotoxin signalling through the host TLR4/MD2 inflammation pathway (*7, 65, 177, 178*). The *pagP* gene is transcriptionally activated through the bacterial PhoP/PhoQ virulence signal transduction network in response to membrane-perturbing antimicrobial peptides (*66*), immunological agents PagP helps the bacteria to resist (*113, 115, 118*). PagP thus provides a target to develop anti-infective agents, but it also provides a tool to synthesize novel vaccine adjuvants and endotoxin antagonists (*176*).

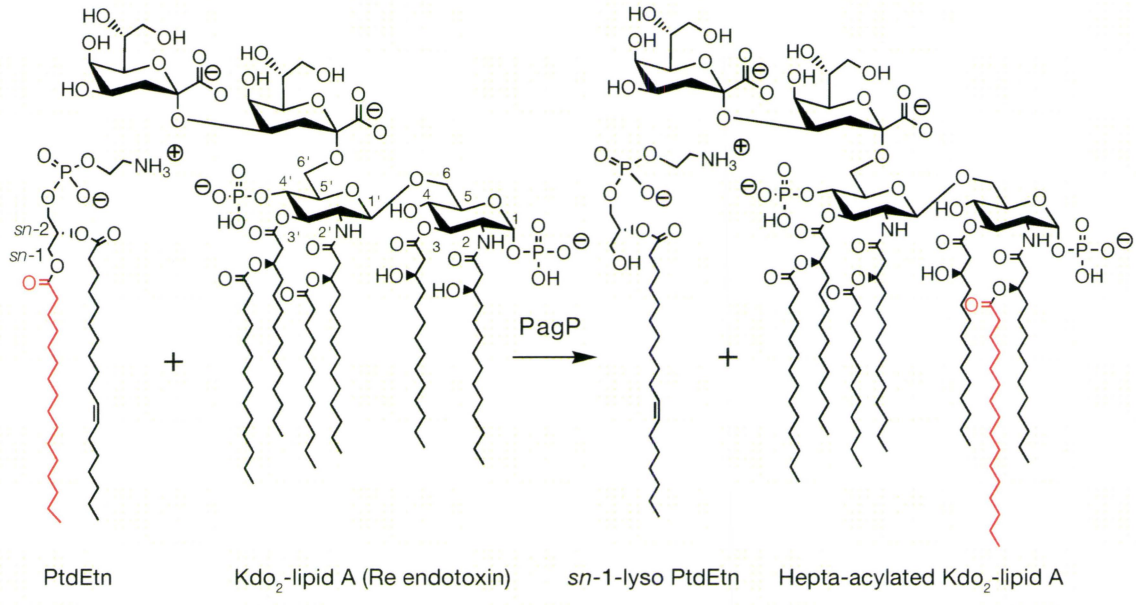

Figure 3.1 Reaction catalyzed by the phospholipid::lipid A palmitoyltransferase PagP: *Escherichia coli* PagP transfers a palmitoyl group from the *sn*-1 position of a phospholipid, such as phosphatidylethanolamine (PtdEtn), to the free hydroxyl-group of the *N*-linked *R*-3-hydroxymyristate chain on the proximal glucosamine unit of lipid A. One of the simplest lipid A acceptors for PagP in the outer membrane is known as Kdo_2-lipid A or Re endotoxin.

Despite the widespread importance of integral membrane enzymes of lipid metabolism in signal transduction and membrane biogenesis processes (*179, 180*), their molecular mechanisms for lipid substrate interrogation remain poorly understood because many such enzymes have proven recalcitrant toward functional detergent extraction from the membrane environment. In contrast, PagP is a heat-stable 161-amino acid β-barrel enzyme, which intrinsically lacks Cys residues, and can reversibly unfold and refold in a defined detergent miceller enzymatic assay system (*126*). The structure of *E. coli* PagP has been solved both by solution nuclear magnetic resonance spectroscopy and by X-ray crystallography to reveal an eight-stranded antiparallel β-barrel preceded by an *N*-terminal amphipathic α-helix (*110, 111*). Like other β-barrel membrane proteins, the center of the PagP β-barrel is relatively rigid with more flexible external loops where the active site residues are localized. PagP alternates between two dynamically distinct states likely representing dormant and active conformations in the outer membrane environment (*123, 125, 181*). PagP activity is triggered by perturbations to outer membrane lipid asymmetry that allow the phospholipid and lipid A substrates to access the active site from the external leaflet by a lateral lipid diffusion mechanism through two gateways in the β-barrel wall known as the crenel and embrasure, respectively (*182*). X-ray studies of the inhibited PagP enzyme identified a single detergent molecule buried within the rigid β-barrel core, a site that corresponds with the palmitoyl group binding pocket known as the hydrocarbon ruler (Figure 3.2A) (*111*).

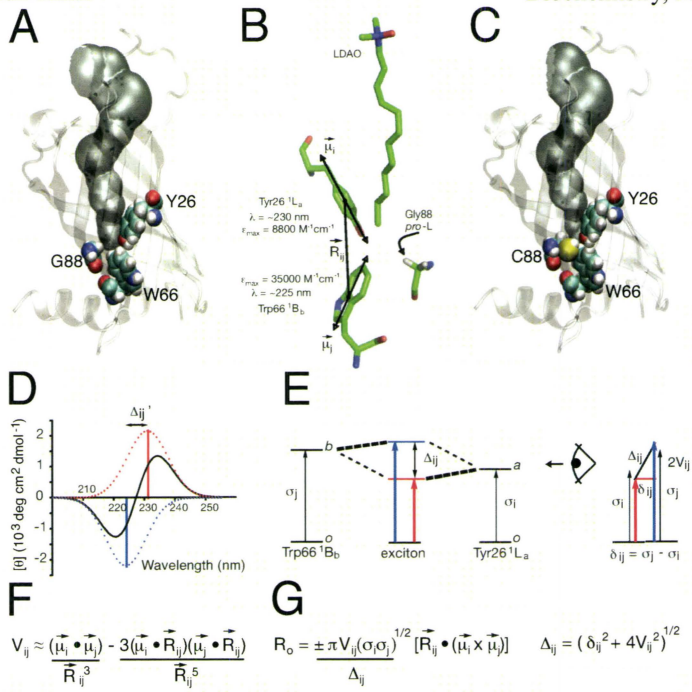

Figure 3.2 Structural relationships between the PagP hydrocarbon ruler and the aromatic exciton. A) Structure of PagP emphasizing the hydrocarbon ruler floor residues G88, Y26, and W66. Contouring identifies the hydrocarbon ruler after the bound LDAO molecule has been removed from the crystal structure (PDB: 1THQ). B) The polarization axes for the interacting Tyr26 ^1L_a (μi) and Trp66 ^1B_b (μj) electric dipole transition moment vectors are shown to emphasize their relationships with the interchromophore distance vector R_{ij}. C) A structural perturbation hypothesis is proposed to explain the properties of Gly88Cys PagP, where a rotation of the Cys side chain perturbs the aromatic exciton partners. D) The PagP exciton couplet (solid curve) is the resultant of two Cotton effects (red and blue dotted curves) having equal magnitude, opposite sign, and separated by the exciton splitting energy Δ_{ij}, denoted as Δ_{ij}' on the wavelength scale. E) The $\pi \rightarrow \pi^*$ transitions from ground states o to excited states *a* (for Tyr26 ^1L_a) or *b* (for Trp66 ^1B_b) are split by a non-degenerate exciton interaction into two new exciton states. Tyr26 and Trp66 each dominate the low (red) and high-energy (blue) exciton states, respectively. The splitting energy is related to the interaction energy V_{ij} through an orthogonal relationship with the difference in the transition energies of the two chromophores $\delta_{ij} = (\sigma_j - \sigma_i)$. F) V_{ij} can be given by the point-dipole approximation. G) The theoretical rotational strengths (R_o) of non-degenerate exciton split CD Cotton effects are proportional to V_{ij}, Δ_{ij}, and a scalar triple product involving the vector connecting the centers of the two transitions and the dipole transition moments. Δ_{ij} is given by the Pythagorean theorem and reduces to $2V_{ij}$ when identical chromophores interact in the degenerate exciton condition where $\sigma_j = \sigma_i$, which in combination reduce the constant term in the equation for R_o to $\pm\pi\sigma/2$.

PagP is exquisitely selective for a 16-carbon palmitate chain because its hydrocarbon ruler excludes lipid acyl chains differing in length by a solitary methylene unit (*111, 126*). Mutation of Gly88 lining the hydrocarbon ruler floor can modulate lipid acyl chain selection (*126*). Appropriate amino acid substitutions shorten the selected acyl chain by a degree predictable from the expected rise in the hydrocarbon ruler floor. The architecture of the acyl chain-binding pocket can be monitored by a spectroscopic signature in circular dichroism (CD) known as an aromatic exciton couplet (*126, 147*). The exciton is a consequence of excited state delocalization between Tyr26 and Trp66 aromatic $\pi \rightarrow \pi^*$ transitions, which occurs when these residues adopt the precise geometric arrangement found near the hydrocarbon ruler floor in folded wild-type PagP (Figure 3.2B). The exciton couplet manifests a positive ellipticity maximum at 232 nm together with an equivalent negative ellipticity maximum at 218 nm (Figure 3.2D). This non-degenerate exciton couplet is the resultant of two Cotton effects where the rotational strength (R_o) and splitting energy (Δ_{ij}) each depend on an orthogonal relationship between the interaction energy (V_{ij}) and the energy difference between the two chromophore electric dipole transition moments (δ_{ij}) (Figure 2E-G) (*140*). The 218 nm signal in the PagP far-ultraviolet (far-UV) CD spectrum is the further resultant of combining the negative exciton band with the stronger Cotton effect for the polypeptide $n \rightarrow \pi^*$ transition, which also arises at 218 nm due to the β-barrel structure (*126*). The aromatic exciton thus provides an intrinsic spectroscopic probe, which was first detected by its

disappearance in response to a local structural perturbation associated with PagP hydrocarbon ruler manipulation.

In our prior investigations of PagP hydrocarbon ruler mutants where Gly88Ala, Gly88Cys-*S*-methyl and Gly88Met were constructed, acyl chain selection was predictably shifted toward C15, C13 and C12 acyl chains, respectively, and with the expected unitary methylene resolution. Furthermore, these mutants displayed the exciton and the thermal melting temperature (T_m) of 88°C characteristic of wild-type PagP (*126*). We had anticipated the Gly88Cys mutant would behave as a dedicated myristoyltransferase, but the acyl chain resolution was compromised because it selected both C14 and C15 acyl chains. Additionally, the Gly88Cys enzyme failed to display the exciton and had its T_m decreased to 67°C (*126*). Since Cys residues are usually accommodated easily in a membrane environment (*170*), our findings suggested the sulfhydryl group in Gly88Cys PagP might be surprisingly too polar for its surrounding hydrophobic milieu. Instead of projecting up from the hydrocarbon ruler floor, we reasoned the sulfhydryl group might be stabilized by rotating around its χ_1 dihedral angle to interact with the aromatic exciton partners (*126*). This rotation would simultaneously disrupt the geometric requirements to establish the exciton and partially lower the hydrocarbon ruler floor to explain the observed selection of both C14 and C15 acyl chains (Figure 3.2C).

If the Cys sulfhydryl group is too polar for the hydrocarbon ruler microenvironment, we might predict substitution of Gly88 with Ser or Thr, to introduce a decidedly more polar hydroxyl group (*170*), would elicit a destabilizing effect at least comparable to that observed for Gly88Cys PagP. An alternative hypothesis is that the dissociation constant for ionization of the Cys88 thiol is lowered in the hydrophobic interior of PagP, resulting in the burying of a negatively charged thiolate anion. Initially, this seemed unlikely as the pKa for a Cys thiol generally lies well above pH 8.0 (*183*), which is the optimal pH for PagP activity and is typically used in our PagP folding procedure. We fold PagP into detergent solutions from a denatured state in guanidine-HCl (Gdn-HCl), where the pKa of the Cys thiol has previously been reported to be 9.4 (*162*). We have recently substituted key Pro residues at two flanking sites in the PagP transmembrane domain with Cys to establish a lateral diffusion mechanism for gating lipids between the membrane and the hydrocarbon ruler (*182*). Cys oxidation to form disulfide bonds in this study required the membrane permeable oxidizing reagent copper phenanthroline because the hydrophobic environment appears to stabilize the thiols and thereby raises their pKa. Assuming that the Cys located in the floor of the hydrocarbon ruler similarly experiences a hydrophobic environment, we did not anticipate that the thiol pKa would instead be suppressed. However, the pKa for the Cys thiol in folded proteins is among the most variable of ionizable groups (*184*), indicative of a high sensitivity to its surrounding microenvironment. The hypothesis that a buried thiolate anion is the source of the compromised acyl chain selectivity in Gly88Cys PagP could be

tested if the pH optimum for enzymatic activity is sufficiently broad across the actual Cys pKa.

Here we report that Gly88Ser and Gly88Thr PagP enzymes display the aromatic exciton and thermal stability characteristics of wild-type PagP, and function predictably as dedicated myristoyltransferases. However, folding Gly88Cys PagP from Gdn-HCl into detergent was found to lower the thiol pKa in support of a model for electrostatic perturbation from a buried thiolate anion. This thiolate can be neutralized by an appropriate pH adjustment to resolve the structural perturbation and restore methylene-unit resolution. Hydrocarbon ruler-exciton coupling in Gly88Cys PagP thus reveals how lipid acyl chain selection can be modulated through a thiol-thiolate ionization mechanism.

3.2 Experimental Procedures:

3.2.1 DNA Manipulations and Purification of Protein and Lipid A:

PagP mutants were constructed with the primer sets listed in Table 2, using the pETCrcAHΔS vector, which incorporates a *C*-terminal His-6 tag and lacks the *N*-terminal signal peptide to allow for expression as insoluble aggregates in *E. coli* BL21(DE3) as described (*110*). Plasmids were purified using the QIAprep spin miniprep kit (Qiagen). Site-directed mutagenesis reactions were conducted using the Quickchange protocol (Stratagene). To confirm the absence of any spurious mutations in the plasmid, sequencing from the T7 primers (Table 2) was performed.

To express the native *pagP* gene in bacterial cells under control of its endogenous promoter, the wild-type copy in plasmid pACPagP (*123*) was transferred to pBad18 (*185*) by restriction enzyme digestion with *Sal*I and *Hind*III followed by ligation and transformation using standard molecular biology procedures (*186*). Designated mutations were then constructed as described above, but with sequencing performed using the pBAD primers (Table 2). Native PagP and its mutant derivatives were expressed in *E. coli* NR754Δ*pagP*, an *araD*$^{+}$ revertant of *E. coli* MC4100 (*40, 187*), into which the Δ*pagP::kan* allele was moved from the Keio collection (*188*) by P1 transduction (*189*) followed by excision of the *kan* cartridge (*190*).

Cells were grown in LB medium at 37°C with 100 μg/mL or tetracycline at 12.5 μg/mL as appropriate (*186*). PagP was purified from cytoplasmic inclusion bodies according to an established procedure (*126*). In order to purify radiolabelled lipid A bearing two units of 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo_2-lipid A), a 10 mL culture of *E. coli* WBB06 was grown on [^{32}P] orthophosphate as described (*155*). Lipid A acylation was evaluated *in vivo* by [^{32}P] orthophosphate labelling with lipopolysaccharide isolation followed by mild-acid hydrolysis of the ketosidic Kdo-lipid A linkage for analysis by thin layer chromatography (TLC) as described (*191*). Cells were treated with ethylenediaminetetraacetic (EDTA) to stimulate PagP activity as described (*123*). Non-radioactive 250 mL cultures were grown to A_{600} = 0.8 and treated with EDTA prior to harvesting for lipid A isolation as described (*192*). The dried non-radioactive lipid A samples were dissolved in 2:3:1 (chloroform, methanol, water) and

stored at -20°C for no more than 12 hours before the samples were analyzed by electrospray ionization mass spectrometry (ESI-MS) in the negative ion mode. Modelling of PagP side-chain conformations at position 88 was performed as described (*126*).

3.2.2 Mass Determination and Protein Folding:

Precipitated protein samples were dissolved in a solution of 1:1 1% formic acid/acetonitrile just prior to ESI-MS (*124, 126*). The sample concentration was maintained at 1 ng/μL and was injected directly onto a Waters/Micromass Q-TOF Ultima Global (a quadrupole time-of-flight) mass spectrometer. The spectra were reconstructed using MassLynx 4.0 with the MaxEnt 1 module. ESI-MS performed on PagP and its derivatives (*124, 126*) yields experimental protein masses accurately matching theoretical predictions (Table 4).

After mass determination, the remainder of the precipitated protein fractions (~50 mg) were solubilized in 5 mL of 10 mM Tris-HCl (pH 8.0), 6 M Gdn-HCl and diluted dropwise (~1 drop per 2 seconds) into a stirring solution of a 10 fold molar excess of 10 mM Tris-HCl (pH 8.0), 0.5% lauryldimethylamine-*N*-oxide (LDAO) at room temperature and then left to stir overnight at 4°C. The sulfhydryl group of Gly88Cys PagP was maintained in its reduced state by addition of 20 mM β-mercaptoethanol. The solution was then applied to a 4 mL Ni^{+2}-ion column resin bed charged with 50 mM $NiSO_4$ and equilibrated with 10 mM Tris-HCl (pH 8.0), 0.1% LDAO, 5 mM imidazole. The column

was washed with 10 column volumes of the equilibration buffer and 10 column volumes of 10 mM Tris-HCl (pH 8.0), 0.1% LDAO, 20 mM imidazole. The proteins were eluted with 2 mL of 10 mM Tris-HCl (pH 8.0), 0.1% LDAO, and 250 mM imidazole. The samples were then dialyzed against 10 mM Tris-HCl (pH 8.0), 0.1% LDAO. The protein concentration was determined using the bicinchoninic acid assay kit (Pierce) (*151*) or the Edelhoch method (*152*). The folded samples were resolved by 15% sodium dodecyl sulfate polyacrylamide gel electrophoresis with heated and unheated samples in adjacent lanes. The heated samples were boiled at 100°C for 10 min prior to loading to distinguish the folded β-forms from the heat denatured α-forms as described previously (*126*). Folded PagP proteins were stored at -20°C. Gly88Cys PagP was also folded as above, but including different pH values of 5.0, 6.0, 7.0, 7.2, 7.4, 7.6, 7.8, 8.0 and 9.0. Tris was used in all instances, but the pH remained stable for all samples.

3.2.3 Spectroscopic Analysis:

Samples to be analyzed by far-UV CD were maintained at a concentration of 0.3 mg/mL in 10 mM Tris-HCl (pH 8.0), 0.1% LDAO. A cuvette with a path length of 1 mm was used in an Aviv 215 spectrophotometer, which was linked to a Merlin Series M25 Peltier device for temperature control. For each sample 3 accumulations were averaged at a data pitch of 1 nm and a scanning speed of 10 nm/min. The temperature was maintained at 25°C and data sets were obtained from 200 nm to 260 nm. Thermal denaturation profiles were obtained by measuring the loss of minimal ellipticity of the

sample. The proteins were heated from 20°C to 100°C, unless indicated otherwise, at 218 nm with a temperature slope of 2°C/min and a response time of 16 seconds.

A similar analysis was performed for Gly88Cys PagP folded at different pH values. In one experiment, far-UV CD and a thermal profile at 218 nm were compared for samples at pH 5.0, 6.0, 7.0, 8.0 and 9.0. In a second experiment, identical analyses were conducted for Gly88Cys folded at pH 7.0, 7.2, 7.4, 7.6, 7.8 and 8.0. In another experiment, Gly88Cys PagP folded at pH 7.8 or 7.2 was diluted in 10 mM Tris-HCl, 0.1% LDAO buffered at pH 7.2 or 7.8, respectively, before far-UV CD scans were obtained at 25°C, after which the samples were heated to 76°C and then cooled back to 25°C. A second far-UV CD scan was acquired at this stage. A final thermal profile was then generated by heating the samples to 100°C. The Gly88Cys PagP proteins folded at pH 7.0 or pH 8.0 were also diluted in 10 mM Tris-HCl, 0.25% *n*-dodecyl-β-D-maltoside (DDM) buffered at pH 8.0 or 7.0, respectively. Far-UV CD scans from 220 nm to 240 nm were conducted at 5 min intervals for 1 hr. The data were reported by comparing changes is the magnitude of the positive peak at 232 nm.

3.2.4 Acyltransferase Assays:

In a 0.5 mL microcentrifuge tube, sufficient Kdo_2-lipid A was added to achieve a concentration of 10 μM in a final assay volume of 25 μL. A trace amount of [^{32}P]-Kdo_2-lipid A was then added in order to achieve 200 cpm/μL. Sufficient phosphatidylcholine (PtdCho) of defined acyl chain composition was added to attain a concentration of 1 mM. Diacyl PtdCho's contained identical saturated acyl chains varying in length from C10 to

C17 at both the *sn*-1 and *sn*-2 positions, whereas mixed acyl chain PtdCho's contained C14 and C16 acyl chains at the *sn*-1 and *sn*-2 positions, respectively, or *vice versa*. These constituents were dried under a gentle N_2 stream and subsequently dissolved in 22.5 μL of a reaction cocktail containing 0.1 M Tris-HCl (pH 8.0), 10 mM EDTA and 0.25% DDM. The reaction was initiated by adding 2.5 μL of PagP at a sufficient concentration to achieve a linear reaction profile. All reactions were carried out at 30°C and were stopped by directly spotting 4 μL of the reaction mixture to the origin of a Silica Gel 60 TLC plate. The TLC plate was resolved in $CHCl_3$/pyridine/88% formic acid/H_2O (50:50:16:5; v/v) within a tightly sealed glass tank. The constituents of the tank were allowed to equilibrate for a period of three hours prior to exposure of the TLC plate. After the plate was dried, it was exposed to a Molecular Dynamics PhosphorImager screen overnight to visualize the reaction products, which were quantified using ImageQuant software. Wild-type PagP folded in 0.1% LDAO (pH 8.0) was also diluted in a reaction cocktail of 100 mM Tris-HCl, 0.25% DDM and 10 mM EDTA at a pH range of 5, 6, 7, 8, 9 and 10. Similarly, both wild-type and Gly88Cys PagP were folded at pH 7 and 8 and assayed at the same values to determine the effects of pH upon acyl chain selectivity.

3.3 Results and Discussion:

The PagP hydrocarbon ruler floor was adjusted by creating Gly88 substitutions using site-directed mutagenesis with the primer sets shown in Table 2, and the isopropyl-

β-D-thiogalactopyranoside-inducible PagP expression plasmid pETCrcAHΔS (*110*). The expressed proteins possess a *C*-terminal hexahistidine tag and lack the *N*-terminal signal peptide to target PagP for secretion to the outer membrane. The proteins were expressed as insoluble aggregates, which could be dissolved in Gdn-HCl, purified by Ni^{2+}-ion affinity chromatography, and folded by dilution into the detergent LDAO (*126*). Mutations were confirmed both by DNA sequencing of the mutant plasmids and by ESI-MS of the purified proteins (Table 4).

3.3.1 Gly88Ser and Gly88Thr PagP Mutants are Dedicated Myristoyltransferases:

To test the hypothesis that the Gly88Cys PagP structural perturbation associated with exciton loss, 21°C reduction in apparent T_m, and compromised acyl chain selectivity is a consequence of sulfhydryl group polarity within the hydrophobic hydrocarbon ruler environment (*126*), we additionally created the polar Gly88Ser and Gly88Thr PagP substitutions. Far-UV CD analysis of the folded mutant proteins clearly identified the positive 232 nm ellipticity known to be derived from the intact aromatic exciton (Figure 3.3A). Additionally, the 218 nm component of the exciton was observed by its loss at temperatures above 40°C, consistent with the previously described thermal sensitivity of

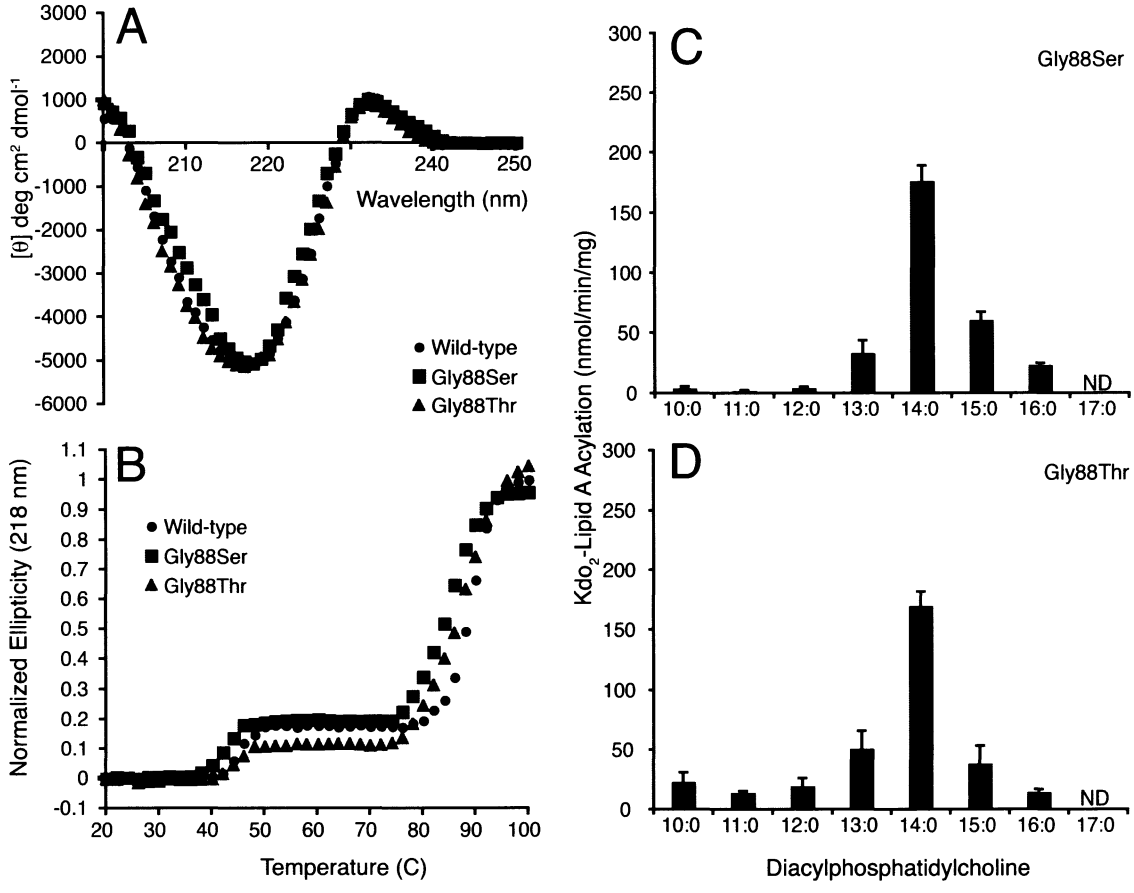

Figure 3.3 Gly88Ser/Thr PagP mutants reveal that polarity alone does not perturb acyl chain selection: A) Far-UV CD identifies the β-barrel and exciton signatures characteristic of wild-type PagP. B) Thermal melts at 218 nm reveal exciton loss preceding β-barrel unfolding. C) and D) Hydrocarbon ruler measurements reveal unitary acyl chain selectivity catalyzed by Gly88Ser and Gly88Thr PagP folded at pH 8. Synthetic diacyl-PtdCho's with saturated acyl-chains varying in single increments from 10 to 17 carbon atoms were used as donors for the enzymatic acylation of Kdo_2-lipid A. ND, not detected.

the exciton (*126*). The apparent T_m of 88°C for wild-type PagP β-barrel unfolding was reduced by no more than 5°C in the Gly88Ser and Gly88Thr mutants (Figure 3.3B).

PagP phospholipid:lipid A palmitoyltransferase activity was monitored *in vitro* using a defined detergent miceller enzymatic assay with TLC separation of radioactive lipid products (*126*). The inhibitory LDAO detergent used during PagP folding was exchanged by dilution into DDM to support enzymatic activity. Wild-type PagP is highly selective for a palmitoyl group at the *sn*-1-position in a glycerophospholipid, but largely unspecific for the polar head group (*70*). We employ the palmitoyl donor di-16:0-PtdCho and the palmitoyl acceptor [^{32}P]-Kdo_2-lipid A. When challenged with a spectrum of donor acyl chain lengths, wild-type PagP is a dedicated palmitoyltransferase whereas Gly88Cys PagP utilizes both myristoyl and pentadecanoyl groups (*126*), but we found that the Gly88Ser and Gly88Thr PagP enzymes each functioned as dedicated myristoyltransferases (Figure 3C,D). These findings indicate that Cys polarity cannot explain the extinguished exciton, reduced apparent T_m, and compromised acyl chain selectivity displayed by Gly88Cys PagP (*126*) because the Ser/Thr hydroxyl group, which is more polar than the Cys sulfhydryl group (*170*), appears to be easily tolerated within the hydrocarbon ruler environment.

3.3.2 Exciton-monitored Titration Reveals pKa Suppression of the Gly88Cys PagP Thiol:

To test the alternative hypothesis that the structural perturbation associated with the Gly88Cys substitution is a consequence of a buried thiolate anion, we folded Gly88Cys PagP at a range of values between pH 5 and 9 (Figure 3.4A,B). The absence of the exciton in Gly88Cys PagP at pH 8 was expected (*126*), but the mutant was surprisingly indistinguishable from wild-type PagP when folded at values below pH 8. The destabilization of Gly88Cys PagP at pH 8 and above was marked by a substantial 21°C drop in apparent T_m, but this drop was undetectable at pH 7 and below. In contrast, the Gly88Ser and Gly88Thr PagP mutants displayed a modest 5°C reduction in apparent T_m (Figure 3.3B), consistent with the greater polarity of the hydroxyl versus sulfhydryl group (*170*).

To identify the thiol pKa inflection point, we repeated this experiment by folding Gly88Cys PagP at smaller increments between values of pH 7 and pH 8 (Figure 3.4 C,D). By following the positive ellipticity at 232 nm (Figure 3.4E) or the apparent T_m values (Figure 3.4F), we could identify a sharp midpoint in both cases at pH 7.5, thus identifying the pKa for Cys thiol ionization. These findings indicate that the Gly88Cys PagP sulfhydryl group is easily accommodated within the hydrocarbon ruler, whereas the buried thiolate anion is highly destabilizing.

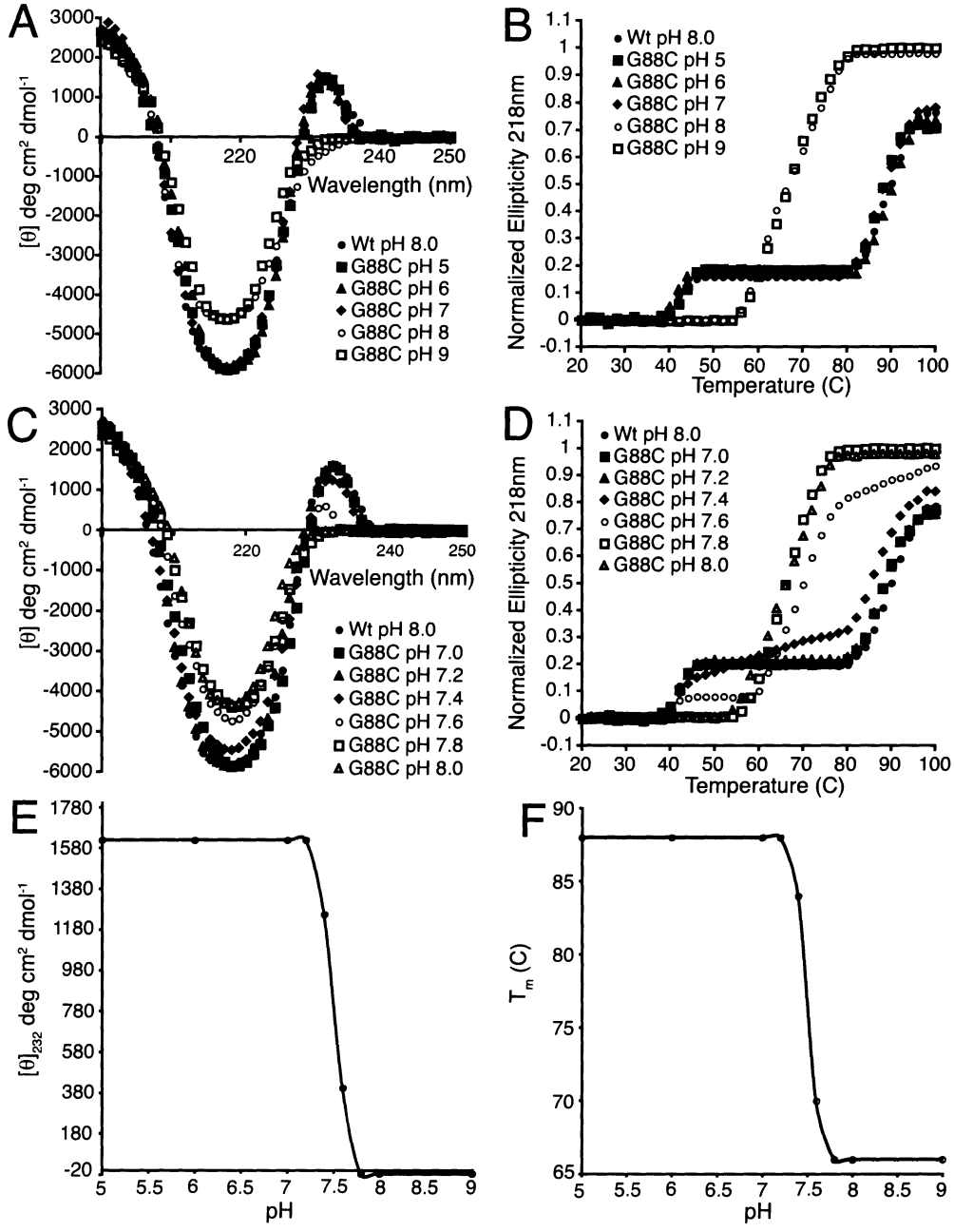

Figure 3.4 Exciton-monitored titration of the Gly88Cys PagP thiol reveals a suppressed pKa: Far-UV CD and thermal melts for wild-type PagP folded at pH 8 and for Gly88Cys PagP at a range of pH values. A) and C) Far-UV CD wavelength scans are shown. B) and D) Thermal melts at 218 nm are shown. E) pH dependence of the molar ellipticity at 232 nm. F) pH dependence of the apparent T_m values.

3.3.3 Neutralizing the Thiolate Anion by Protonation in LDAO requires PagP Thermal Unfolding:

The inhibitory LDAO molecule bound in the hydrocarbon ruler presumably shields the thiol group of Gly88Cys PagP from exposure to bulk solvent. In unfolded PagP, we have previously validated the Cys thiol pKa reported in Gdn-HCl at a value of ~9.4, which was determined by measuring the second order rate constant for Cys thiol methylation using the alkylating agent methyl-*p*-nitrobenzenesulfonate (*126, 162*). Additionally, when PagP is folded by dilution from Gdn-HCl into LDAO, the T_m for unfolding was previously determined by differential scanning calorimetry to be 88°C (*126*). The midpoint for β-barrel unfolding in far-UV CD thermal melts was also found to be 88°C when the 218 nm ellipticity was monitored, but aggregation of the unfolded protein at temperatures above 80°C in this LDAO detergent system renders these precipitation curves irreversible (*126*). In contrast, the Gly88Cys PagP mutant folded at pH values of 7.8 and above displayed a thermal unfolding midpoint of only 67°C (Figure 3.4B,D), so we reasoned that reversible folding might be observed if the system is not heated beyond 80°C. In this case, it should be possible to expose the buried Cys thiolate to solvent at pH 7.2 and then refold the enzyme upon cooling to validate the observed pKa of 7.5.

When Gly88Cys PagP was folded at pH 7.8 and diluted into the same LDAO detergent system at pH 7.2, the exciton remained absent from the far-UV CD spectrum

thus indicating that the buried thiolate is indeed shielded from bulk solvent (Figure 3.5A). Upon heating the sample to 75°C, the thermal melt confirmed the midpoint for β-barrel unfolding at 67°C (Figure 3.5B). Subsequent cooling demonstrated that folding was freely reversible and established the necessary thermodynamic equilibrium to identify 67°C as a true T_m. Interestingly, the exciton was reacquired below 50°C and became fully evident in the wavelength scan after cooling was complete (Figure 3.5 A,B). This observation indicates that exposure of the destabilizing thiolate anion to the lower external pH upon unfolding allowed for a now neutralized thiol to become buried upon refolding, thus restoring the exciton at the hydrocarbon ruler floor. Subsequent heating to 100°C revealed the thermal unfolding profile characteristic of Gly88Cys PagP folded at pH 7.2 (Figure 3.4D and 3.5B). This result was a consequence of the pH difference created upon dilution, rather than a structural reorganization associated with PagP unfolding and refolding, because the exciton was not reacquired when the experiment was repeated by dilution at the same pH 7.8 (Figure 3.5 C,D). The reverse experiment where Gly88Cys PagP was folded at pH 7.2 and diluted into the same LDAO detergent system at pH 7.8 showed that the exciton remained intact even after 75°C heating and cooling (Figure 3.5 E,F) consistent with the 88°C β-barrel T_m, which precludes unfolding without the consequence of aggregation. Indeed, the residual ellipticity observed after heating to 100°C (Figure 3.5E) represents intermolecular β-structure associated with PagP aggregation (*126*).

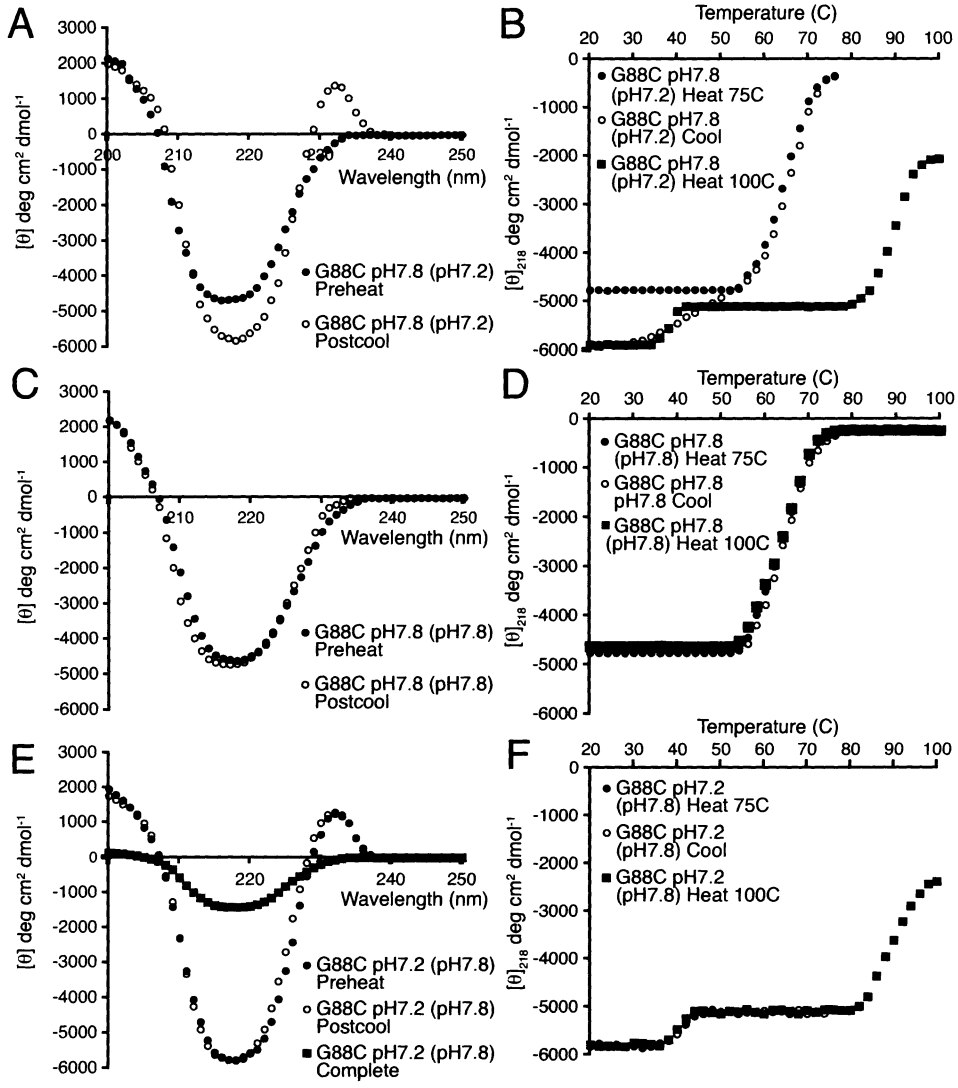

Figure 3.5 Gly88Cys PagP in LDAO must unfold to neutralize the thiolate anion: PagP folded in LDAO at pH 7.8 was diluted into LDAO at pH 7.2, heated to 75°C, cooled, and reheated to 100°C. A) Far-UV CD and B) thermal melts at 218 nm reveal the acquisition of the exciton after reversible unfolding and refolding. PagP folded in LDAO at pH 7.8 was diluted into LDAO also at pH 7.8, heated to 75°C, cooled, and reheated to 100°C. C) Far-UV CD and D) thermal melts at 218 nm reveal the absence of the exciton after reversible unfolding and refolding. PagP folded in LDAO at pH 7.2 was diluted into LDAO at pH 7.8, heated to 75°C, cooled, and reheated to 100°C. E) Far-UV CD and F) thermal melts at 218 nm reveal the retention of the exciton and the absence of thermal unfolding before heating to temperatures above 80°C.

3.3.4 Gly88Cys PagP Folded in DDM Exposes its Thiol/Thiolate to Bulk Solvent:

We have previously established that LDAO binds to the PagP hydrocarbon ruler and thereby inhibits enzymatic activity (*111*). However, diluting the enzyme into certain detergents, which are too bulky to fit within the narrow hydrocarbon ruler pocket, can create a miceller environment capable of supporting PagP activity because such detergents do not compete with substrate binding (*125*). Since the Gly88Cys PagP sulfhydryl group is clearly shielded from bulk solvent when folded in LDAO, we reasoned that the detergent DDM, which is used routinely in our defined detergent micellar enzymatic assay for PagP (*126*), should expose the thiol to solvent without the requirement of first thermally unfolding the enzyme. Indeed, when Gly88Cys PagP folded in LDAO at pH 7.0 or pH 8.0 was diluted into DDM at pH 8.0 or pH 7.0, respectively, the positive exciton ellipticity at 232 nm could be lost or acquired over a period of about 50 minutes (Figure 3.6A), thus revealing a slow solvent exchange not apparent when the protein remained folded in LDAO (Figure 3.5). No ellipticity change was detected when these Gly88Cys PagP proteins were monitored at the same pH in which they were initially folded (not shown).

3.3.5 Neutralization of the Thiolate Anion by Protonation Restores Unitary Acyl Chain Selection:

Diluting wild-type PagP into DDM at a range of pH values revealed the optimum for enzymatic activity of pH 8.0, but 80% of this maximal activity was still retained at pH

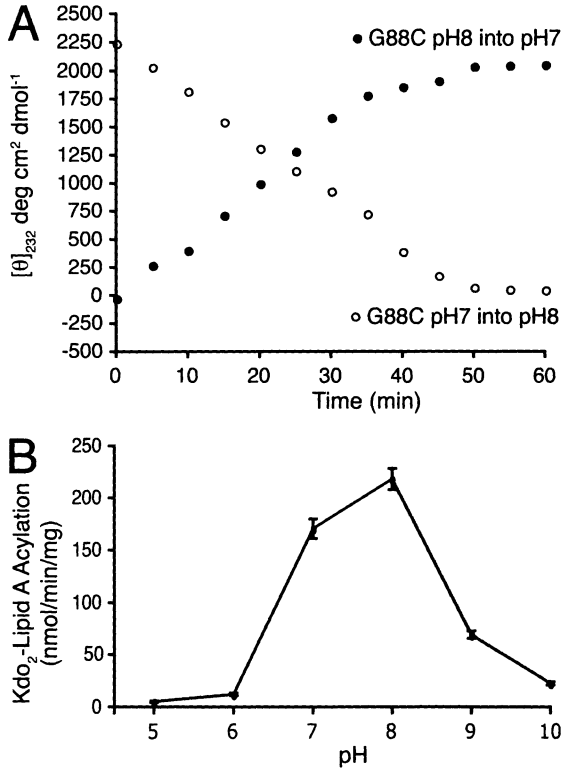

Figure 3.6 Gly88Cys PagP in DDM can expose the thiol/thiolate to solvent: A) Gly88Cys PagP folded in LDAO at pH 8 or pH 7 was diluted into DDM at pH 7 or pH 8, respectively, and the molar ellipticity at 232 nm monitored for 60 min. B) Wild-type PagP folded in LDAO at pH 8 was diluted into the defined detergent micellar enzymatic assay system using DDM at a range of pH values and the acylation of Kdo_2-lipid A by dipalmitoyl-PtdCho was monitored.

7.0 (Figure 3.6B). Since wild-type PagP exhibited substantial activity at both pH 7 and pH 8, and Gly88Cys PagP displayed a pKa for its lone thiol of 7.5, we were in a position to establish if the structural perturbation associated with the buried thiolate anion is responsible for the compromised acyl chain selectivity. When folded and assayed at pH 8.0, the hydrocarbon ruler measurements revealed the preferences of wild-type or Gly88Cys PagP for C16 or both C14 and C15 acyl chains, respectively (Figure 3.7A,B), as noted previously (*126*). However, folding and assaying the same enzymes at pH 7.0 restored methylene-unit resolution to Gly88Cys PagP, which then functioned as a dedicated myristoyltransferase, despite the fact that wild-type PagP remained unaltered as a palmitoyltransferase (Figure 3.7C,D).

3.3.6 Engineered PagP Myristoyltransferases retain Regiospecificity:

Wild-type PagP is highly selective for a palmitate chain, but it also selects the palmitate regiospecifically from the *sn*-1 position in the phospholipid glycerol backbone (*70*). To evaluate whether the engineered PagP myristoyltransferases display similar regiospecificity, we utilized mixed acyl chain PtdCho's having palmitate and myristate esterified to the *sn*-1 and *sn*-2 positions, respectively, or *vice versa*. When challenged with these mixed acyl chain PtdCho's, wild-type PagP effectively utilized only palmitate at the *sn*-1 position (Figure 3.8) and any residual activity from its regioisomer corresponded to the measured background utilization of myristate (Figure 3.7A). Exactly the opposite behaviour was displayed by the engineered myristoyltransferases, thus

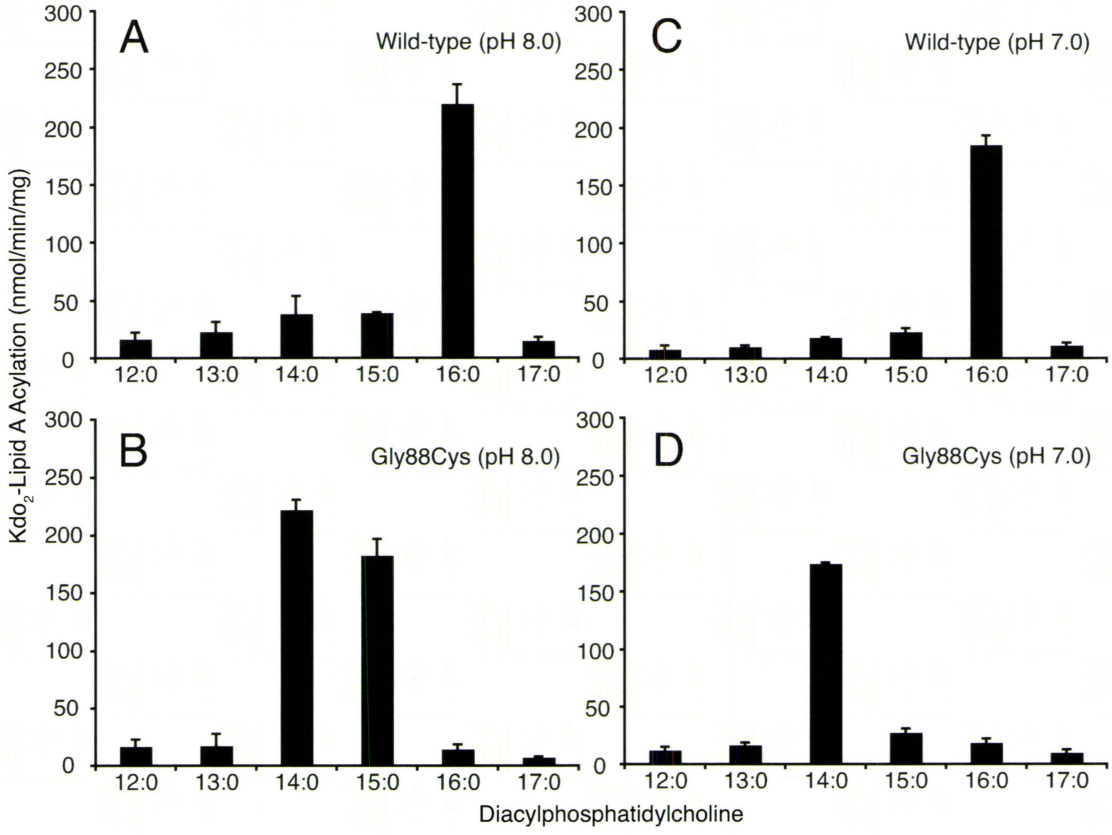

Figure 3.7 A thiol-tiolate ionization mechanism can modulate lipid acyl chain selection: Hydrocarbon ruler experiments were performed for wild-type PagP at pH 8 A) or pH 7 C), and for Gly88Cys PagP at pH 8 B) or pH 7 D). Synthetic diacyl-PtdCho's with saturated acyl-chains varying in single increments from 12 to 17 carbon atoms were used as donors for the enzymatic acylation of Kdo_2-lipid A.

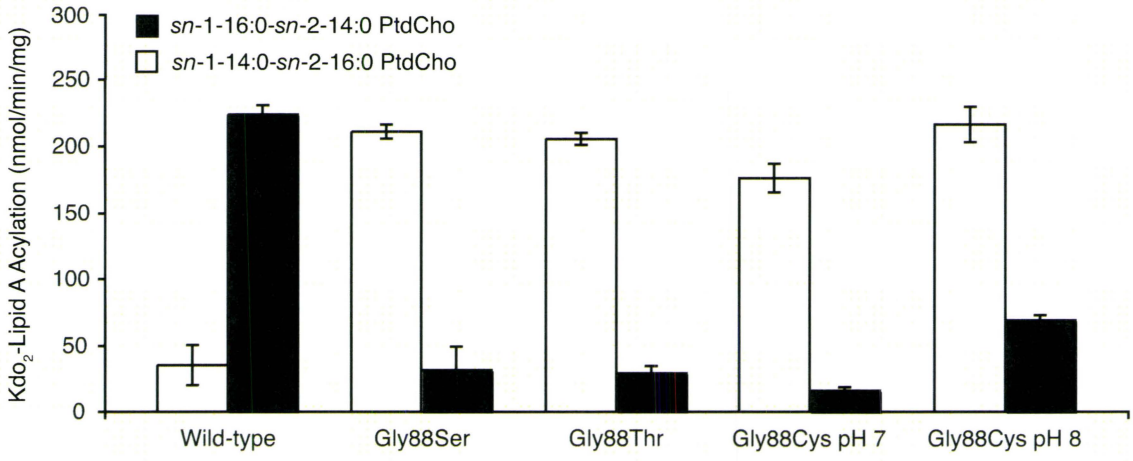

Figure 3.8 Regiospecificity of PagP palmitoyltransferase and myristoyltransferase activity: The indicated enzymes were assayed for Kdo_2-lipid A acylation using *sn*-1-palmitoyl-*sn*-2-myristoyl-PtdCho or *sn*-1-myristoyl-*sn*-2-palmitoyl-PtdCho.

establishing that hydrocarbon ruler modulation influenced acyl chain selectivity without affecting enzyme regiospecificity (Figure 3.8).

3.3.7 Gly88Asp, Asn, Glu, Gln, His and Lys PagP Mutants Retain the Exciton:

If the observed structural perturbation associated with the Gly88Cys thiolate anion is a general effect of burying charge within the PagP hydrocarbon ruler, we reasoned that similar observations should be obtained for other charged amino acid substitutions. We engineered the acidic amino acid substitutions Gly88Asp/Glu together with their corresponding amides Gly88Asn/Gln, and the basic amino acid substitutions Gly88His/Lys (Table 4, and Figure 3.9). Interestingly, the exciton remained intact in all cases (Figure 3.9A), despite the presence of substantial destabilization indicated by approximate 15°C and 10°C reductions in apparent β-barrel T_m for the charged and amide substitutions, respectively (Figure 3.9B). Hydrocarbon ruler experiments revealed that the Gly88Asp/Glu/His mutants could not support substantial enzymatic activity for any of the substrates tested, whereas the remaining substitutions could acylate Kdo_2-lipid A with maximal activities approaching 70% of wild-type PagP. The Gly88Asn substitution exhibited broad acyl chain selectivity centered over C12 (Figure 3.9C), which was approximately consistent with what might be expected if the polar branched Asn side chain projected into the hydrocarbon ruler floor in its extended conformation. However, the Gly88Gln substitution exhibited unusually broad acyl chain selectivity with a peak

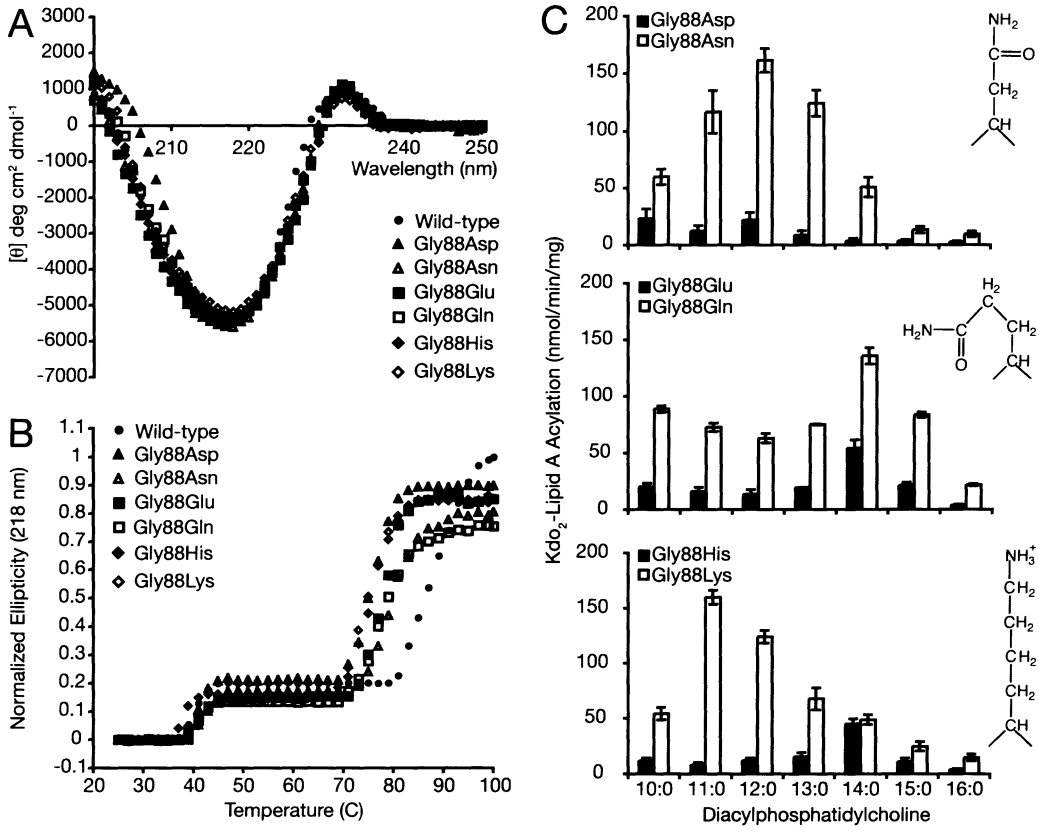

Figure 3.9 Gly88Asp/Asn/Glu/Gln/His/Lys PagP mutants retain the exciton: A) Far-UV CD identifies the β-barrel and exciton signatures characteristic of wild-type PagP. B) Thermal melts at 218 nm reveal exciton loss at 40°C preceding β-barrel unfolding. C) Hydrocarbon ruler measurements reveal broad acyl chain selectivity in branched or charged substitutions, but not in substitutions that are both branched and charged. Synthetic diacyl-PtdCho's with saturated acyl-chains varying in single increments from 10 to 16 carbon atoms were used as donors for the enzymatic acylation of Kdo_2-lipid A. Asn/Lys and Gln amino acid side chains are shown in extended and rotated conformations, respectively.

over C14 (Figure 3.9C), which could only be rationalized in terms of a Gln side chain that had rotated back upon itself. The Gly88Lys substitution shared with the Gly88Cys thiolate substitution the property of being unbranched, charged, and accommodating acyl chains shorter than the one predicted from the length of the substitution itself (Figure 3.9C). Branching and charge thus appear to represent two distinct destabilizing factors when amino acid substitutions at position 88 are introduced within the PagP hydrocarbon ruler. Only groups that were both branched and charged were clearly incompatible with enzymatic activity.

In contrast to these mutants, only the Gly88Cys thiolate introduced a more pronounced 21°C reduction in thermal stability, which was achieved distinctly at the exciton's expense (Figures 3.4 and 3.5). The exciton-proximal Cys thiolate might fit uniquely within the hydrocarbon ruler floor as compared with the more distally positioned perturbing moieties. For example, the β-branched methyl group of Gly88Thr PagP appeared to be accommodated just as easily as Gly88Ser (Figure 3.3), thus identifying at this position an extra degree of steric freedom likely capable of accommodating a Cys thiolate side chain rotation to bring it into contact with the aromatic exciton partners as previously proposed (Figure 3.2C) (*126*). The apparent Gln side chain rotation occurring more distally within the PagP hydrocarbon ruler (Figure 3.9C) drastically compromised acyl chain selection even though the exciton remained intact. This latter point speaks to the local nature of the observed structural perturbation responsible for the loss of the exciton in Gly88Cys PagP. The distinct consequences of

thiolate ionization in Gly88Cys PagP reveal not just a general effect of buried negative charge, but also a role for the local context of steric and electronic constraints within the hydrocarbon ruler environment.

3.3.8 PagP Myristoyltransferase Mutant Activity can be Detected *in vivo*:

In exponentially growing *E. coli* cells, phospholipids are esterified with roughly 40% palmitate, 32% palmitoleate, 16% *cis*-vaccenate, 5% cyclopropane fatty acids, and another 5% is attributed to myristate (*193, 194*). Therefore, we chose to examine whether or not our engineered PagP myristoyltransferases could incorporate myristate into lipid A *in vivo*. We cloned the native *pagP* gene under control of its endogenous promoter into plasmid pBad18 (*185*) to generate pPagP. The Gly88Cys, Ser, and Thr substitutions were then generated by site-directed mutagenesis to produce pG88C, pG88S, and pG88T, respectively. Each plasmid was transformed into a wild-type strain of *E. coli* K12 in which the chromosomal copy of the *pagP* gene had been deleted. Cells were grown to exponential phase and treated prior to harvesting with EDTA, which stimulates PagP activity *in vivo* by promoting migration of phospholipids into the external leaflet where they can access the active site by a lateral lipid diffusion mechanism (*123, 182*). The *E. coli* outer membrane lipids are normally asymmetrically organized with lipopolysaccharide restricted to the external leaflet and phospholipids restricted to the inner leaflet, which thus maintains PagP in a latent state (*181*).

We first evaluated the incorporation of a seventh acyl chain into the hexa-acylated 1,4'-*bis*-phosphorylated lipid A species using [^{32}P] labelling for analysis by TLC (*123, 191*) (Figure 3.10A). The pG88C transformant was cultured at both pH 7 and pH 8. Substantial production of a hepta-acylated lipid A species was detected in all PagP-expressing cells except for that of the pG88S transformant. Although the Gly88Ser mutant functions as a dedicated myristoyltransferase *in vitro* (Figure 3.3C), it appears that this mutant is not functional *in vivo* perhaps due to a defect in expression and/or outer membrane assembly. To evaluate the nature of the acyl-chains incorporated into lipid A of EDTA-treated cells, we isolated non-radioactive lipid A for analysis by ESI-MS (*192*) (Figure 3.10B-F). The *E. coli* Δ*pagP* cells produced hexa-acylated lipid A, which displays a mass of 1797 Da with associated sodium adduct of mass 1819 Da (Figure 3.10B). Transformation with pPagP was expected to result in the incorporation of a palmitate chain at position 2 (*113, 191*), consistent with the observed mass of 2035 Da with associated sodium adduct of mass 2057 Da (Figure 3.10C). Interestingly, the pG88C transformant grown at pH 7 incorporated a new species of mass 2007 Da, consistent with myristate without any associated sodium adduct (Figure 3.10D). Only a trace of palmitate was apparent despite its 8-fold excess over myristate in the *E. coli* phospholipid pool (*194*). When this same mutant was grown at pH 8, palmitate dominated over myristate (Figure 3.10 E) suggesting that the buried thiolate anion associated with PagP folding at pH 8 *in vitro* (Figure 3.7B) also compromises lipid acyl chain selection *in vivo*. Transformation with pG88T resulted in myristate incorporation

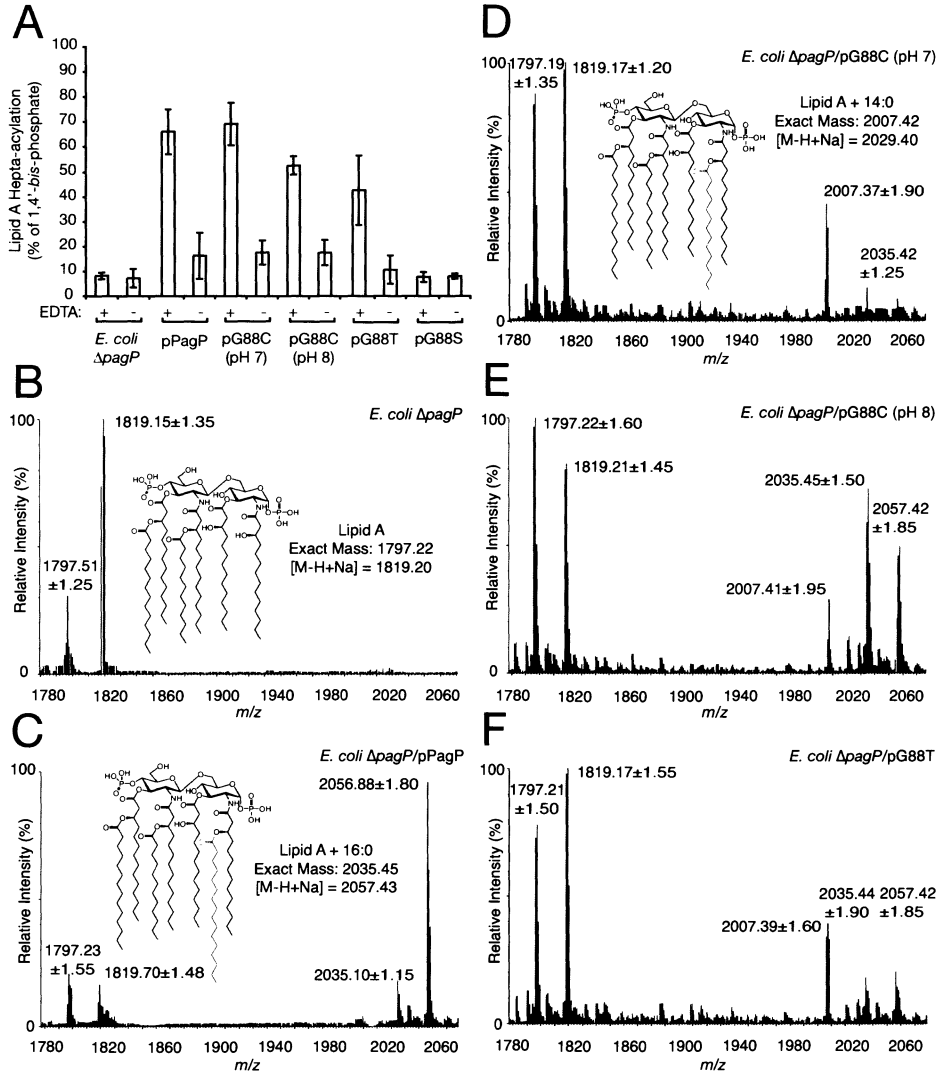

Figure 3.10 Lipid A myristoylation detected *in vivo*. A) The [32P]-labeled lipid A 1,4'-bis-phosphate was resolved by TLC as hexa- and hepta-acylated species. The percentage of hepta-acylation in exponentially grown cultures harvested with and without EDTA treatment is shown for the pagP deletion strain (*E. coli* ΔpagP) and its transformants with plasmids pPagP, pG88C (grown at pH 7 and pH 8), pG88T, and pG88S. B) ESI-MS of the lipid A 1,4'-bis-phosphate isolated from EDTA-treated E. coli ΔpagP. C) ESI-MS of the lipid A 1,4'-bis-phosphate isolated from EDTA-treated E. coli ΔpagP/pPagP. D) ESI-MS of the lipid A 1,4'-bis-phosphate isolated from EDTA-treated *E. coli* ΔpagP/pG88C (pH 7). E) ESI-MS of the lipid A 1,4'-bis-phosphate isolated from EDTA-treated *E. coli* ΔpagP/pG88C (pH 8). F) ESI-MS of the lipid A 1,4'-bis-phosphate isolated from EDTA-treated *E. coli* ΔpagP/pG88T.

with a clear detection of palmitate (Figure 3.10F) consistent with the slight destabilization associated with the Gly88Thr substitution (Figure 3.3B) compared to Gly88Cys at pH 7 (Figure 3.4B). These findings indicate that trace amounts of myristate present in *E. coli* phospholipids can be selectively incorporated into lipid A by PagP myristoyltransferase mutants provided they assemble in outer membranes and are sufficiently stable to exclude the excess palmitate encountered in cellular phospholipids.

3.4 Concluding Remarks:

We have demonstrated that the floor of the hydrophobic acyl chain-binding pocket in a Gly88Cys PagP mutant serves to suppress the thiol pKa, which resulted in a local structural perturbation from a buried thiolate anion. Neutralization of the thiolate anion by protonation restored normal function and stability to the mutant enzyme. The question as to why the hydrophobic pocket would suppress the thiol pKa when the expected effect is to stabilize the protonated thiol, and thus raise the pKa, remains. In a recent study of acetoacetate decarboxylase, the pKa of Lys115 was perturbed downward by 4.5 log units due to the localization of the ε-amino group near the floor of a hydrophobic pocket formed by a β-barrel domain (*195*). The reduced dielectric constant of this microenvironment, which is reminiscent of the PagP hydrocarbon ruler, probably lowers the pKa of Lys115 through a desolvation effect where the uncharged amino group is better stabilized in a hydrophobic milieu. The same reasoning applied to Cys88 in PagP predicts that the low dielectric environment of the hydrocarbon ruler should instead

raise the Cys pKa by stabilizing the neutral thiol (*183*). A comprehensive investigation of perturbed pKa values for catalytic groups in enzyme active sites has revealed only two mechanisms by which the pKa of a Cys thiol can be suppressed (*183*). Interaction with the N-terminal dipole of an α-helix or formation of an ion pair between a Cys thiolate and a His imidazolium ion can both suppress the thiol pKa, but neither of these factors are apparent near the floor of the PagP hydrocarbon ruler. Nevertheless, thiols are particularly sensitive to local electrostatics, which have been previously argued to lower the pKa in otherwise hydrophobic environments (*196, 197*). The thiol in Gly88Cys PagP is flanked by the low dielectric environment of the hydrophobic acyl chain-binding pocket above and by numerous dipoles among the aromatic exciton partners and a largely polar β-barrel interior below. Therefore, details of the electrostatics located adjacent to the non-polar milieu within the PagP interior might eventually reveal a novel mechanism to suppress the pKa of Cys thiols in proteins. To our knowledge, the use of an identified aromatic exciton couplet as a probe to monitor the pKa of an ionizable functional group has not been previously described in enzymology.

4.0 Molecular Mechanism for Lateral Lipid Diffusion Between The Outer Membrane External Leaflet and a β-Barrel Hydrocarbon Ruler

Work presented in this chapter was published in the peer-reviewed journal *Biochemistry* in 2009: Khan, M.A., Bishop, R.E. "Molecular mechanism for lateral lipid diffusion between the outer membrane external leaflet and a beta-barrel hydrocarbon ruler". *Biochemistry*. 2009; 48(41):9745-56.

4.1 Introduction:

The molecular mechanisms by which integral membrane enzymes recognize lipid substrates impact signal transduction and membrane biogenesis within all biological cells (*132, 179, 198*). Despite recent progress in the structural biology of integral membrane proteins (*199*), only a handful of known structures are enzymes of lipid metabolism. Along with the lipid cofactor-utilizing respiratory enzymes, only two other enzymes of eicosanoid biosynthesis (*200-202*) and a diacylglycerol kinase (*203*) represent solved subunit structures of bitopic transmembrane α-helical lipid metabolic enzymes. Another group of monotopic lipid enzyme structures engage the hydrophobic core of the bilayer using α and/or β motifs to bury into one membrane leaflet only (*135, 180*). Fundamental questions concerning lipid recognition by membrane enzymes remain largely unexplored (*180*), but robust β-barrel enzymes from the outer membranes of Gram-negative bacteria are providing interesting answers (*98, 104, 109, 181*).

The outer membrane palmitoyltransferase PagP catalyzes palmitate transfer from the *sn*-1 position of a phospholipid to the hydroxyl group of the *R*-3-hydroxymyristate chain at position 2 of lipid A (endotoxin) (Figure 4.1A) (*28, 70*). PagP is a heat-stable

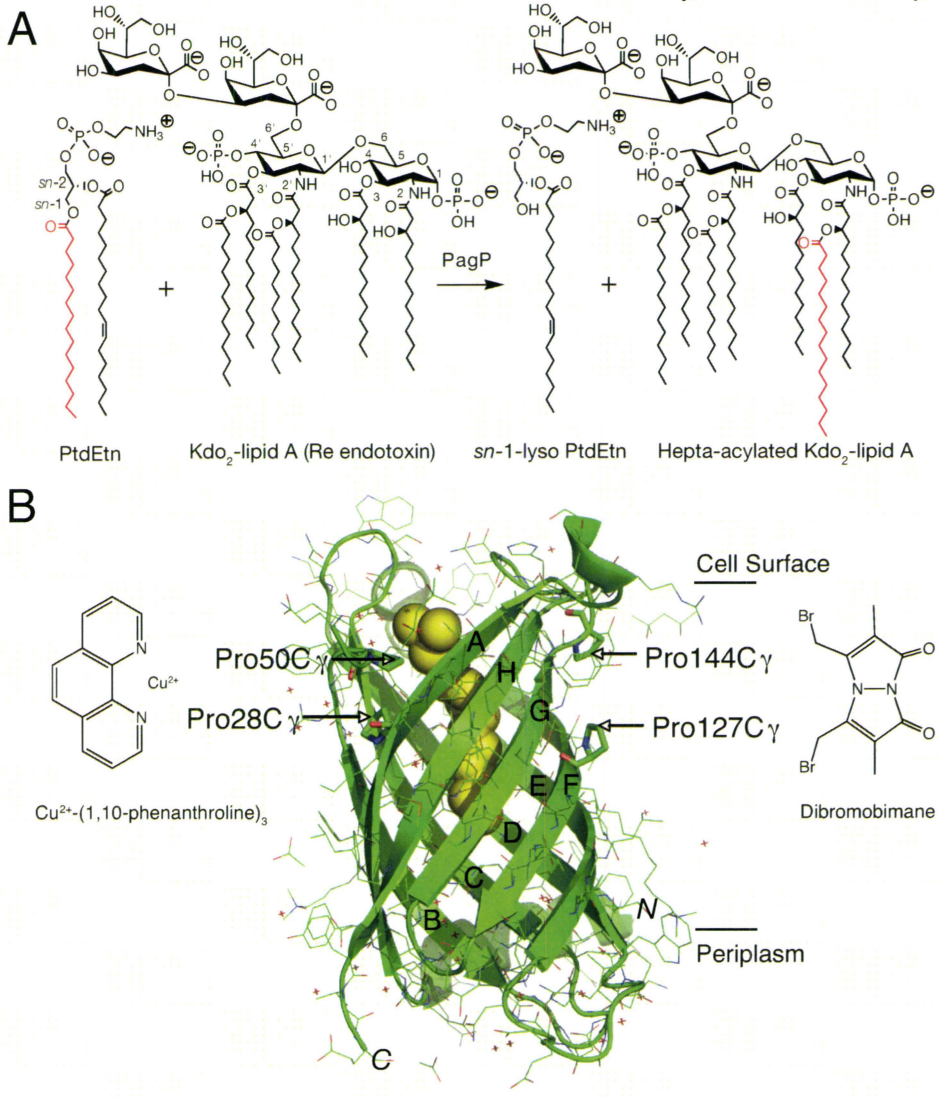

Figure 4.1 A lipid A palmitate chain is selected by a β-barrel interior hydrocarbon ruler: (A) PagP transfers a palmitoyl group from the *sn*-1 position of a phospholipid, such as phosphatidylethanolamine (PtdEtn), to the free hydroxyl-group of the *N*-linked *R*-3-hydroxymyristate chain on the proximal glucosamine unit of lipid A. One of the simplest lipid A acceptors for PagP in the outer membrane is known as Kdo_2-lipid A or Re endotoxin. (B) PagP is an 8-stranded antiparallel β-barrel (green) with its hydrocarbon ruler delineated by an internalized LDAO detergent molecule (yellow), which is aligned with the external leaflet of the bacterial outer membrane. Pro to Cys substitutions flanking the PagP embrasure (Pro28Cys/Pro50Cys) and crenel (Pro127Cys/Pro144Cys) indicate cysteine sulfhydryl groups will localize near Cγ of the corresponding proline residues. Oxidation of the PagP embrasure can be driven by copper phenanthroline, whereas the crenel can by alkylated by dibromobimane.

161-amino acid β-barrel enzyme, intrinsically lacking cysteine residues, which can fold reversibly in a defined detergent micellar enzymatic assay system (*126*). In *Escherichia coli* and related pathogenic Gram-negative bacteria, PagP combats host immune defences by restoring the bacterial outer membrane permeability barrier (*114, 115, 118, 150*), and attenuating endotoxin signalling through the host TLR4 inflammation pathway (*7, 65, 177, 178*). The *pagP* gene is transcriptionally activated through the bacterial PhoP/PhoQ virulence signal transduction network in response to membrane-perturbing antimicrobial peptides (*66*), immunological agents PagP helps the bacteria to resist (*113, 115, 118*). PagP thus provides a target to develop anti-infective agents, but it also provides a tool to synthesize novel vaccine adjuvants and endotoxin antagonists (*176*).

PagP is exquisitely selective for a 16-carbon palmitate chain because its hydrocarbon ruler excludes lipid acyl chains varying in length by a solitary methylene unit (*111, 126*). The PagP hydrocarbon ruler is delineated by a detergent-binding hydrophobic pocket buried within the 8-stranded antiparallel β-barrel structure (Figure 4.1B) (*111*), and mutation of Gly88 can raise the hydrocarbon ruler floor to correspondingly shorten the selected acyl chain (*126*). The localization of the PagP hydrocarbon ruler within the outer membrane external leaflet must be rationalized with the asymmetric lipid distribution of the bacterial outer membrane. The external leaflet is lined by lipopolysaccharide molecules anchored by their lipid A moiety, whereas the periplasmic leaflet is lined by phospholipid molecules (*3, 149*). Hydrophobic antibiotics and detergents freely permeate phospholipid bilayers, but they cannot penetrate the outer

membrane because neighbouring negatively charged LPS groups are ionically bonded by bridging divalent cations (*121, 204, 205*). PagP lies dormant in the outer membrane when lipid asymmetry restricts the phospholipid palmitoyl donor to the inner leaflet, but permeability defects involving aberrant external leaflet phospholipid migration directly trigger lipid A palmitoylation (*40, 44, 123, 150, 206*).

Once a phospholipid molecule arrives in the outer membrane external leaflet, the mechanism by which the palmitoyl group transits between the membrane and PagP interior hydrocarbon ruler remains to be elucidated. A driving force for outer membrane β-barrel protein folding comes from hydrogen bonding between backbone amide proton and carbonyl oxygen atoms within the non-polar membrane environment (*78, 207*). The interiors of typical β-barrel membrane proteins are shielded from membrane lipids because of this continuous β-strand hydrogen bonding, and proline residues lacking an amide proton to donate a hydrogen bond are usually excluded from transmembrane β-strands (*208*). Nevertheless, proline residues are strikingly localized within the PagP transmembrane β-strands at two sites flanking Pro28 in strand A and Pro50 in strand B (Figure 4.1B), which creates a β-bulge beneath the dynamic L1 surface loop (*110, 125*), and flanking Pro127 in strand F and Pro144 in strand G (*111*). Resembling structures in the turret of a medieval castle, we refer to these regions of diminished transmembrane β-strand hydrogen bonding in PagP as the embrasure, for the window between strands A and B, and as the crenel, for the indentation between strands F and G.

Although the embrasure and crenel provide potential routes for lateral lipid diffusion between the hydrocarbon ruler and outer membrane external leaflet, the PagP crystal structure indicates physical openings to facilitate lateral lipid exchange are blocked by amino acid side chains (*111*). The question as to whether conformational changes gate lipid access, or if an altogether different mechanism extrudes the palmitoyl group out of the membrane plane and over the β-barrel wall, remains. Here we mutate prolines flanking the PagP embrasure and crenel to cysteines without compromising protein folding and activity *in vitro*. The influence on PagP enzymology of physical barricades between the flanking cysteines is wholly consistent with a lateral diffusion lipid gating mechanism. These and other mutational and enzymological results support a model where lipid acyl chains diffuse laterally between the biological membrane and the interior of an integral membrane lipid-metabolic enzyme.

4.2 Experimental Procedures:

4.2.1 DNA Manipulations and Protein Purification:

Mutagenesis was performed using the primer sets shown in Table 2 and the pETCrcAHΔS vector as described previously (*110*). PagP was expressed using isopropyl-β-D-thiogalactopyranoside (IPTG) induction in *E. coli* BL21(DE3) for purification and folding as previously reported (*126*). Protein samples precipitated from water prior to detergent folding were dissolved in a solution of 50:50 1% formic acid:acetonitrile just prior to electrospray ionization mass spectrometry (ESI-MS). The

sample concentration was maintained at 1 ng/μl and injected directly onto a Waters/Micromass Q-TOF Ultima Global (a quadrupole time-of-flight) mass spectrometer. The spectra were reconstructed using the MassLynx 4.0 MaxEnt 1 module. Detergent folded samples were precipitated for ESI-MS by adding 1.25 ml of PagP at 1 mg/ml in 10 mM Tris-HCl (pH 8.0) and 0.1% lauroyldimethylamine-*N*-oxide (LDAO) to 3.75 ml of 8 M guanidine-HCl (Gdn-HCl) and dialyzing overnight against water. To avoid intermolecular cross linking between Cys carrying mutants during folding, precipitated samples were initially solubilized in 5 ml of 10 mM Tris-HCl (pH 8.0), 6 M Gdn-HCl, 10 mM β-mercaptoethanol (βME) and heated at 50°C for two hours. The folding buffer also contained 5 mM βME (*126*). Folded proteins were concentrated through Ni^{+2}-affinity chromatography and dialyzed against 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO to remove imidazole and βME. βME was not needed for Cys carrying mutations after folding because the sulfhydryl groups were protected from air oxidation within the detergent micelle. All protein stock solutions were maintained at 2 mg/ml concentration determined using the Edelhoch method (*152*) or bicinchoninic acid assay (*151*).

4.2.2 Copper Phenanthroline Oxidation:

Cu^{+2}-(1,10-phenanthroline)$_3$ ($Cu(OP)_3$) can drive disulfide bond formation in membrane environments (*127, 209*). To obtain a final concentration of 10 mM $Cu(OP)_3$ in 10 ml, 19.8 mg of 1,10-phenanthroline monohydrate was dissolved in 0.5 ml ethanol

and the volume was adjusted with 9.5 ml of dH_2O prior to dissolving 13.22 mg of $Cu(II)SO_4$. The oxidation was carried out with 1 mM $Cu(OP)_3$ at 4° C for 2 hours in a 15 ml falcon tube using 5 ml of 0.1 mM PagP in 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO. The reaction was terminated by addition of 20 µl of 500 mM ethylenediaminetetraacetic acid (EDTA) and the mixture was dialyzed overnight against 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO. Heating protein in the presence of 500 mM βME at 50°C for 2 hrs reduced the disulfide bridge (*162*).

4.2.3 Dibromobimane Alkylation:

Dibromobimane (bBBr) introduces a fluorophore across thiol groups lying within 5 to 7 Å of each other (*128*). To generate a 100 mM bBBr solution, 0.035 g of bBBr was dissolved in 1 ml of acetonitrile. PagP at 0.1 mM in 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO was incubated at 37°C for 1 hour with 1 mM bBBr. The reaction was terminated by addition of 5 mM βME and dialyzed against 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO. Fluorescence spectra were obtained using a Varian Carry Eclipse fluorescence spectrophotometer. Protein samples in 10 mM Tris-HCl (pH 8.0) and 0.1 % LDAO at a concentration of 0.50 mg/ml were excited at 393 nm and emission spectra were obtained for a range of 400-600 nm.

4.2.4 Quantitation of Sulfhydryl Groups Using Ellman's Reagent:

Ellman's reagent 5, 5'-dithio-*bis*-(2-nitrobenzoate) (DTNB) was freshly dissolved in 50 mM sodium acetate at 2 mM and kept at 4°C. Free thiols in PagP were reacted

under either native (0.1% LDAO) or denaturing (6M Gdn-HCl) conditions using 0.1 mM DTNB in 100 mM Tris-HCl pH 8.0. After 10 minutes incubation at room temperature, the absorbance at 412 nm was measured using a Varian Cary 50 Bio-UV-visible spectrophotometer and the free thiol concentration estimated using the extinction coefficient of 13, 600 M^{-1} cm^{-1} (*210*).

4.2.5 SDS-PAGE Analysis:

The folded PagP samples were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using a non-commercial 16% Tris-glycine gel prepared under reducing and non-reducing conditions (*211*). Prior to loading the gel, samples were diluted 1:1 with Laemmli solubilization buffer with or without 10 mM dithiothreitol (DTT) for reducing conditions and non-reducing conditions, respectively. The gel was run at 120 mV and stained using Coomassie Blue. After staining for 2 hours the gel was destained overnight in 10% acetic acid. Each lane was loaded with 6 μg of PagP unheated or boiled at 100°C for 10 min.

4.2.6 CD Spectroscopy:

Samples analyzed by near and far ultraviolet (UV) circular dichroism (CD) were maintained at a concentration of 1.0 mg/ml (near-UV CD) and 0.3 mg/ml (far-UV CD) in 10 mM Tris-HCl (pH 8.0) and 0.1% LDAO using a cuvette of 1 mm path length. The samples were analyzed using an Aviv 215 spectrophotometer, which was linked to a Merlin Series M25 Peltier device for temperature control. For each sample 3

accumulations were averaged at a data pitch of 1 nm and a scanning speed of 10 nm/min. The temperature was maintained at 25°C and data sets were obtained from 200-260 nm for far-UV CD and 250-300 nm for near-UV CD. Thermal denaturation profiles were obtained by heating the samples from 20°C to 100°C at 218 nm with a temperature slope of 2°C/min and a response time of 16 seconds.

4.2.7 Palmitoyltransferase Assays:

Details for the preparation of ^{32}P-Kdo_2-lipid A, which carries two units of 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo), for thin layer chromatography (TLC) assays have been previously described (*126*). Phospholipase assays were carried out using phosphatidylcholine (PtdCho) in a volume of 25 µl with di-[1-^{14}C]-16:0-PtdCho to achieve a final concentration of 20 µM (4000 cpm/µl). The lipid was dried under a stream of N_2 (g) and dissolved in a 22.5 µL reaction buffer containing 0.1 M Tris-HCl (pH 8.0), 10 mM EDTA and 0.25% *n*-dodecyl-b-D-maltoside (DDM). The reactions were started by adding 2.5 µl of PagP to give 1 mg/mL in the assay conducted at 30°C. Reactions were terminated by adding 12.5 µl of the reaction mixture to 22.5 µl of 1:1 $CHCl_3$/MeOH to generate a 2-phase mixture from which 5 µl of the lower organic phase provided 10,000 cpm for spotting onto a silica gel 60 TLC plate. The experiment was carried out over five hours with the reaction being spotted once every hour. The TLC plates were developed in a $CHCl_3$:MeOH:H_2O (65:25:4 v/v) solvent system equilibrated in sealed glass tank. The plates were exposed overnight to a PhosphorImager screen and

developed the following day with a Molecular Dynamics Typhoon 9200 PhosphorImager. Reactions were supplemented with 10% v/v glycerol or 100 μM 16:0 and 14:0 monoglycerides where indicated. The phospholipase A_2 control (4 mU/μL) was incubated with 10 mM $CaCl_2$ in place of EDTA.

4.3 Results:

Proline residues were substituted by cysteine, both singly and in pairs, using site-directed mutagenesis with the primer sets shown in Table 1 and the IPTG-inducible PagP expression plasmid pETCrcAHΔS (*110*). The expressed proteins possess a *C*-terminal hexahistidine tag and lack the *N*-terminal signal peptide to target PagP for secretion to the outer membrane. The proteins are expressed as insoluble aggregates, which can be dissolved in Gdn-HCl, purified by nickel-affinity chromatography, and folded by dilution into the detergent LDAO (*126*). Mutations were confirmed both by DNA sequencing of the mutant plasmids and by ESI-MS of the purified proteins. ESI-MS performed on PagP and its derivatives (*124, 126*) yields experimental protein masses accurately matching theoretical predictions (Table 2).

4.3.1 Barricading the PagP Embrasure and Crenel by Cysteine Cross-Linking:

Based on the PagP structure previously crystallized from LDAO, we expected cysteine sulfhydryl groups in Pro→Cys double mutants to localize near C_g of the corresponding proline residues (Figure 4.1B) (*111*). On that basis, the embrasure sulfhydryl groups of Pro28Cys/Pro50Cys PagP should localize in a dynamic region of the

protein (*110, 125*) within sufficient proximity to form an intramolecular disulfide bond. In contrast, the crenel sulfhydryl groups of Pro127Cys/Pro144Cys PagP should localize in a structured region separated by a gap of ~5 Å thus requiring a local polypeptide backbone repositioning to establish a 2.05 Å disulfide bond (*111*). Interestingly, no spontaneous cysteine oxidation occurs in the detergent micelle environment for either folded Pro28Cys/Pro50Cys PagP or folded Pro127Cys/Pro144Cys PagP, and only in the former mutant could the reaction be driven forward by addition of the membrane permeable oxidizing reagent $Cu(OP)_3$ (Figure 4.1B) (*127, 209*).

ESI-MS could not reliably measure the 2.02 Da decrease associated with PagP embrasure intramolecular disulfide bond formation (Table 4), but the full mass spectrum clearly excludes any intermolecular dimer formation following $Cu(OP)_3$ oxidation (Figure 4.2 A). To validate intramolecular disulfide bond formation, we took advantage of the intrinsically dissymmetric disulfide bond $n \rightarrow \sigma^*$ transitions, which, together with contributions from interacting aromatic side chain $\pi \rightarrow \pi^*$ transitions, can manifest near-UV CD ellipticity centered around 260 nm (*143, 212*). A dependence of the CD rotational strengths on the disulfide bond dihedral angle means either positive or negative ellipticity can manifest itself in the near UV (*213-215*). Intriguingly, PagP embrasure oxidation using $Cu(OP)_3$ introduced positive ellipticity at 260 nm, and subsequent reduction by excess β-ME occurred only when combined with a 50°C heat treatment (Figure 4.2 B) (*162*).

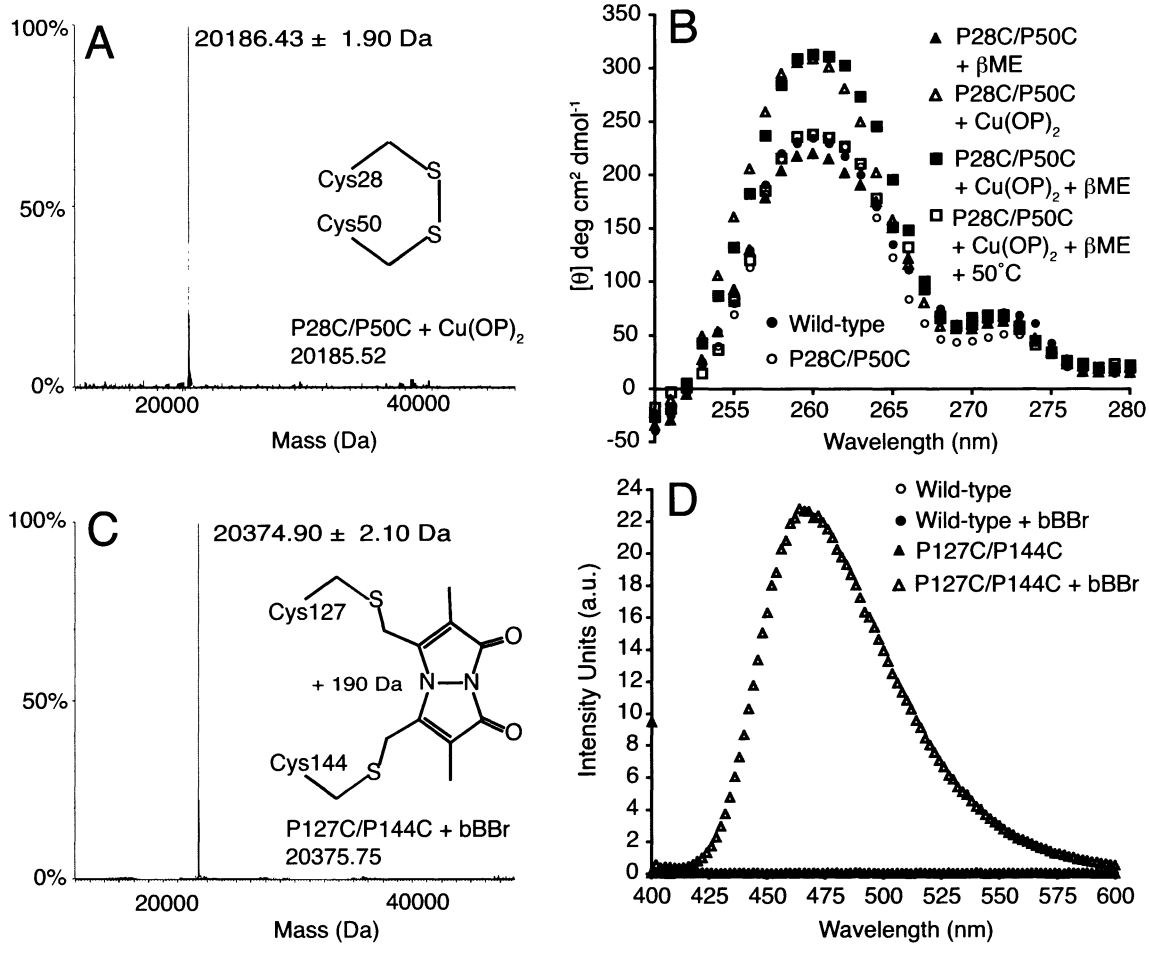

Figure 4.2 Barricading the PagP embrasure and crenel by cysteine cross-linking: (A) ESI-MS identifies a single monomeric peak after oxidation of the PagP embrasure. (B) Near-UV CD identifies reversible oxidation and reduction of the PagP embrasure disulfide bond. (C) ESI-MS identifies a single monomeric peak after bimane alkylation of the PagP crenel. (D) Excitation at 393 nm reveals the emission spectrum for the bimane fluorophore.

Our failure to oxidize the PagP crenel sulfhydryl groups suggests insufficient conformational dynamics exist to bridge the ~5 Å gap. We opted instead to introduce a barricade using irreversible alkylation with the homobifunctional cross-linking reagent bBBr, which is ideally suited to bridge gaps of ~5 Å (*128-131*). Intramolecular bimane cross-linking of the PagP crenel was validated by ESI-MS detection of the expected 190 Da increase (Table 4) with no intermolecular dimer formation (Figure 4.2C). Fluorescence imparted by bBBr bisalkylation was detected with an excitation at 393 nm resulting in a λ_{max} of 466 nm (Figure 4.2 D). The 477 nm λ_{max} reported previously for soluble bBBr-modified proteins (*128*) suggests an 11 nm blue-shift is associated with bimane cross-linking of the PagP crenel, possibly reflecting fluorophore localization within a more hydrophobic environment.

4.3.2 Embrasure and Crenel Barricades do not Compromise PagP Structural Integrity:

SDS PAGE provides a facile means to evaluate the effectiveness of PagP folding *in vitro*. PagP folded in LDAO can be diluted into SDS where the β-barrel structure is maintained until heating induces a denatured state dominated by α-helical structure (*126*). When resolved by SDS PAGE, the folded β-form of PagP migrates faster in some gel systems (*70, 126*), and slower in others (*216, 217*), but quantitative folding manifests a single band regardless of the gel system employed. According to this criterion, the PagP embrasure and crenel Pro→Cys mutants were all completely folded, as were the

double mutants before and after chemical treatments with $Cu(OP)_3$ and bBBr (Figure 4.3). Under non-reducing conditions only the unfolded α-forms possessing free sulfhydryl groups were capable of migrating as disulfide-linked dimers. Consistent with our observed $Cu(OP)_3$ requirement to drive embrasure disulfide bond formation in LDAO (Figure 4.2B), the sulfhydryl groups in the folded b-form PagP Pro→Cys mutants were all protected from spontaneous air oxidation (Figure 4.3), possibly due to shielding within a more hydrophobic micellar milieu.

The PagP far-UV CD spectrum provides a measure of secondary structural content and thermal stability (*126, 143*), but also a measure of tertiary structural integrity because the hydrocarbon ruler floor is coupled to a rare signature in CD known as an aromatic exciton couplet (*147*). The exciton is a consequence of excited state delocalization between Tyr26 and Trp66 aromatic $\pi \rightarrow \pi^*$ transitions, which in their folded geometry near the hydrocarbon ruler floor interact to manifest positive ellipticity at 232 nm together with an equal magnitude enhancement of the negative ellipticity at 218 nm. The 218 nm signal in PagP thus results from the superpositioning of the negative exciton Cotton effect upon that for the stronger polypeptide $n \rightarrow \pi^*$ transition, which also arises at 218 nm due to the PagP β-barrel structure (*126*). The exciton contribution at 218 nm is thermally sensitive because it disappears in response to heating above 40°C, whereas the $n \rightarrow \pi^*$ transition is lost only at temperatures above 80°C (Figure 4.4). Protein aggregation occurring at temperatures above 80°C precludes reversible folding in this LDAO detergent system, but the midpoint of 88°C corresponds to a true

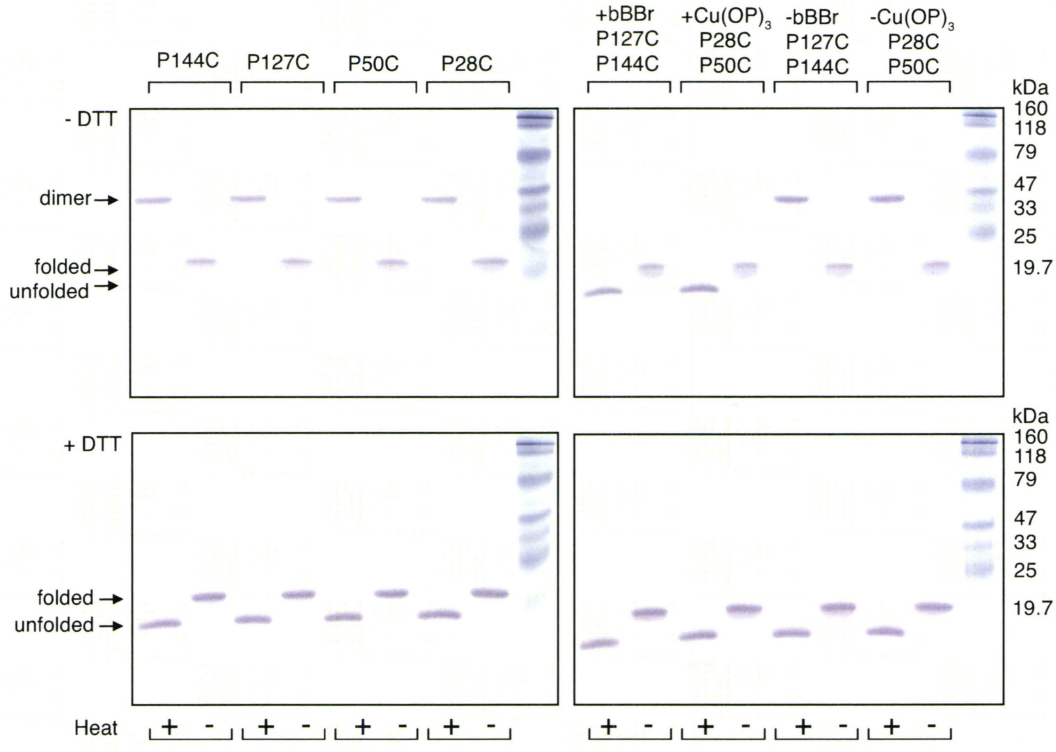

FIGURE 4.3 SDS-PAGE of PagP embrasure and crenel: Pro to Cys single mutants and the double mutants with and without heat denaturation and under dithiothreitol (DTT) reducing or non-reducing conditions.

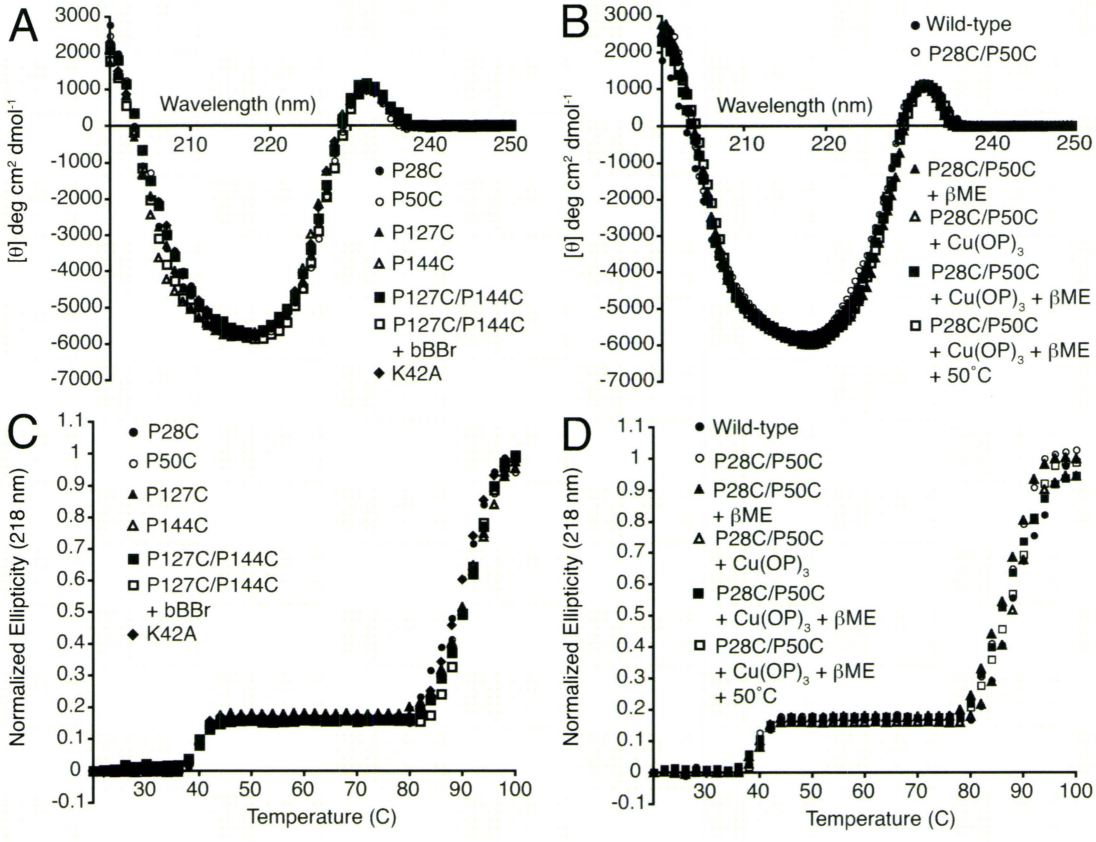

FIGURE 4.4 Far-UV CD spectroscopy performed on PagP embrasure and crenel Pro to Cys mutants, before and after embrasure oxidation/reduction and crenel alkylation, and including the Lys42Ala substitution: Wavelength scans are shown in panels (A) and (B). Thermal melts are shown in panels (C) and (D).

melting temperature established previously by differential scanning calorimetry (*126*). Importantly, the embrasure and crenel Pro→Cys mutants, and the double mutants before and after chemical treatments with $Cu(OP)_3$ and bBBr, all display far-UV CD spectra and 218 nm thermal unfolding profiles indistinguishable from wild-type PagP (Figure 4.4). Furthermore, the intrinsic thermal stability of the PagP β-barrel satisfactorily protects the enzyme from denaturation during 50°C βME reduction of the oxidized embrasure (Figures 4.4B and 4.4D). We conclude PagP tertiary structural integrity and stability has not been compromised by the embrasure and crenel Pro→Cys substitutions, by the reversible oxidation/reduction of the embrasure, or by the irreversible alkylation of the crenel.

4.3.3 Lys42Ala PagP is a Dedicated Phospholipase *in vitro*:

PagP phospholipid:lipid A palmitoyltransferase activity is monitored *in vitro* using a defined detergent micellar enzymatic assay with TLC separation of radioactive lipid products (*126*). The inhibitory LDAO detergent used during PagP folding is exchanged by dilution into DDM to support enzymatic activity. PagP is highly selective for a palmitoyl group at the *sn*-1-position in a glycerophospholipid, but largely unspecific for the polar head group (*70*). We employ the palmitoyl donor di-16:0-PtdCho and the palmitoyl acceptor ^{32}P-Kdo_2-lipid A. In the absence of a lipid A acceptor, a slow phospholipase reaction occurring at 20% the rate of lipid A palmitoylation can also be monitored *in vitro* using di-[1-^{14}C]-16:0-PtdCho (Figure 4.5) (*218*). Although three PagP

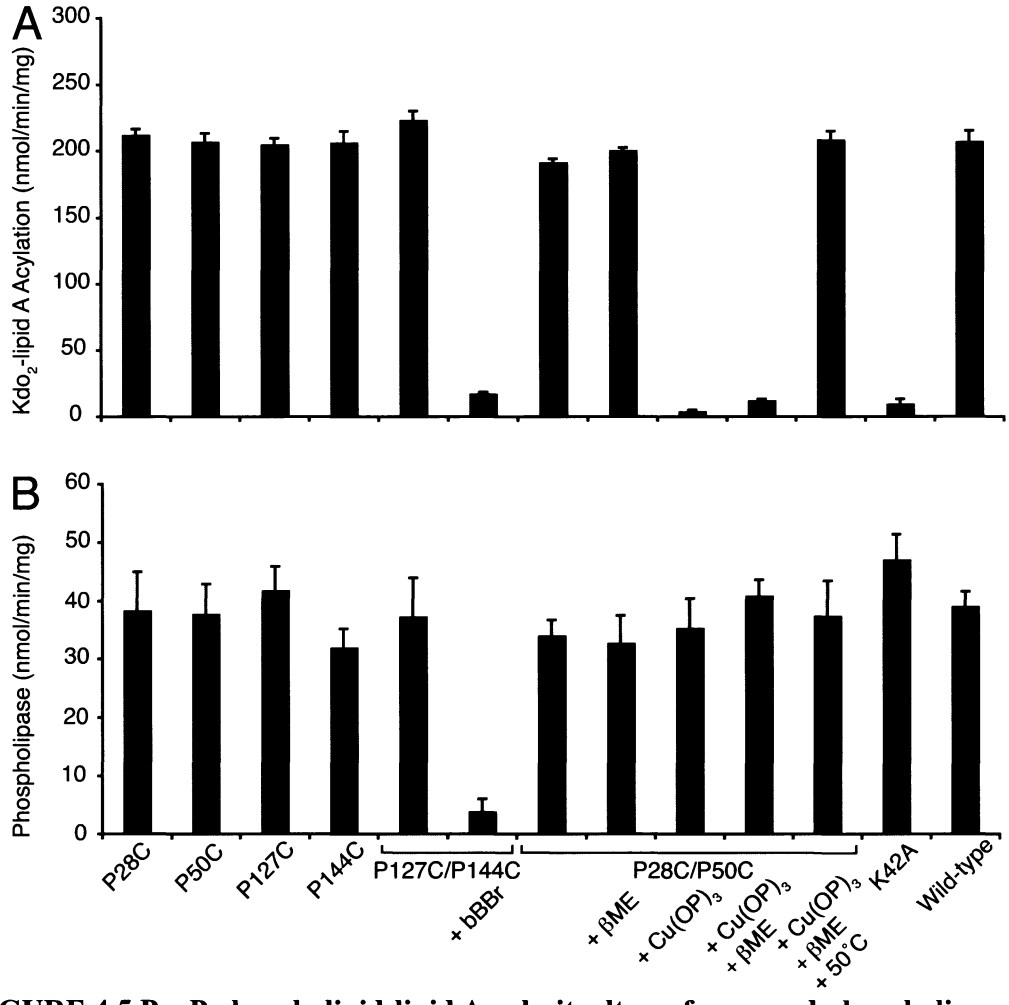

FIGURE 4.5 PagP phospholipid:lipid A palmitoyltransferase and phospholipase activity measured for PagP embrasure and crenel Pro to Cys mutants, before and after embrasure oxidation/reduction and crenel alkylation, and including the Lys42Ala substitution: Specific activity for the palmitoyltransferase reaction is shown in panel (A) and for the phospholipase reaction in panel (B).

amino acid residues were shown previously by mutagenesis to be essential for *in vitro* lipid A palmitoylation (*110*), a subsequent survey of all conserved cell surface-exposed dichroism functional groups identified 11 such residues (unpublished results). Uniquely, the Lys42Ala substitution revealed the only PagP mutant to function *in vitro* as a dedicated phospholipase (Figure 4.5). Lys42Ala PagP was purified and its structural integrity validated here (Tables 1 and 2, Figures 4.4A and 4.4C) because it remains the only known PagP specificity determinant for lipid A recognition. Lys42 localization in the disordered L1 surface loop thus identifies the PagP embrasure as a likely locus for lipid A access to the hydrocarbon ruler.

4.3.4 Lateral Lipid Diffusion Through the PagP Embrasure and Crenel:

The PagP embrasure and crenel double mutants, together with each single Pro→Cys mutant, all display wild type lipid A palmitoylation and phospholipase activity (Figure 4.5). Exposure to βME does not stimulate PagP activity, but closing the embrasure by $Cu(OP)_3$ oxidation eliminates lipid A palmitoylation without affecting phospholipase activity, thus substantiating Lys42Ala mutant predictions. Subsequent exposure of the oxidized embrasure to excess βME fully restores lipid A palmitoylation (Figure 4.5), but only after 50˚C heat treatment needed to reduce the embrasure disulfide bond (Figure 4.2B). Barricading the crenel with bBBr inhibits both lipid A palmitoylation and phospholipase reactions (Figure 4.5) without compromising PagP structure (Figure 4.4 A and 4.4 C), thus implicating the crenel as a locus for

glycerophospholipid discrimination. The retention of slow phospholipid hydrolysis after embrasure closure indicates water molecule access remains unimpeded. PagP does not function as a phospholipase *in vivo* where it is continually exposed to LPS (*176*), but *in vitro* phospholipid hydrolysis implies the existence of a "hydrophilic funnel" to channel water from bulk solvent to the active site as proposed for other membrane-intrinsic lipid enzymes employing a soluble cosubstrate (*180*).

To further validate this model, we took advantage of findings PagP can palmitoylate various non-specific miscible and fatty alcohols *in vitro* (*218*). Addition of methanol, ethanediol, propanediol, or glycerol to the phospholipase reaction results in production of the corresponding palmitoyl esters. The fatty alcohols Triton X-100 and various monoglycerides are also palmitoylated when added to the phospholipase reaction, and the palmitoylation of lysophospholipids can result in palmitoyl group exchange between different classes of phospholipids *in vitro* (*218*). We reasoned embrasure closure should distinguish palmitoyl group transfer to acceptors arriving via a lateral route versus a hydrophilic funnel. Using glycerol and its palmitoylated or myristoylated monoglycerides as acceptors, we demonstrate palmitoylation of only the fatty alcohols is completely impeded by embrasure closure (Figure 4.6). Palmitoyl group transfer to water remains unimpeded, whereas embrasure closure renders glycerol just 2-fold less effective as a palmitoyl acceptor. These findings indicate phospholipid donor palmitoyl groups within the PagP hydrocarbon ruler can be resolved slowly by water or miscible alcohols arriving from the bulk solvent phase *in vitro*, but either lipid A or non-specific fatty

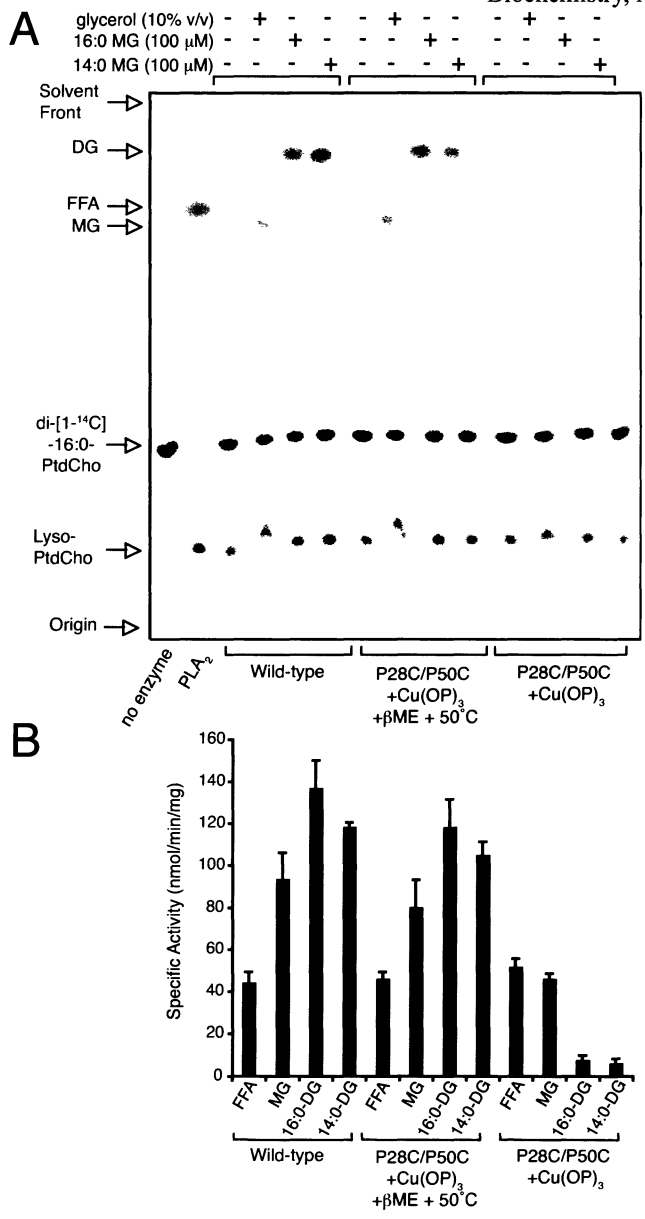

FIGURE 4.6 PagP palmitoylation of non-specific miscible and lipid alcohols: (A) Representative TLC of the phospholipase reaction supplemented with and without glycerol and its palmitoylated (16:0) or myristoylated (14:0) monoglycerides (MG). Phospholipid hydrolysis liberates palmitate as the free fatty acid (FFA), which is indicated by reaction with phospholipase A2 (PLA2). Addition of glycerol generates 16:0 MG, whereas addition of 16:0 or 14:0 MG generates the corresponding diglycerides (DG). (B) Specific activities for phospholipid hydrolysis (FFA), glycerol palmitoylation (MG), and monoglyceride palmitoylation (16:0 DG and 14:0 DG) for wild-type PagP are compared to PagP having the embrasure either oxidized or reduced.

alcohol acceptors are more rapidly engaged via a strict lateral route taken through the embrasure.

4.4 Discussion:

Amphiphilic detergents possessing both polar and non-polar functional groups often disrupt lipid bilayer structural features essential for the normal function of integral membrane lipid metabolic enzymes, and many such enzymes prove recalcitrant toward functional detergent extraction from the membrane environment. The PagP detergent micellar enzymatic assay faithfully replicates key conditions *in vivo* because enzyme activity is triggered by lipid redistribution associated with a breach in the outer membrane permeability barrier (*40, 44, 123, 150, 206*). Detergent micelles clearly accommodate PagP and both of its natural substrates (*70, 126*), but the enzyme's exposure to local polar and non-polar environments also matches those anticipated to exist *in vivo*. Shielding from air oxidation only in the folded state of introduced cysteine sulfhydryl groups spanning the embrasure and crenel (Figure 4.3), combined with blue-shifting of the crenel-spanning bimane fluorophore (Figure 4.2D), indicate these membrane-intrinsic regions experience a hydrophobic environment also in the micellar milieu. The expected membrane surface exposure of the active site also faithfully anticipates the bulk solvent access revealed by slow phospholipid hydrolysis observed *in vitro* (Figures 4.5 and 4.6). Since bacterial protein secretion is generally coupled to disulfide redox enzymology (*219*), periplasmic chaperones (*220*), outer membrane β-

barrel docking machinery (*221, 222*), and dedicated outer membrane lipid transport systems (*206, 223*), addressing PagP lateral lipid diffusion *in vivo* is a considerable challenge. By reconstituting this basic process *in vitro*, we now provide a first validation for the PagP crenel and embrasure as *bona fide* structural adaptations to control lateral-lipid diffusion between the outer membrane external leaflet and a β-barrel interior hydrocarbon ruler.

We propose a model where lateral palmitoyl group diffusion within the PagP hydrocarbon ruler is gated during phospholipid entry via the crenel and during lipid A egress via the embrasure (Figure 4.7). A non-sequential acyl-enzyme mechanism would be obligatory if only a single lipid gateway existed for PagP hydrocarbon ruler access, but the presence of two such gateways indicates a sequential ternary complex mechanism is equally plausible (*224*). The crenel and embrasure represent distinctly different regions in terms of protein structural dynamics, perhaps reflecting known differences in PagP substrate selectivity. The PagP crenel is highly structured (*111*) and incapable of spanning the ~5 Å gap with an intramolecular disulfide bond (Figures 4.1 B and 4.2), but we also expect it to be stringently selective for a phospholipid palmitoyl group, with *sn*-1-palmitoyl phosphatidic acid providing the simplest of the known efficient palmitoyl donors (*70*) (Figure 4.5). In contrast, the PagP embrasure is highly dynamic (*110, 125*) and remarkably unselective insofar as it recognizes multiple lipid A substructures *in vitro* including the most basic diacylglucosamine-1-phosphate moiety (lipid X) (*70*). This dynamic embrasure region is also clearly responsible for the palmitoylation of non-

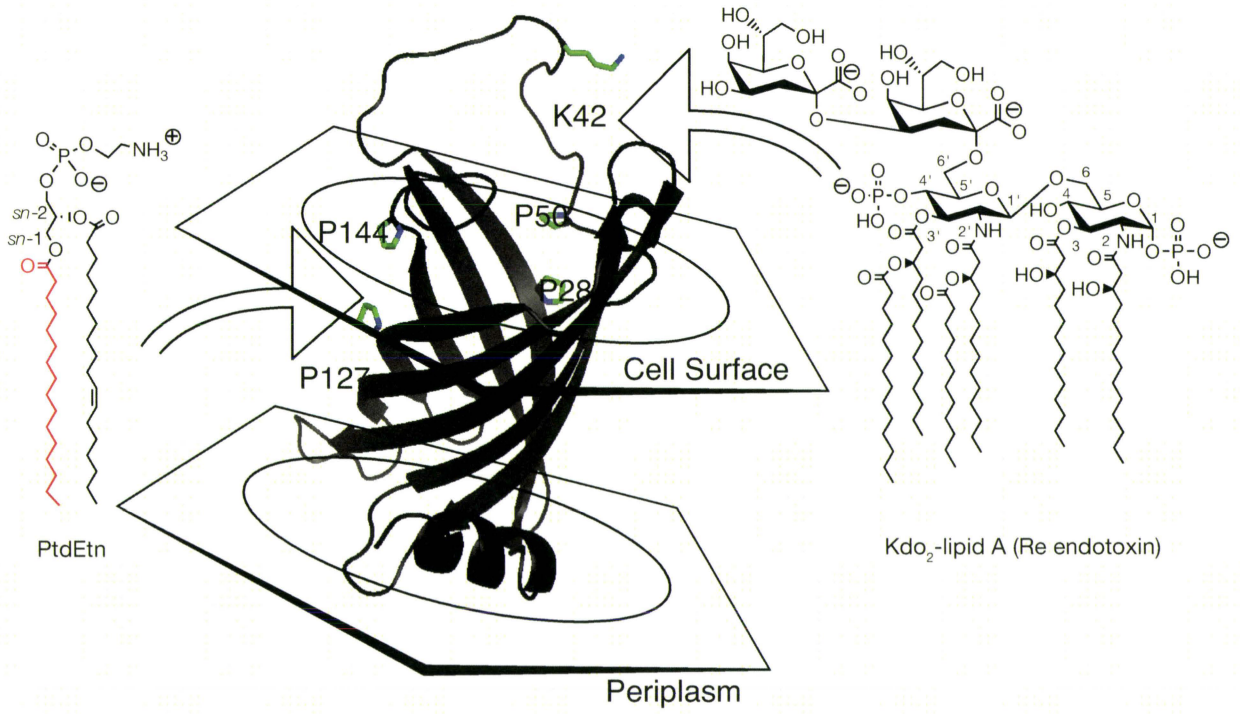

FIGURE 4.7 A model for lateral lipid diffusion between the PagP β-barrel interior hydrocarbon ruler and the outer membrane external leaflet: Donor phospholipid palmitoyl groups diffuse laterally into the hydrocarbon ruler through the crenel flanked by prolines 144 and 127. Lipid A acceptors diffuse laterally into the hydrocarbon ruler through the embrasure flanked by prolines 28 and 50. The ε-amino group of lysine 42 in the dynamic L1 surface loop flanking the embrasure likely provides a single lipid A specificity determinant capable of neutralizing negative charge on Kdo_2-lipid A. Disordered residues including lysine 42 in the L1 loop region were introduced and energy minimized in the PagP model shown.

specific lipid alcohols (Figure 4.6). After a conformational ordering in and around the L1 loop induces a catalytically competent PagP state (*125*), the single embrasure-spanning Lys42 might suffice as sole specificity determinant to neutralize negative charge on Kdo_2-lipid A (Figure 4.7).

We have previously argued the LPS-replete outer membrane external leaflet represents an environment of too low selective pressure to demand stringent lipid A selectivity from PagP (*176*), but selective pressure likely acts to protect the phospholipid palmitoyl donor from unproductive hydrolysis. An induced-fit substrate binding mechanism might provide not only a means to shield substrates from water molecules (*225*), but also a means to gate lipid access between the membrane and hydrocarbon ruler. The aforementioned ordering of residues in and around the L1 surface loop only occurs in a detergent system known to support PagP activity (*125*). The crystal structure in the inhibitory LDAO (*111*) likely represents a latent state before PagP is triggered into an active state by perturbations to outer membrane lipid asymmetry (*181*). Additional structural data derived from PagP in novel detergent systems (*217*) might eventually reveal the mechanisms by which lipid substrates and/or their analogs induce the conformational changes governing both catalysis and hydrocarbon ruler gating.

Although highly conserved or invariant proline residues flank the PagP crenel and embrasure, we successfully managed to fold all of the Pro→Cys mutants *in vitro* (Figure 4.3). Proline residues are clearly not essential to enable PagP hydrocarbon ruler lateral

lipid diffusion, but the diminished transmembrane β-strand hydrogen bonding is likely critical. The outer membrane β-barrel transporter FadL channels long-chain fatty acids from bulk solvent directly into the outer membrane external leaflet through another embrasure structure displaying diminished hydrogen bonding, but achieved instead by a proline-independent shifting of the local β-strand registration (*226*). Under control of the β-barrel protein LptD (Imp) (*223*), lateral lipid diffusion during LPS delivery into the outer membrane external leaflet could conceivably require channeling through a crenel structure. Our findings thus have general significance for outer membrane biogenesis processes in bacteria, chloroplasts and mitochondria (*89, 227*).

5.0: PagP Crystallized from SDS/Cosolvent Reveals the Route for Phospholipid Access to the Hydrocarbon Ruler

Work presented in this chapter was submitted to the peer-reviewed journal *Structure* in 2009: Jose Antonio Cuesta-Seijo, Chris Neale, M. Adil Khan, Joel Moktar, Christopher Tran, Russell E. Bishop, Régis Pomès, Gilbert G. Privé. "PagP crystallized from SDS/cosolvent reveals the route for phospholipid access to the hydrocarbon ruler". The manuscript is currently under revisions. Jose Antonio Cuesta-Seijo in the laboratory of Gilbert Privé (U of T) performed the crystallography studies, while Chris Neale in the laboratory of Dr. Régis Pomès (U of T) conducted the MD simulations. With the help of Joel Moktar, a fourth year undergraduate research student in the laboratory of Dr. Russell Bishop, I constructed, purified and analyzed the mutants implicated in phospholipid gating. Therefore our contributions are limited to figure 5.7.

5.1 Introduction:

PagP is an integral membrane enzyme that transfers a palmitoyl group from the *sn*-1 position of a glycerophospholipid to the lipid A (endotoxin) moiety of lipopolysaccharide (LPS). The donor phospholipid, LPS acceptor and enzyme active site are all located in the external-facing leaflet of the Gram negative outer membrane (Fig. 5.S1) (*70*). The products of the transacylation reaction are a lysophospholipid and a hepta-acylated form of lipid A with attenuated endotoxic properties. The enzyme provides bacterial resistance to host immune defenses involving cationic antimicrobial peptides, attenuates the host inflammatory response to infection triggered by the TLR4 pathway, and is a virulence determinant to infection by certain pathogens (*176*). The

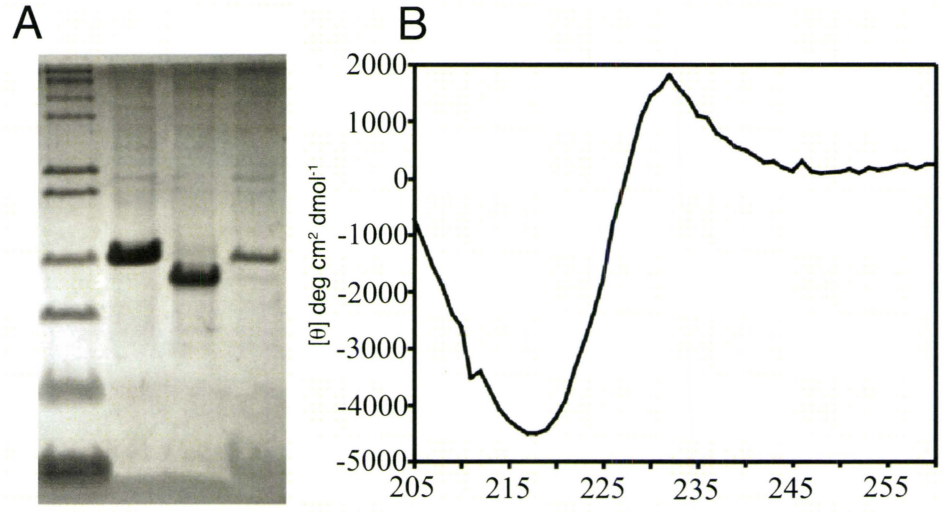

Figure 5.S1 Refolding of PagP in SDS/MPD:
Left. 18% SDS-PAGE of the protein crystallization stock. From left to right, Mark 12 molecular weight standards, 1 μL of the crystallization protein stock, 5 μg of PagP unfolded in SDS alone and 2 μg of PagP folded in LDAO. The gel was stained with Simply Blue (Invtrogen).
Right. CD spectrum of the sample prior to the concentration step (see Methods). The maximum at 232 nm is indicative of the correct formation of the hydrocarbon ruler, the minimum at 218 nm of the β–sheet structure.

protein belongs to the PhoP/PhoQ-regulon, which controls the expression of bacterial genes needed for covalent lipid A modifications induced during infection by antimicrobial peptides (*66*). PagP is a potential target for the development of anti-infective agents and a tool for the synthesis of lipid A-based vaccine adjuvants and endotoxin antagonists (*176*).

As an enzyme of lipid metabolism, PagP is remarkable in its ability to select a 16-carbon saturated palmitate chain to the exclusion of all other acyl chain types present within the membrane pool of glycerophospholipids (*70*). The enzyme is an 8-stranded β-barrel (*110*) whose catalytic mechanism likely involves cycling between two dynamically distinct states (*125*), although the details of the reaction remain unknown. The crystal structure of PagP determined in the presence of the non-denaturing detergent lauryldimethylamine oxide (LDAO) (*111*) revealed a detergent molecule bound in the interior of the barrel with the head group pointing towards the extracellular side of PagP. This detergent molecule marks the position of the hydrocarbon ruler, which selects acyl chains in the glycerophospholipid substrate with methylene unit resolution (*111, 126*).

Here we describe the 1.4 Å crystal structure of PagP in sodium dodecylsulfate (SDS) / 2-methyl-2,4-pentanediol (MPD). SDS is strongly denaturing anionic detergent that widely used in biochemical applications to solubilise and unfold proteins (*228*). It is effective at millimolar concentrations and induces non-native structures with high proportions of α-helix, although its mechanism of action is not well understood (*217,*

229-238). MPD is best known in biochemical applications as a protein precipitant in crystallization experiments (*238*). We have recently reported that certain water miscible amphipathic alcohols, notably MPD, can modulate the protein-binding properties of SDS so that the detergent behaves essentially as a "gentle" non-denaturing detergent (*217, 239*). Concentrations of 1 to 2 M MPD protected both soluble and integral membrane proteins from SDS denaturation, and several proteins, including PagP, could be induced to fold from the SDS denatured state upon addition of MPD (*217*). PagP in SDS/MPD could be reversibly denatured and refolded upon heating and cooling cycles, confirming that the folded protein is at thermodynamic equilibrium in this detergent/cosolvent system. We now report that the SDS/MPD detergent system is suitable for the crystallization of integral membrane proteins.

The structure of PagP in SDS/MPD compares favourably with the previously determined crystal structure in LDAO (*111*) and with the solution NMR structures in dodecylphosphocoline (DPC) and β-D-octyl-glucoside (OG) (*110*), but differs in the conformation of the four extracellular loops. These differences indicate regions of flexibility in the protein that may be important for substrate binding and catalysis. Like the parapet that surrounds a turret of a medieval castle, the PagP β–barrel contains indentations known as crenels or crenellations, which could provide gateways for lateral lipid access between the membrane external leaflet and the hydrocarbon ruler. Diminished β–strand hydrogen bonding has identified potential crenels between strands F and G and at the β–bulge between strands A and B. Molecular dynamics simulations and

mutagenesis support a phospholipid access route via the F/G crenel coupled to a movement of extracellular loop L4. Tyrosine 147 in that loop is revealed as a potential gating residue controlling lateral access of phospholipids to the interior hydrocarbon ruler pocket of PagP.

5.2 Materials and Methods:

5.2.1 Refolding:

Washed PagP inclusion bodies (13 mg) were dissolved in 20 ml of 1% SDS, 1M MPD and 10 mM Tris-HCl, pH 8. The solution was heated for two minutes in a boiling water bath and allowed to cool down by letting the water bath slowly return to room temperature (approximately 2 h). Refolding proceeded virtually to completion as judged by a gel shift assay (*217*) and the presence of a positive exciton band in the CD spectra at 232 nm (*126*). The PagP solution was concentrated in a 10-kDa MWCO spin filter (Amicon) resulting in 1.5 ml of a solution with 5.8 mg/ml of PagP. This solution was used directly for crystallization. The protein remained folded for months under these conditions as judged by the band shift assay. The SDS concentration in this stock solution was 3.8% w/v by a SDS assay (*240*).

5.2.2 Crystallization:

Crystals were grown by the hanging drop method. The reservoir solution consisted of 500 µL of 1M MPD, 0.1 M Na_3Citrate buffer pH 5.6, 1.7 M Li_2SO_4 and

0.3M $(NH_4)_2SO_4$. This mixture separates spontaneously into two phases, a hanging drop was prepared by mixing 1 μL of the lower phase (rich in salt) with 1.5 μL of the PagP stock. The hanging drop was a homogeneous mixture at first but upon equilibration with the reservoir it also separated into two phases.

Crystals were obtained only in one of the two phases in the hanging drop. Crystals were allowed to grow at 20° C for three months and then transferred to a temperature of 4° C for four days. A crystal of dimensions 0.4 mm x 0.2 mm x 0.2 mm was frozen directly in its mother liquor by plunging into liquid nitrogen.

5.2.3 Molecular Dynamics Simulations:

The simulation system consisted of a solvated POPC bilayer in which 2 PagP molecules were embedded. The binding pocket of PagP was empty for steered insertion simulations and contained a DP detergent substrate analogue for steered extraction simulations. Construction and 20-50 ns equilibration of the simulation systems are described in the Supplementary Methods.

For each molecule of PagP in each starting conformation, a separate steered insertion simulation of 100 ps was carried out in which the *sn-1* chain of a single phospholipid in the upper bilayer leaflet was forced toward the PagP binding pocket with a force of constant magnitude. Phospholipids were targeted for insertion if at least one of the distal 13 carbons in the *sn-1* chain was within 3 nm of the protein. Similarly, steered

extraction simulations forced the DP molecule away from its initial position in the PagP binding pocket. Additional details are provided in the Supplementary Methods.

The Berger parameters (*241*) for POPC (*242*) and the OPLSAA parameters (*243*) for PagP were properly combined for the GROMACS simulation package (*244*) using the half-ε double-pairlist method that we introduce (Fig. 5.S4). Briefly, the ε values of the 1-4 Lennard-Jones parameters of the lipids were multiplied by an additional factor of 0.5 in the pairtypes section and the list of 1-4 interactions in the pairs section was duplicated. The regular OPLSAA combination rules were then applied.

5.2.4 Supplemental Methods:

5.2.4.1 Overexpression and Purification of PagP inclusion bodies:

E. coli PagPΔHΔS was cloned by the quick change procedure (Stratagene) from the plasmids described elsewhere (*110*) and expressed into inclusion bodies. BL21(DE3) cells were transformed and cell cultures were grown in LB media at 20 °C until the OD_{600} reached 0.6. Expression was induced in by adding IPTG to a concentration of 0.2 mM and growth was allowed to proceed for five hours at 37 °C. Expression of PagP is driven into inclusion bodies under these conditions. The total production of protein as inclusion bodies was approximately 1 g of PagP from 6 L of LB media in shaker flasks. The cells were lysed in an Emulsiflex cell disrupter (Avestin) and inclusion bodies were collected by centrifugation of the lysed cells. The inclusion bodies were washed three times with a buffer consisting of 2% Triton X-100 and 10 mM Tris-HCl, pH 8 followed by

centrifugation at 40000 g. The final white pellet was dissolved in 6M Guanidine hydrochloride and 10 mM Tris-HCl, pH 8 for storage. A 1 ml aliquot containing 13 mg of protein was precipitated by dilution into 25 ml of 10 mM Tris-HCl, pH 8. The precipitate was washed three times by centrifugation and resuspension in 25 mls of the same buffer. No further purification steps were carried on the sample that was used for crystallization.

5.2.4.2 Structure solution and refinement:

The structure was solved by molecular replacement using EPMR (*245*) with the structure of PagP in LDAO, PDB accession code 1THQ (*111*). The structure was refined and extended using alternate cycles of manual building in Coot (*246*) and refinement in Refmac5 (*247*) using the CCP4i graphical interface (*248*) with implicit riding hydrogens. The stereochemical quality of the final models was assessed with MolProbity (*249*) and Procheck (*250*). Restraints for the MPD molecules were modified from the standard Refmac 5 dictionary by removing the torsion and chiral restraints. Restraints for the SDS molecules were adapted from Coiro and Mazza (*251*). Least squares superpositions between the different models were made with the program LSQKab (*252*). Only the α-carbons of the residues that are in β-strands in the SDS/MPD structure as indicated by Procheck were used for the superpositions (88 atoms in total). The final model had the following Ramachandran statistics for the non-glycine and non-proline residues (*250*): 88.1% in the favoured regions and 9.5% in the allowed regions.

5.2.4.3 Cloning, Expression and Refolding of the PagP Mutants:

Wild-type PagP, and the Tyr147Ala and Tyr147Phe mutants were constructed in the pETCrcAHΔS vector, with a *C*-terminal His_6 tag and no signal peptide (*110*) using the QuikChange protocol (Stratagene) and purified as described previously (*126*). The proteins were verified by electrospray ionization mass spectrometry. Wild type PagP had an observed mass of 20175.3 ± 0.6 Da (theoretical mass 20175.5). Tyr147Ala PagP had an observed mass of 20082.1 ± 0.9 Da (theoretical mass 20083.4). Tyr147Phe PagP had an observed mass of 20159.5 ± 0.7 Da (theoretical mass 20159.5).

5.2.4.4 Refolding of the PagP mutants for enzymatic assays:

The protein samples dissolved in 5ml of 6M Gdn-HCl, 10mM Tris-HCl, pH 8.0 were diluted 10 fold into 10mM Tris-HCl pH 8.0, 0.1% LDAO. These samples were allowed to stir overnight at 4 °C and applied to a NiNTA column. The samples were prepared from the columns as previously described in Khan *et al.*, 2007. The concentration of the protein was determined using an extinction coefficient, $\in_{280}$ of 80,463 M^{-1} cm^{-1} for wild-type PagP and 78,211 M^{-1} cm^{-1} for PagPY147A and PagPY147F. The extinction coefficients were determined experimentally using the Edelhoch method (*152*).

5.2.4.5 Electrospray ionization mass spectrometry (ESI-MS):

Precipitated protein samples were dissolved in a solution of 50:50 1% formic acid/acetonitrile just prior to ESI-MS. The sample concentration was maintained at 1 ng/μl and was injected directly onto a Waters/Micromass Q-Tof Ultima Global (a quadrupole time-of-flight) mass spectrometer. The spectra were reconstructed using MassLynx 4.0, MaxEnt 1 module.

5.2.4.6 Enzymatic Assays:

Kdo_2-lipid A and synthetic dipalmitoylphosphatidylcholine (DPPC) were obtained from Avanti Polar Lipids. A detailed description of the preparation of ^{32}P-labelled Kdo_2-lipid A, and thin layer chromatography (TLC)-based assays for lipid A palmitoyltransferase activity were previously described (*126*).

Phospholipase assays were carried out in a volume of 25μl, with sufficient dipalmitoyl-1-^{14}C-DPPC to achieve a final concentration of 20μM (4000 cpm/μl). The lipid was dried under a stream of $N_2(g)$ and was dissolved in 22.5μL reaction buffer containing 0.1M Tris-HCl, pH 8, 10mM EDTA and 0.25% β-D-dodecyl maltoside. The reactions were started by adding 2.5μl of PagP and were conducted at 30°C. Reactions were stopped by adding 12.5μl of the reaction mixture to 22.5μl of a 1:1 $CHCl_3$/MeOH. This generated a 2-phase solution, from which 5μl of the lower phase was spotted on a silica gel 60 plate. The experiment was carried out over five hours, with the reaction being spotted once every hour. TLC plates were developed in a $CHCl_3$:MeOH:H_2O

(65:25:4; v/v) solvent system, which was equilibrated in sealed glass tank for four hours. The plates were exposed overnight to a PhosphorImager screen and developed the following day with a Molecular Dynamics Typhoon 9200 PhosphorImager.

5.2.4.7 Circular Dichroism (CD) spectroscopy:

Samples to be analyzed by CD were maintained at a concentration of 0.3 mg/ml in 10 mM Tris-HCl pH 8.0, 0.1% LDAO and were analyzed using a cuvette of 1mm path length. The samples were analyzed using a Aviv 215 spectropolarimeter which was linked to a Peltier device Merlin Series M25 for temperature control. For each sample 3 accumulations were averaged at a data pitch of 1 nm and a scanning speed of 10 nm/min. The temperature was maintained at 25 °C and data were obtained from 200 nm to 260 nm. Thermal denaturation profiles were obtained by heating the samples from 20 °C to 100 °C at 218 nm with a temperature slope of 1 °C/min and a response time of 3 seconds.

5.2.4.8 Molecular Dynamics Simulations:

Simulations were conducted with version 3.3.1 of the GROMACS simulation package (*244*), modified to apply a biasing force of constant magnitude. The Berger parameters (*241*) for POPC (*242*) and the OPLSAA parameters (*243*) for PagP were combined using the half-ε double-pairlist method that we introduce. The water model was TIP4P (*253*). Periodic boundary conditions were enforced via a rectangular unit cell. Lennard–Jones interactions were evaluated using a group-based twin-range cutoff (*254*) calculated every step for separation distances less than 0.9 nm and every ten steps for

distances between 0.9 and 1.4 nm, when the nonbonded list was updated. Coulomb interactions were calculated using the smooth particle-mesh Ewald method (*255, 256*) with a real-space cutoff of 0.9 nm and a Fourier grid spacing of 0.12 nm. Simulation in the *NpT* ensemble was achieved by isotropic coupling to a Berendsen barostat (*257*) at 1 bar with a coupling constant of 4 ps and separate coupling of the solute and the solvent to Berendsen thermostats (*257*) at 310 K with coupling constants of 0.1 ps. Bonds involving hydrogen were constrained with SETTLE (*258*) and LINCS (*259*) for solvent and solute, respectively. The integration time step was 2 fs.

POPC bilayer construction. A POPC molecule was constructed based on the 1,2-dilauryl-*sn*-glycero-3-phosphoethanolamine (DLPE) crystal structure of Elder *et al.* (*260*) by replacing amide hydrogens with methyls and extending the acyl chains in the *trans* conformation. A POPC bilayer with 160 lipids per leaflet measuring 7.77 nm by 7.96 nm in the bilayer plane was then constructed based on the DLPE crystal symmetry (*260*). The rectangular unit cell was set to 20 nm normal to the plane of the bilayer and solvated with 29,953 TIP4P water molecules. The crystal symmetry was broken using a protocol derived from Takaoka *et al.* (*261*). However, we found that a preliminary low temperature simulation was required to prevent leaflet separation during the 510 K pulse, and that a large excess of water normal to the bilayer was required to accommodate undulations that were induced when, at 510 K, the area per lipid increased more quickly than could be accommodated in the bilayer plane by pressure coupling. Our protocol was 0-1 ns: 310 K, 1-1.1 ns: gradual heating to 510 K, 1.1-1.17 ns: gradual cooling to 360 K,

1.17-5.17 ns: 360 K, 5.17-40 ns: 323 K. At this point, the system dimensions were 10.4 nm by 9.8 nm in the bilayer plane and 12 nm along the bilayer normal. The undulations had reduced significantly by this point and we reduced the unit cell to 8 nm along the bilayer normal, excluding a large number of water molecules. This smaller system was simulated for an additional 50 ns at 310 K. A time-dependent analysis of the order parameters of the acyl chains indicated that the bulk properties of the bilayer had converged after 25 ns (not shown). The coordinates of this bilayer at 25 ns were used for the study of PagP.

PagP all-atom modelling. The available coordinates for PagP, including crystal waters within 5 Å, were taken from the crystal structure presented in this article. We then constructed 1000 candidate models of L1 loop residues 38-45 as random coil using the program Loopy (*262*), keeping the two loop conformations of lowest colony energy. To both of these structures, we modelled the missing sidechains for residues 4, 35, 36, 46, 47, 146, 147, and 148 with SCWRL 3.0 (*263*), during which we found it necessary to manually direct Y46 χ1 to -120° to avoid atomic overlap. A single molecule of dodecyl phosphate (DP) was placed in the PagP binding pocket as a dodecyl sulphate analogue. Parameters for DP were constructed based on those of dodecylphosphocholine (DPC) (*264*) by removing the choline group and adding a partial charge of 0.1 to each of the four oxygens.

Composite system creation. Two copies of PagP having different L1 loop conformations were overlaid on the POPC bilayer with the principle axis of each β-barrel oriented 15° to the bilayer normal, bringing the L3 loop toward the bilayer, in order to align the exposed hydrophobic/hydrophilic surfaces with the hydrophobicity profile of the bilayer as suggested by Ahn *et al. (111)*. To reduce any anisotropic effects that the mobile loop regions of the protein may inflict upon the relative surface tension of each leaflet, the proteins were inserted anti-parallel to one another. Any phospholipid within 0.2 Å was removed, additionally removing a minimal number of phospholipids while ensuring that each leaflet contained the same amount. A surface representation of each protein was then constructed using MSMS (*265*) and PagP-shaped holes were made in the bilayer according to the protocol of Faraldo-Gomez *et al. (266)* in 3 segments of 20 ps each while applying a position restraint along the bilayer normal with a force constant of 1000 kJ mol^{-1} nm^{-2} on phosphorous atoms of all lipids. The strength of the hole-making force was 10, 100, and 500 kJ mol^{-1} nm^{-2} in the first, second and third segments, respectively. The composite system, composed of a solvated POPC bilayer, two protein molecules and associated DP and crystal waters, was neutralized with 8 Na^{+} ions and simulated for 50 ns at 310 K, with the position of protein and DP heavy atoms restrained during the first 5 ns. The rate at which the backbone RMSD to the crystal structure increased was significantly attenuated by 15 ns (not shown), and hence the final 30 ns of this simulation was used to generate conformations with which to initiate non-equilibrium steered MD simulations. Starting configurations for steered insertion of the

sn-1 chain of a POPC molecule into the PagP binding pocket and steered extraction of a DP molecule from the PagP binding pocket were taken at 2 ns and 200 ps intervals, respectively. Prior to the initiation of steered insertion simulations, both DP molecules and two Na^+ ions were removed from the system.

Steered MD simulations. Steered insertion simulations employed a biasing force of the form

$$F = \frac{-K}{\left|x^{ligand} - x^{ref}\right|}\left(x^{ligand} - x^{ref}\right) \quad (1)$$

where x_{ligand} represents the current position of the centre of mass of the distal 13 carbons in the *sn-1* chain of the selected phospholipid and x_{ref} represents the reference position calculated as

$$x^{ref} = x^{PagP} + x_0^{DP} - x_0^{PagP} \quad (2)$$

in which the current position of the centre of mass of the PagP β-barrel, x_{PagP}, is offset by the vector from the initial center of mass of the PagP β-barrel, $x_{PagP,0}$, to the initial centre of mass of the DP molecule that was removed from the binding pocket, $x_{DP,0}$. Only biasing force components in the bilayer plane were applied. The constant K was set to 500 kJ mol^{-1} nm^2 such that the targeted *sn-1* chain is attracted to the PagP binding pocket. For steered extraction simulations, x^{ligand} represents the current position of the centre of

mass of DP and the constant K was set to -500 kJ mol^{-1} nm^2 such that DP is repelled from its initial position.

Half-ε double-pairlist combination rules for Berger lipids and OPLSAA proteins.

During MD simulation, non-bonded interactions between atoms that are separated by exactly three bonds (the so-called 1-4 interactions) are often scaled by some factor from their full strength. This scaling factor differs amongst forcefields and thus the simultaneous use of multiple forcefields requires special attention to ensure that the 1-4 interactions of all molecules are properly scaled. Indeed, combination of the Berger lipid parameters and the OPLSAA protein parameters within the molecular dynamics simulation package GROMACS is complicated by the inability to specify unique parameters for Coulombic 1-4 interactions. It is this deficiency that led Tieleman *et al.* (*267*) to reparametrize the dihedral angle energy functions of the Berger lipids such that the nonbonded component of the 1-4 interactions is accounted for by the new dihedral potentials. While the published reparametrization of dihedral parameters appears to be a valid solution, one would like to avoid reparametrizing every lipid and detergent for which parameters have already been generated using this lipid forcefield. To this end, we suggest a simple method that can be used within GROMACS to combine the above-mentioned forcefields while properly scaling the 1-4 nonbonded interactions of each that is at once simpler and more inclusive. Our method makes use of the fact that (i) it is possible to specify a unique value for the LJ component of the 1-4 interactions, and (ii) the scaling factor for the Coulombic component of the 1-4 interactions of one forcefield

is an integer multiple of the scaling factor used in the other forcefield (1.0 for Berger lipids and 0.5 for OPLSAA). Specifically, in the GROMACS topology file, the ε values of the 1-4 LJ parameters of the lipids are multiplied by an additional factor of 0.5 in the pairtypes section and the list of 1-4 interactions in the pairs section is duplicated. The regular OPLSAA combination rules are then applied. In this way, the LJ and Coulombic 1-4 interactions are both cut in half and then included twice for the lipids, yielding properly scaled 1-4 interactions for both Berger lipids and OPLSAA protein. We refer to this as the half-ε double-pairlist method.

The area per lipid obtained from simulating a bilayer composed of 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC), containing 64 lipids per leaflet solvated by 28.5 SPC (*268*) waters per lipid, using a variety of methods to treat the lipid forcefield. Initial DPPC parameters were taken from Tieleman *et al.* (*269*). The area per lipid derived from the direct use of lipid.itp and ffgmx, a forcefield that does not arbitrarily scale 1-4 interactions, is traced by the red line. Similar results are found when the functional form of the LJ variables are converted from C6/C12 to δ/ε while leaving 1-4 interactions as intended (blue line) and when applying the half-ε double-pairlist method under the OPLSAA scaling rules (purple line). However, the area per lipid is significantly reduced when the unmodified δ/ε formulation of LJ interactions is used while scaling lipid 1-4 interactions according to the rules of OPLSAA (blue line). In this set of conditions, the Coulombic 1-4 interactions of lipids are erroneously assigned only half of their full strength.

5.3 Results:

5.3.1 Crystallization of PagP from SDS/MPD:

E. coli PagP was directed to inclusion bodies by overexpressing the protein without a signal sequence (*110, 111*). Washed inclusion bodies were dissolved in a buffer containing 1% SDS and 1 M MPD. The solution was then heated to 100 °C and the protein was refolded by slow cooling. Refolding was quantitative as judged by an SDS-PAGE assay (*217*). Circular dichroism confirmed that the refolded protein was predominantly β-sheet and exhibited a characteristic exciton couplet at 232 nm indicating the correct formation of the hydrocarbon ruler (Fig. 5.S2) (*126, 217*). No chromatographic steps were performed, and the only purification of the protein occurred during the initial preparation of the insoluble inclusion body material. The protein solution was concentrated to 5.8 mg/ml and contained 3.85% w/v SDS (135 mM) as measured by a methylene blue extraction assay (*240*). The critical micelle concentration (cmc) of SDS in 1 M MPD and 10 mM Tris-HCl is approximately 0.6 mM as measured by isothermal titration calorimetry (John Holyoake and GGP, unpublished data), and so the SDS concentration in the final crystallization stock solution was over 250 fold above its cmc.

The protein was crystallized by the traditional hanging drop method by mixing equal volume solutions of the protein stock and a salt-rich reservoir containing 1 M MPD, and allowing vapor-phase equilibration between the protein droplet and the much

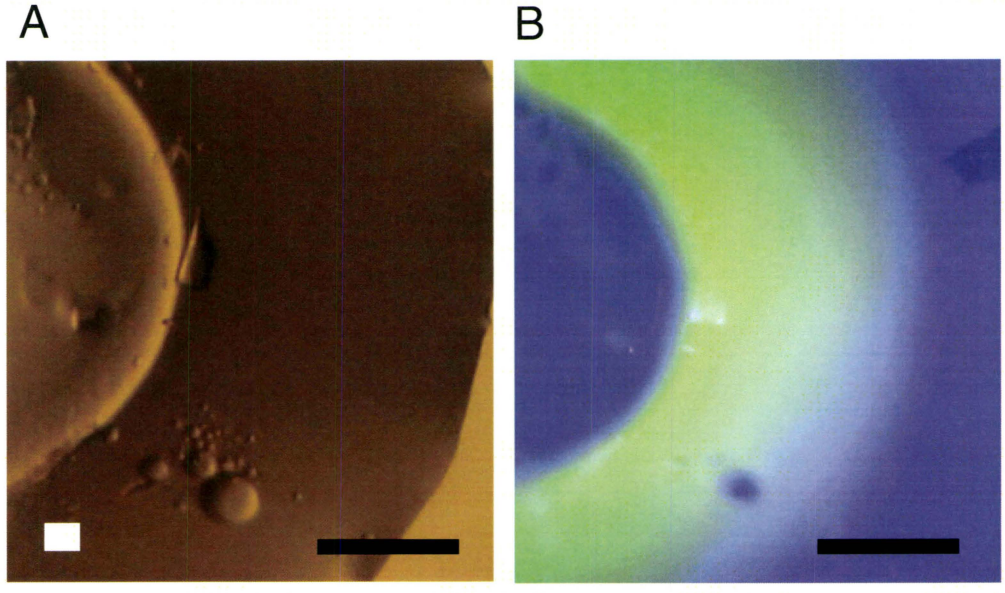

Figure 5.S2 Supplemental Figure:
Left: Phase separation and crystallization of PagP from SDS/MPD.
Right: The same droplet after the addition of the fluorescent lipid NBD-DPPC and viewed under uv light. The scale bar represents 100 μm.

larger volume of the reservoir solution. Two phases formed in the protein droplet as equilibration proceeded. Crystals nucleated at the interface of the two phases and grew exclusively into one of the phases. The fluorescent lipid NBD-DPPC was added to a hanging drop containing crystals and the lipid dye partitioned preferentially into the phase in which the PagP crystals grew, suggesting that the SDS concentration in the phase that supported crystal growth was further enriched in detergent relative to the initial protein stock solution (Fig. 5.S3). Both the protein stock solution and the reservoir solution initially contained 1 M MPD, but because MPD is volatile, the MPD concentration in the phase containing the protein crystals is unknown. However, we located 10 MPD molecules in the final refined structure (see below), indicating a relatively high concentration of the cosolvent in the mother liquor surrounding the crystal.

5.3.2 Structure overview:

Full anisotropic refinement was performed to a resolution of 1.4 Å (Table 5-1). The structure is an eight-stranded β-barrel that spans the membrane region with a leading amphipathic *N*-terminal α-helix that lies at the interface between the periplasm and the outer membrane inner leaflet. The eight strands of the barrel are named A to H and the four extracellular loops are referred to as L1 to L4 (Fig 5.1).

The high resolution of this structure allowed us to trace all of the amino acids in the *N*-terminal α-helix (*216*), which extends from Asn-1 to Trp-17. Loop L1 links strand

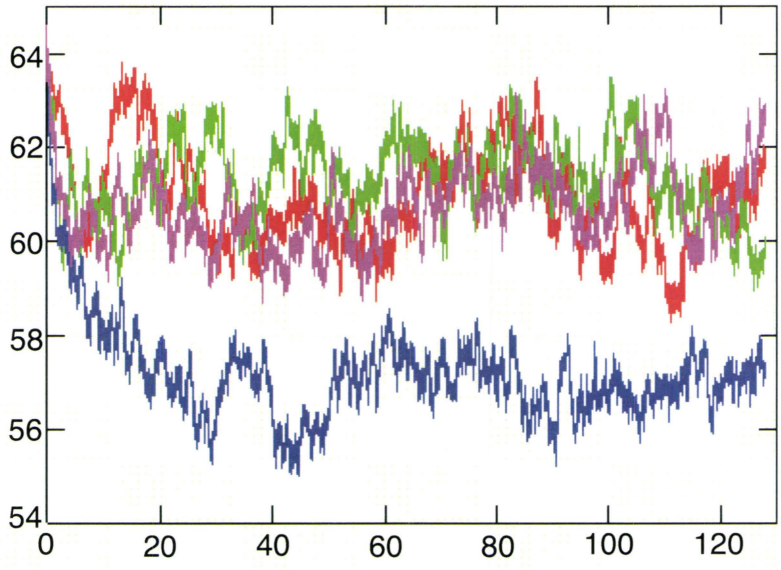

Figure 5.S3 Supplemental Figure:
The area per lipid time series at 323K for simulations employing the forcefield parameters of (red) standard ffgmx, (green) ffgmx converted to ε and δ, (blue) unmodified ε and δ using OPLSAA scaling rules, (purple) the half-ε double-pairlist method for combination with OPLSAA.

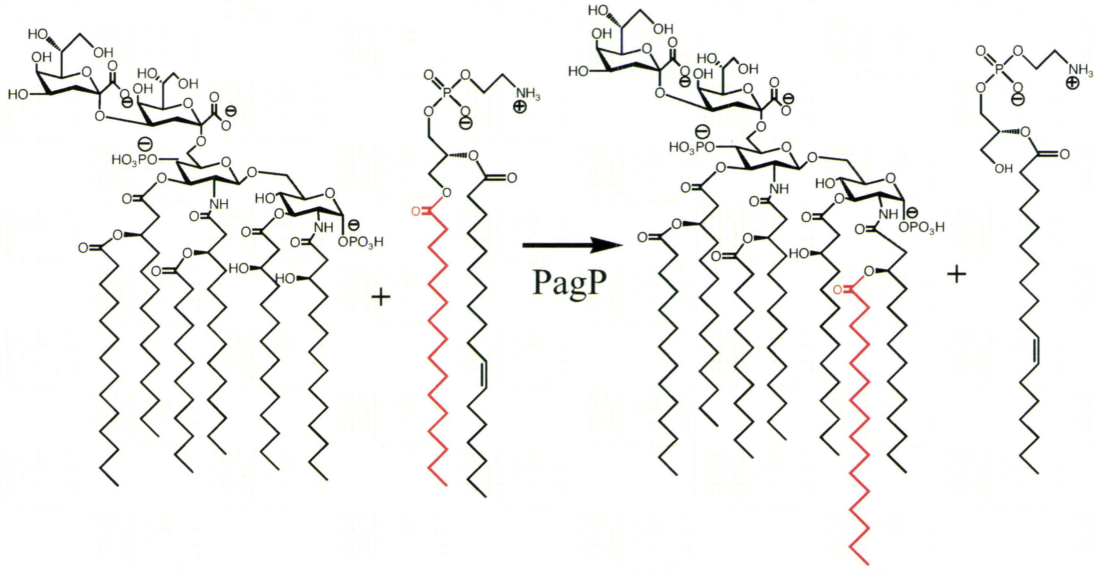

Figure 5.1 Reaction catalyzed by PagP:
PagP catalyzes the transfer of a palmitate chain (shown in red) from the *sn*-1 position of phosphatidylethanolamine (PtdEtn) to lipopolysaccharide (LPS), shown here as Kdo_2-lipid A (Re endotoxin). This lipid consists of lipid A linked at the 6'-position with two units of 3-deoxy-D-*manno*-2-octulosonic acid (Kdo). Lipid A is a β-1',6-linked disaccharide of glucosamine that is acylated with *R*-3-hydroxymyristate chains at the 2, 3, 2' and 3' positions, and phosphorylated at the 1 and 4' positions. Acyloxyacyl linkages with laurate and myristate chains at the 2' and 3'-positions, respectively, provide the constitutive hexa-acylated lipid A, which is a potent endotoxin. A regulated proportion of lipid A in *E. coli* contains a palmitate chain (16:0) in acyloxyacyl linkage at position 2 (in red in the figure), which yields a hepta-acylated molecule with attenuated endotoxic properties.

A to strand B and is partly disordered, and amino acids 38-45 from this loop are not included in the final model. This loop was also not fully modeled in the LDAO crystal structure (*111*) and is highly dynamic in the NMR structures (*110, 125*), indicating that this region is intrinsically disordered, at least in the absence of a lipid bilayer.

The overall barrel structure of PagP in SDS/MPD resembles that of PagP in LDAO by x-ray crystallography (*111*) and in DPC or OG micelles by NMR (*110*). Least squares superposition of the Ca atoms from the β-sheet residues in strands A-H of PagP results in an RMSD of 0.70 Å between the SDS/MPD and the LDAO structures (Fig. 5.1). The structures of the short periplasm-facing loops are also well conserved between the structures, however deviations greater than 1.2 Å are seen in the four extracellular loop regions (Fig. 5.1*B*). These extracellular loops are also the most dynamic regions of the protein as measured by NMR (*110, 125*).

5.3.3 PagP/SDS interactions:

The final model contains six SDS molecules with an average B factor of 47 $Å^2$. In all cases the B factors are higher for the head group than for the aliphatic chain, with three of the head groups modeled at reduced occupancy. Five of the SDS molecules are in contact with the exterior of the protein in the membrane-spanning region of PagP and are bound into crevices in the surface of the molecule (Fig. 5.2*A*). The sulfate head groups do not make any consistent types of interactions with the protein and are consequently less ordered in the crystal. Thus, bound SDS molecules behave like other detergents in which

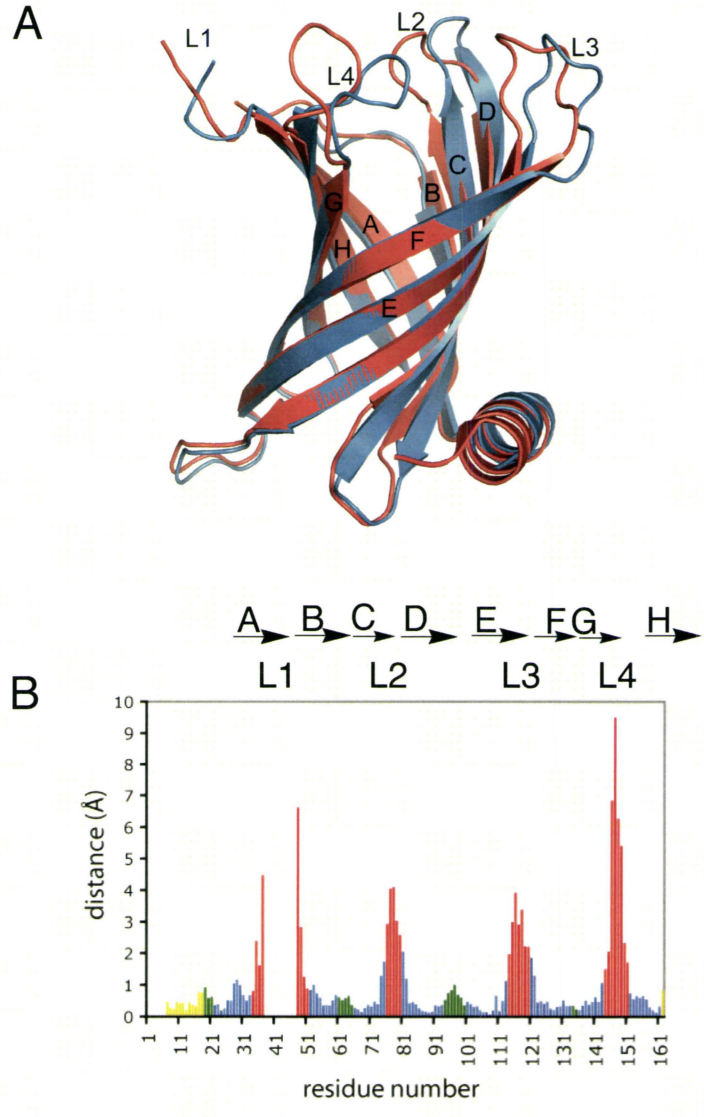

FIGURE 5.2 Differences between the SDS/MPD and the LDAO crystals of PagP: (A) Superposition of the SDS/MPD (red) and LDAO (blue) crystal structures of PagP. Only the β-sheet residues were used in the least-squares fitting. (B) Distances between the α-carbons of the SDS/MPD and the LDAO crystal structures of PagP. Residues in β-strands are in blue, and only those were used for the superposition. The N-terminal helix is shown in yellow, the periplasm-facing loops are in green and the extracellular loops L1-L4 are in red.

binding to the protein is determined primarily by the aliphatic chain (*270, 271*). The structure of PagP crystallized from LDAO contains five ordered detergent molecules, four of which are common with the SDS binding sites (Fig. 5.2*A*). The detergents molecules seen the SDS/MPD structure are thus bound in general detergent binding sites that likely correspond to lipid acyl chain binding sites in the natural outer membrane environment of PagP. No significant SDS-SDS contacts are observed in the crystal and any micellar structures that may be present in the crystal are not ordered.

The remaining SDS molecule is found in the interior of PagP. The acyl chain binding pocket in the center of the β-barrel is occupied by a detergent molecule in a similar fashion to that observed in the LDAO-based structure (*111*). This pocket is part of the hydrocarbon ruler of PagP that selects for palmitate chains in the *sn*-1 position of substrate phospholipids (*126*). The nine distal carbons of the SDS molecule are found to be tightly held and have the lowest atomic B-factors for this molecule. The B factors for carbons nearer to the headgroup increase significantly and the electron density becomes tubular in contrast to the zigzagging pattern of the lower part of the tail. The SDS in the pocket overlaps very well with the LDAO molecule in the terminal eight carbons in the acyl chain, but the paths diverge nearer to the headgroup region. As a result, the anionic SDS head group is located in a different position than the zwitterionic head group of LDAO (Fig. 5.2*B*). The head group of SDS is closer to loop L1 and farther from L2, relative to the position of the LDAO headgroup.

5.3.4 MPD in the crystals:

The final model includes ten MPD molecules (Fig. 5.2*A*). Seven of these are in conformations in which the two hydroxyl groups form intramolecular hydrogen bonds, such that the MPD molecules display a hydrophilic face and a hydrophobic face (*238*). Within the ordered molecules, there are no significant interactions between the MPD and SDS. A previous study also found no interactions between SDS and MPD in lysozyme crystals (*272*), and the present study expands this result to crystals of membrane proteins in which SDS is the stabilizing detergent.

Seven of the ten MPD molecules are bound to the hydrophobic membrane-exposed surface of PagP or to residues marking the transition between the hydrophobic and hydrophilic surfaces. All seven are also in contact with hydrophilic residues of a second molecule and thus play important roles in building the crystal lattice. The MPD molecules are found in two types of environments in the crystal, either filling in small deep pockets in the protein's surface or sandwiched between relatively flat surfaces of two symmetry equivalent molecules. Four of the MPD molecules stack directly with π electronic clouds of the Trp17, Tyr23, Trp89 and Trp117 aromatic side chains of PagP.

5.3.5 Catalytic center and accessibility of the hydrocarbon ruler pocket:

The conserved residues that are necessary for the catalytic activity of PagP include His33 in the extracellular loop L1 and Asp76 and Ser77 in L2 (*110*). In the SDS/MPD structure, His33 is 13.0 Å away from Asp76 and 14.8 Å away from Ser77

(distances between α-carbons; the distance between Oγ of Ser77 and Nε2 of His33 is 8.5 Å). This is closer than in the LDAO structure where these distances are 16.4 Å and 17.0 Å respectively. This difference is mostly due to a rigid-body tilting of loop L2 towards L1, but the residues are still too far apart to form a catalytic center.

An analysis of the SDS crystal structure shows that relative to the LDAO structure, extracellular loop L4 moves away from the barrel interior, resulting in a large gap (a crenel) between strands F and G (Fig. 5.3*A*). This gap opens a path to the top part of the hydrocarbon ruler that was blocked by the side chain of Tyr147 in the LDAO structure (6). A smaller crenel is present between strands B and C (Fig. 5.3*B*). As a result, we sought to further analyze possible access routes into and out of the barrel interior.

We analyzed the possible acyl chain entry and exit routes by performing non-equilibrium molecular dynamics simulations. PagP in the conformation from the SDS/MPD crystal was placed in an explicit phospholipid bilayer. No significant reorganization of the PagP backbone was observed during the simulations. Acyl-chain entry was assessed by applying a biasing force to the *sn*-1 chain of a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) molecule from the upper leaflet of the bilayer, directing it towards the empty PagP binding pocket. Of 1205 such simulations, 55 resulted in the *sn*-1 chain entering the binding pocket between strands F and G (the F/G route), five resulted in the *sn*-1 chain jumping out of the bilayer over the top of the protein near strands B to E, and in 1145 simulations the protein blocked the steered

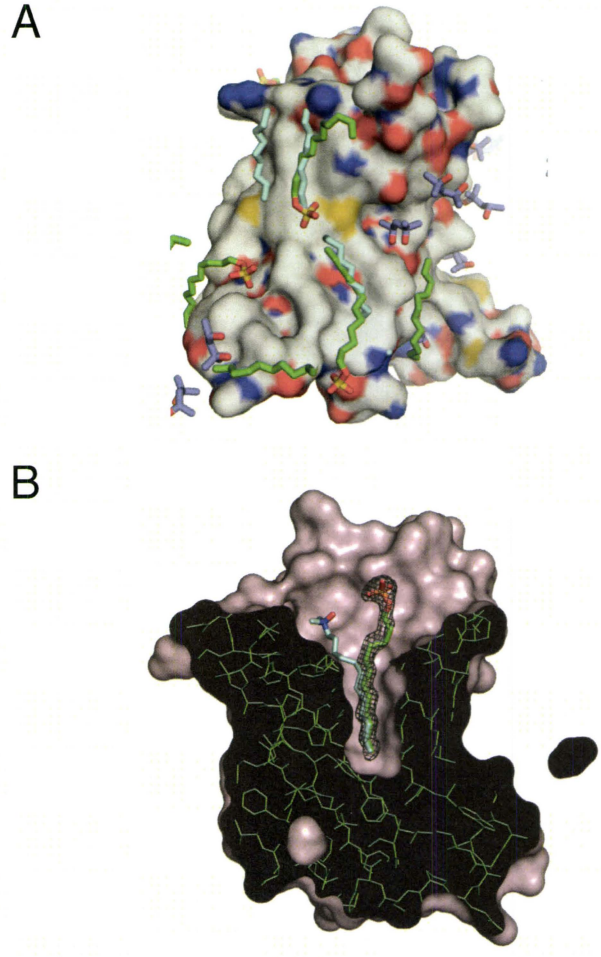

FIGURE 5.3: SDS and MPD environment in the crystal: (A) Externally associated SDS and MPD molecules are shown on the PagP solventaccessible surface. PagP is colored with white carbons, red oxygens, blue nitrogens and yellow sulfurs. The approximate position of the bilayer is indicated. The MPD molecules are shown with purple carbons and red oxygens. The SDS molecules are show with green aliphatic tails. The LDAO detergent coordinates from PDB ID 1THQ are included after superposition of the proteins, and are shown with cyan carbons. (B) The surface of PagP from the SDS/MPD structure is shown in pink, and a cut through the structure shows the pocket of the hydrocarbon ruler. The SDS molecule in the pocket is shown in stick representation with green carbons. The mesh shows the electron density for SDS in the 2Fo-Fc map contoured at 1.2 . The LDAO molecule in the pocket of the LDAO crystal PDB ID 1THQ is shown with cyan carbons for comparison.

phospholipid from approaching the binding pocket more closely than other phospholipids in the first solvation shell (Fig. 5.4*A*).

Next, acyl-chain exit was assessed by applying a biasing force to a dodecyl phosphate (DP) molecule placed in the binding pocket of PagP, compelling it to exit in the bilayer plane. Of 302 such simulations, 95 resulted in detergent extraction by the F/G route, 26 by the B/C route, and 181 extraction attempts ended with the detergent still inside the binding pocket (Fig. 5.4*B*). The fact that the B/C route was identified as a viable route of acyl-chain extraction, but not insertion, is likely due in part to the tilt of PagP in the bilayer plane, which lowers the F and G strands into the hydrophobic region of a POPC bilayer more than it lowers the B and C strands. Analogous simulations starting from the LDAO crystal structure produced results very similar to the simulations from the SDS structure, with additional rare phospholipid chain entry between the A and B strands.

In order to investigate the putative F/G route in the LDAO structure, we conducted MD simulations of the LDAO crystal structure in an explicit POPC bilayer using a non-equilibrium variant of the weighted ensemble method of Huber and Kim (*273*). This evaluates the possibility of acyl-chain entry through the F/G route of the LDAO structure based on thermally-driven conformational fluctuations. Briefly, the system was iteratively simulated in 0.2 ps segments, accepting simulation segments in which the donor chain spontaneously moved closer to the binding pocket, and discarding

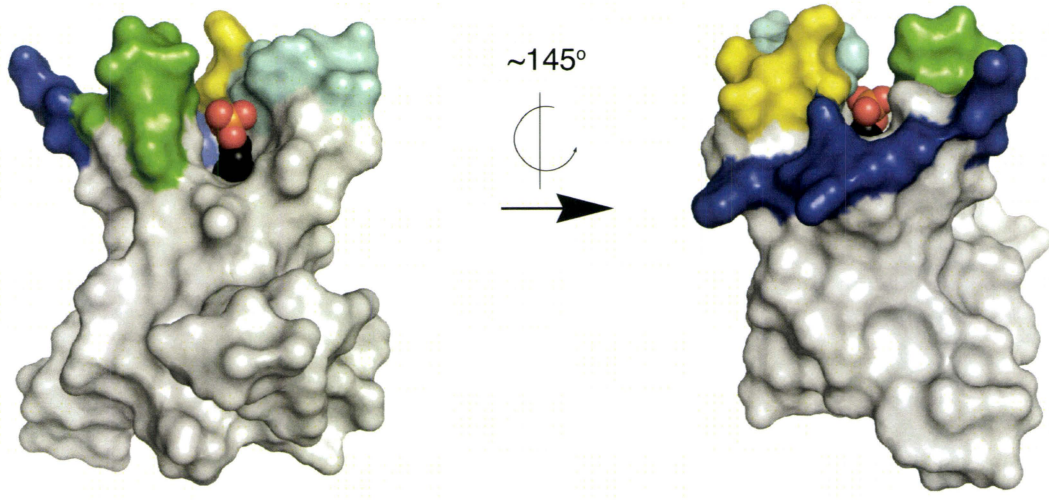

FIGURE 5.4: Crenels as possible routes in and out of the hydrocarbon ruler pocket: PagP is shown in a surface representation with the β–barrel and the periplasmic loops in white. Loop L1 is dark blue, L2 is yellow, L3 is cyan and L4 is green. The SDS molecule buried within the hydrocarbon ruler pocket is shown as a space filling model with black carbons, orange sulphur and red oxygens. The SDS molecule is clearly visible through a crenel between strands F and G in (A) and through a crenel between strands B and C in (B). The F/G crenel is larger and deeper in the membrane than the B/C crenel.

them otherwise. During this simulation, 350 of 400,000 simulation segments successfully progressed towards a bound state, yielding a biased trajectory of 70 ps. These simulations showed a disruption of the hydrogen bond between the phenolic oxygen of Tyr147 and the amide hydrogen of Leu125. This was followed by a rotation of the Tyr147 sidechain towards Arg114, opening a hole in the F/G route through which the *sn*-1 chain of the donor POPC molecule entered the PagP binding pocket (Fig. 5.5). The progress of the acyl chain toward the pocket was impeded until Tyr147 moved away from Leu125.

Finally, we tested the specific role of Tyr147 as a putative gateway residue. Mutants Tyr147Ala and Tyr147Phe had similar CD spectra and denaturation profiles to wild-type PagP (Fig. 5.6*A,B*), indicating that these mutations did not cause significant structural perturbations of the enzyme. *In vitro* reaction rates were measured for acylation of Kdo_2-lipid A and for the phospholipase activity. The latter assay was carried out in the absence of lipid A acceptor, in which case the rate is three to four times slower than the lipid A acylation reaction rate. For both the acyltransferase and the phospholipase activities, the reaction rate is 2-3 fold faster for the Tyr147Ala and Tyr147Phe mutants relative to the wild type (Fig. 5.6 *C,D*), suggesting a role for the phenolic hydroxyl group of tyrosine 147 in controlling phospholipid gating through the F/G crenel.

5.4 Discussion:

In this article we report the use of SDS/MPD as a detergent system for membrane protein crystallization. We have used the SDS/MPD system to refold lysozyme (a β-

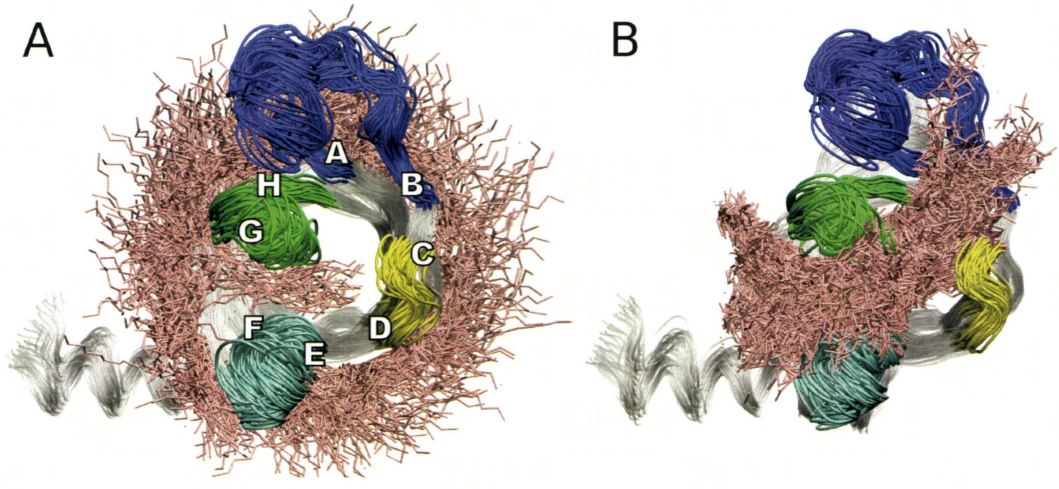

FIGURE 5.5 Plausible routes of ligand entry and exit to and from the PagP binding site based on steered MD simulations: The strands of the β-barrel are labeled A-H and the extracellular loops L1, L2, L3, and L4 are shown in dark blue, yellow, light blue, and green, respectively. Steered acyl chains are overlaid in pink. Acyl chains that jumped out of the bilayer, and all unbiased lipids, are omitted for clarity. (A) The distal 13 carbons of the *sn*-1 chain of biased POPC molecules after 1205 insertion attempts. (B) Detergent heavy atoms within 6 Å of PagP during 302 extraction attempts. Only the F/G route is available for insertion while the F/G and the B/C route can be used for extraction of acyl chains from the hydrocarbon ruler.

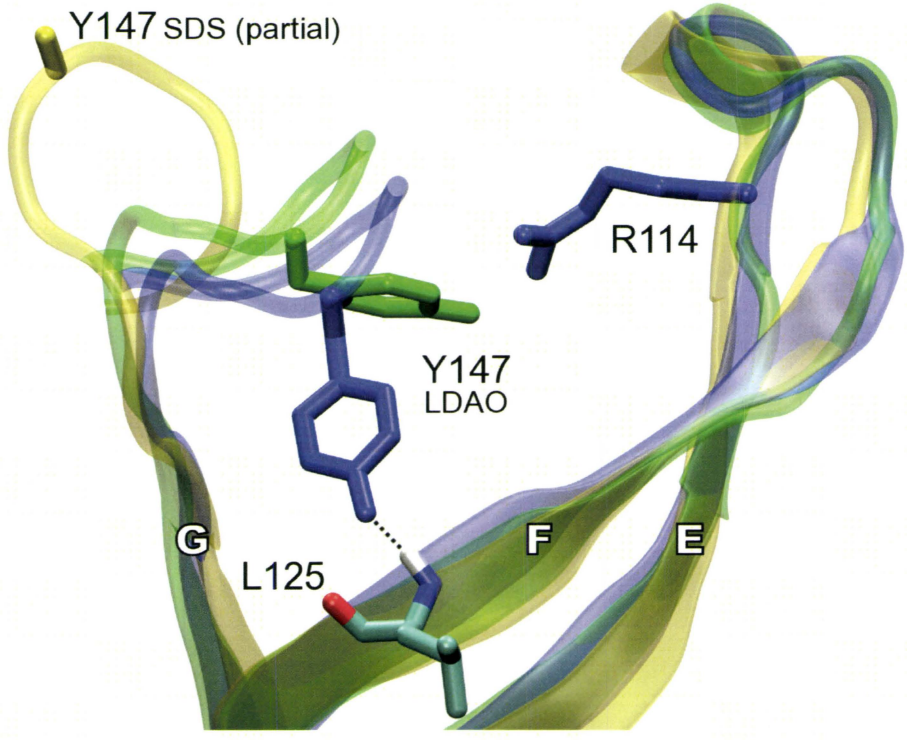

Figure 5.6 Side view of the putative ligand entry route at the F/G crenel:
Only PagP strands E, F and G, including loops L3 and L4, are shown. The phenolic oxygen of Y147 makes a hydrogen bond to the backbone amide hydrogen of Leu125 and obstructs the F/G route in a MD-equilibrated structure based on the LDAO (dark blue), but not the SDS (yellow), crystal form. During non-equilibrium MD sampling based on the weighted-sampling method, Tyr147 rotates toward Arg114 (green), allowing the *sn*-1 chain of the donor POPC molecule to enter the binding pocket (not shown). Atoms in Leu125, shown only for the MD-equilibrated structure based on the LDAO crystal form, are colored red, dark blue, light blue, and white, for oxygen, nitrogen, carbon, and hydrogen, respectively. The hydrogen bond is shown as a broken black line.

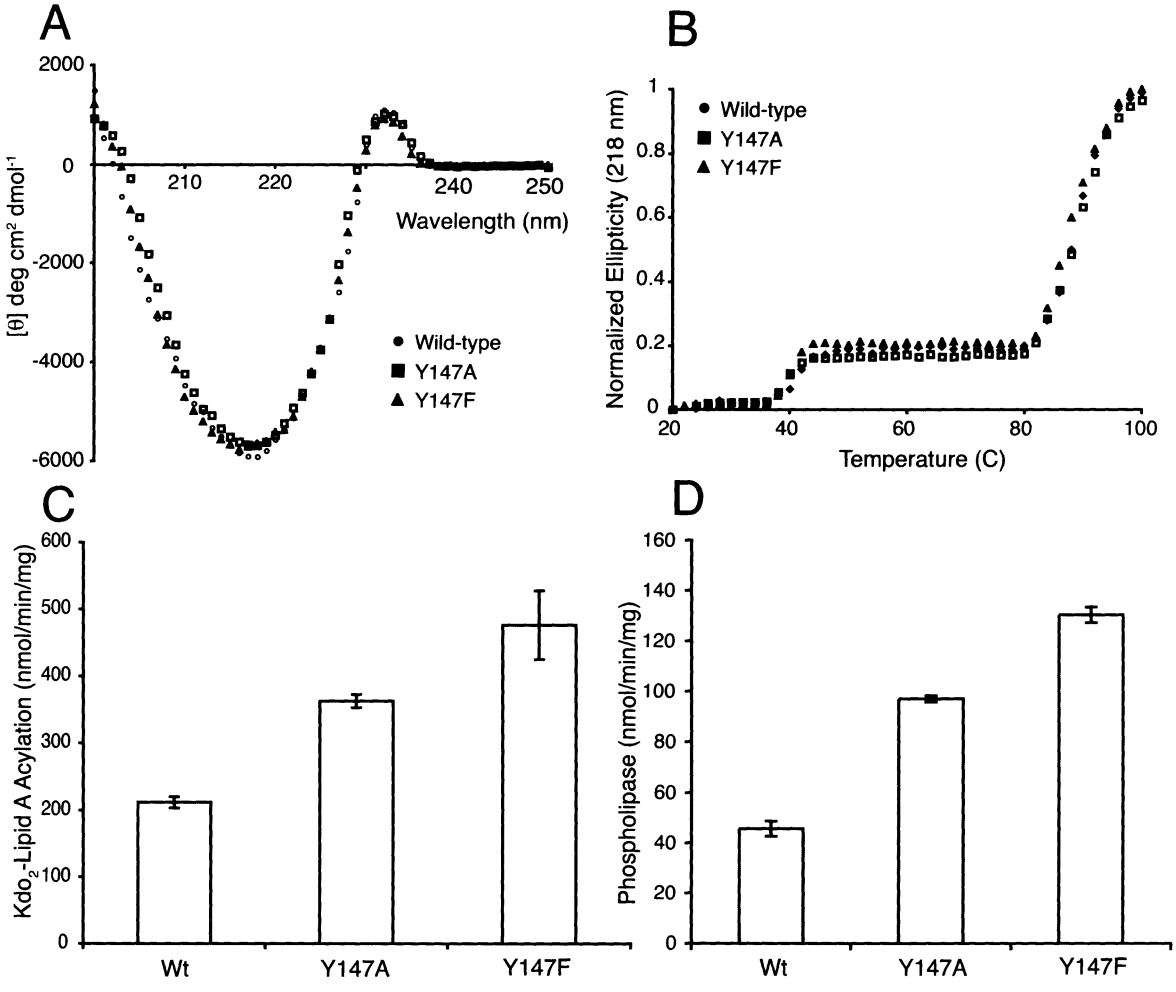

Figure 5-7 Analysis of wild-type PagP and the Y147A and Y147F mutants: Neither mutation has a significant effect on PagP structure or stability, as indicated by far UV CD spectra (A) and thermal denaturation by CD at 218 nm (B). Specific activities *in vitro* for lipid A palmitoyltransferase activity (C) and the phospholipase activity (D).

strand soluble protein), carbonic anhydrase (an α-helical soluble protein) and PagP (a β-strand integral membrane protein) from the SDS-denatured state, indicating that the SDS/MPD system displays properties of a relatively mild detergent of wide applicability (*217*). SDS alone has strongly denaturing properties, but the addition of MPD alters the physicochemical properties of the detergent such that the SDS micelles no longer associate strongly with polypeptides (*217*). It is unlikely that SDS adversely affects the PagP structure presented here because we observe only weak associations between the proteins and ordered detergent monomers.

In the structure of PagP in SDS/MPD, loops L2 and L3 have moved closer to the barrel axis relative to the LDAO structure. The largest difference, however, is the movement of loop L4 away from the acyl chain binding pocket. This creates a significant gap between strands F and G, opening a route to the internal hydrocarbon-binding site. Our MD simulations are consistent in showing that this is the most favoured route for acyl chain entry from the membrane phase. The increased phospholipase activity of the Tyr147 mutants provides independent experimental support that the phospholipid chain enters the barrel via the F/G route. The reaction rates for the phospholipase activity and the acyl transferase activity are affected at the same ratios by the Tyr147Ala and Tyr147Phe mutations, further suggesting that acyl chain entry is the rate limiting step in both reactions.

A recent report on the FadL fatty acid transporter (*226*) shows that hydrophobic compounds can exit this β –barrel structure via a portal-type of opening in the side of the barrel wall. The structures of the FadL family of proteins are distinct from the PagP proteins, and the type of opening in the barrels is different. Nevertheless, in both cases, irregularities in the strand structures provide lateral access routes to and from the membrane hydrocarbon phase. This suggests that there are multiple ways in which β-barrel membrane proteins can facilitate the passage of hydrophobic compounds, notably acyl chains, between the bulk lipid phase and the protein interiors.

6.0 Conclusions and Future Directions

6.1 General Conclusion:

The conclusions made in the previous chapters provide a comprehensive picture of PagP-substrate interaction. We propose that the enzyme rests in the OM waiting for conditions to promote the disruption of membrane lipid asymmetry, which permits phospholipids to migrate into the outer leaflet (*3*). PagP thus gains access to both its substrates and catalyzes the acyltransferase reaction (*123*). The phospholipid donor and the lipid A acceptor enter the PagP active site via lateral diffusion. The phospholipid enters the enzyme through the crenel between β strands F and G (*274*), where Tyr147 behaves as the gatekeeper. Similarly, lipid A gains access through the A/B embrasure (*274*). The active site of the enzyme also includes a hydrophobic pocket that resides in the β barrel interior (*111*), and a string of catalytically significant residues located at the exterior face (*110*). The hydrophobic pocket functions as the hydrocarbon ruler and confers selection of a palmitate chain. This selection is stringent, and a change in the length of the donor acyl chain by as little as a single methylene unit decreases the activity of PagP several fold (*111, 126*). This acyl chain resolution can be modulated by mutation of Gly88 lining the hydrocarbon ruler floor. An intrinsic spectroscopic signature arising from an exciton interaction between Tyr26 and Trp66 at the base of the hydrocarbon ruler is disrupted by the Gly88Cys substitution (*126*). This Cys residue exhibits a suppressed side chain pKa resulting in a buried thiolate anion. This suppression of the Cys pKa in a hydrophobic environment was unexpected and suggested a unique mechanism to

suppress a thiol pKa. The mechanism by which the palmitoyl group is transferred from the phospholipid donor to the lipid A acceptor remains to be elucidated.

6.2 Future Experiments:

While we have provided detailed analysis for many aspects of the model, several molecular details are still outstanding. 1) What is the kinetic mechanism of the enzyme? 2) Do the substrates bind independently of each other or is this event ordered? 3) What is the catalytic mechanism of PagP?

6.2.1 PagP Kinetic Mechanism:

Kinetic analysis of membrane enzymes in detergent micelle systems can be a difficult undertaking. The *in vitro* enzyme assay reported previously for PagP analysis is conducted in DDM which forms large micelles of 72 kDa (*275*). Initial results using this detergent system suggest that the micelle may support phospholipid and lipid A concentrations in the high micro-molar range. This may prove to be a saturating concentration for PagP, since the enzyme in the assay is kept at a low micro-molar scale (*126*).

Membrane intrinsic enzymes of lipid metabolism can be described by the models of interfacial kinetics (*276*) or surface dilution kinetics (*277*). Our defined micellar condition for PagP activity reveal Michaelis Menten behaviour that should allow for steady state kinetic analysis to define parameters k_{cat} and K_m for both substrates.

PagP may operate through a ping-pong mechanism where substrate A binds to the enzyme forming an acyl-enzyme intermediate and release of product A (*224*). This intermediate is resolved by substrate B resulting in formation of product B. Alternatively, PagP might function via ternary complex formation (sequential mechanism), where both substrates are required to bind the enzyme before catalysis can occur (*224*). Initial results indicate that PagP operates via the latter, but we are currently refining conditions for initial velocity experiments.

6.2.2 Order of Substrate Binding:

To establish the order of substrate binding, the classical approach is to conduct product inhibition studies. Both the hepta-acylated lipid A and the *sn*-1 lysophospholipid products can be purified from the PagP reaction cocktail by using reversed phase chromatography (*278*). The *sn*-1 lysophospholipid is unstable and rearranges into the *sn*-2 regioisomer (*279*). Nishijima *et al* reported that *sn*-1 lysophospholipids can be maintained at pH 6.4 (*280*). The optimal pH for PagP activity is 8.0 and it reduces by 95% at pH 6.0. Furthermore, the PagP embrasure is a non-specific substrate entry site and is does not select for a particular acyl chain acceptor (*274*). Therefore, the lysophospholipid product might plausibly be reacylated through the embrasure. Consequently, product inhibition studies might not be possible in PagP.

6.2.3 Catalytic Mechanism of PagP:

A comprehensive description of the catalytic mechanism of PagP will have two major components: 1) which residues are critical in catalysis and 2) what structural changes occur during catalysis. Previously three conserved residues His33, Asp76 and Ser77, which map to the extracellular face of PagP, were suggested to be key in catalysis (*110*). It was suggested that PagP's catalytic mechanism might employ a classical catalytic triad. This was later disproved due to the large distances between these residues apparent in the PagP crystal structure (*111*) and the enzyme's resistance to common inhibitors of serine-esterases. Presently, this list has grown to include Tyr38, Asn47, Glu48, Glu82, Thr112, Arg114 and Asn151, which are also located at the extracellular face of the enzyme. Some of these residues may be significant in the catalytic electron relay, while others may play a role in substrate binding. For example, Lys42 located in the L1 loop was shown to be involved in lipid A binding, since the mutant was unable to catalyze the acyltransferase reaction, and transformed PagP into a dedicated phospholipase (*274*). We are currently using the *in vitro* acyltransferase and phospholipase assays to determine specific roles of these ten residues.

The crystal structure of PagP and the first NMR structure were resolved in LDAO and DPC, respectively (*110, 111*). Both of these detergent systems do not support enzyme activity, since they occupy the hydrocarbon ruler and out compete the phospholipid substrate. A second NMR study of PagP refolded in Cyfos7, a detergent that supports

PagP catalysis because it is too bulky to occupy the hydrocarbon ruler, showed a mixed population of two distinct conformations: an inhibited state, termed "relaxed" (R) and an active state called "tense" (T) (*125*). We have since made several attempts at isolating the T state of PagP. Currently we are employing a two-tiered approach to irreversibly lock PagP in its active conformation, which can then be examined through NMR and crystallography. A PagP-substrate complex can be obtained using natural lipid substrates alongside a catalytically inactive version of the enzyme i.e. a Ser77 mutant (*150*). Alternatively, wild-type PagP can be used in conjunction with non-hydrolysable substrates like the ether analogue of dipalmitoylphosphotidycholine. NMR and crystallography analysis of a PagP substrate complex would require lipids to be in high milli-molar concentrations, since proteins are maintained at low milli-molar levels. Therefore, achieving saturating substrate concentrations poses a significant challenge, and other techniques that may reveal structural information associated with an enzyme-substrate complex must be considered.

CD spectroscopy requires proteins to be in a low micro-molar range (*126*). A high micro-molar concentration of substrates that can be attained in DDM and/or Cyfos7 will therefore be saturating for PagP. Many of the residues implicated in catalysis are localized to the extracellular loop regions. The L1 loop houses 18 residues, many of which are highly conserved. In the absence of substrates this loop exhibits dynamic behaviour and presents itself as a disordered region in both NMR and X-ray crystallography studies. Therefore, it stands to reason that there may be significant

structural changes involved at the extracellular face of PagP upon binding of the substrates in order to bring the catalytic residues in close proximity with the substrates. The resulting changes in secondary structures as a result of catalysis might be detectable through use of CD. We are presently at the preliminary stages of these experiments and are refining conditions for *in vitro* PagP refolding in Cyfos7 and DDM.

6.3 Current Considerations:

6.3.1 Comparison of *in vitro* models with the *in vivo* condition:

It remains to be seen if the conclusions drawn from these *in vitro* studies hold true in bacterial cells. Several differences are obvious upon comparisons of *in vitro* versus *in vivo* conditions. For example, Trp66 aside from its critical role as an exciton partner for Tyr26, was also shown to be an *in vitro* folding determinant (*126*). While the mutant cannot be refolded in detergent micelles, it assembled correctly in *E. coli* membranes. This result indicates that the PagP folding process in bacterial cells likely involves periplasmic chaperones. Several chaperones have been reported, for example Skp and SurA (*89, 281, 282*), which may be involved in *in vivo* PagP folding. These potential associations can easily be identified through techniques such as co-immunoprecipitation or gel shift assays.

Lastly, the Gly88Met mutant showed that PagP can be converted into an efficient lauroyl transferase (*111, 126*). To determine the limits of the PagP hydrocarbon ruler, we propose to systematically decrease the depth of the hydrophobic pocket by introducing

linear, aliphatic chains, whose lengths vary in increments of single methylene units. In the absence of naturally occurring amino acids that possess linear, unbranched aliphatic side chains longer then Met, we will conduct site-specific alkylation of the Gly88Cys mutant. Previously, we successfully added a methyl group to the free sulfhydryl group of Gly88Cys using methyl-4-nitrobenzenesulfonate without generating any side reactions (*162*). We have since been able to obtain *n*-propyl-4-nitrobenzenesulfonate, *n*-butyl-4-nitrobenzenesulfonate, *n*-pentyl-4-nitrobenzenesulfonate and *n*-hexyl-4-nitrobenzenesulfonate from Dr. Fraser Hof at the University of Victoria, British Columbia. We hypothesize that as a result of these alkyl additions, the acyl chain preference of PagP will shift to C11, C10, C9 and C8, respectively. This approach might convert LDAO into a detergent system that supports PagP activity. This is significant since conditions for x-ray crystallography of PagP in LDAO have been reported (*111*). LDAO is a zwitterionic detergent that has a 12-carbon long acyl chain, which occupies the hydrocarbon ruler and inhibits phospholipid binding, thereby inhibiting all enzyme activity. Therefore, by shortening the acyl-chain selection of PagP below C12, we would eliminate this competitive detergent inhibition.

6.3.2 General relevance to other membrane proteins:

Enzymes of lipid metabolism are wide spread and are not exclusive to bacterial membranes. They have been identified in eukaryotic systems as well and are not limited to the membrane milieu, but can also be found in the cytosol. For example, there are a

number of known acyltransferases that are key in cellular metabolic processes, but their complex structural-functional relationships have eluded researchers for several decades. The *in vitro* PagP model might provide critical insights into the study of these enzymes. For instance, lecithin:retinol acyltransferase (LRAT) is a microsomal protein that is essential in the phototransduction cascade of the vision cycle. It catalyzes the transfer of the fatty acid from membrane-associated phosphatidylcholine to retinol (*283, 284*). Similar to PagP, LRAT is a membrane bound protein that is selective for the *sn*-1 position of the phosphatidylcholine donor. It is reported to primarily catalyze the transfer of a palmitate chain *in vivo*, but unlike PagP, no hydrocarbon ruler has been suggested. Therefore, selection of a palmitate chain might primarily be due to the abundance of dipalmitoylphospholipids.

Lecitin:cholesterol acyltransferase (LCAT) is a eukaryotic protein that is responsible for the production of cholesteryl esters in plasma and promoting the formation of high-density lipoproteins (*285*). First described in 1962, the enzyme is primarily expressed in the liver and in small amounts in the brain and testes (*286-289*). In humans, inactivating mutations in the LCAT are associated with pathological conditions called familial LCAT deficiency disease or Fish-Eye disease. A structural model for LCAT based on its homology with the α/β- hydrolase fold family of proteins predicts the presence of a catalytic triad, which is formed by conserved Ser181, Asp345 and His377 residues (*290*). In PagP, conserved His33, Asp76 and Ser77 are critical in catalysis but do not arrange in a catalytic triad. Similarly, LCAT shows preference for the *sn*-2 position of

phosphotidylcholine, which in PagP is switched to the *sn*-1 position. Human LCAT, unlike PagP is not a palmitoyltransferase but preferentially acts on phospholipids containing 18:1 or 18:2 acyl chains (*291, 292*). Despite this difference, substrate specificity exhibited by LCAT suggests that a mechanism similar to the PagP hydrocarbon ruler may exist.

Lastly, lipoproteins are essential components of bacterial cell membranes. The *Escherichia coli* apolipoprotein *N*-acyltransferase (LNT) transfers an acyl group from a glycerophospholipid to the free α-amino group of the *N*-terminal cysteine of apolipoproteins resulting in the mature form (*293*). In Gram-negative bacteria this *N*-acylation is required for engaging the Lol machinery, which either transfers the lipoprotein to the inner leaflet of the outer-membrane or leaves it tethered to the outer leaflet of the inner membrane (*294*). Similar to PagP, LNT is selective for an acyl chain at the *sn*-1 position of the phospholipid donor. While the structure of the membrane bound LNT is unknown at this time, recent experimental evidence shows that the reaction mechanism proceeds via an acyl-enzyme intermediate (*293*). In PagP, initial findings indicate that the enzyme functions via the formation of a ternary complex. There also presumably exists a hydrophobic pocket that is part of the enzyme active site and serves as a binding site for the lipid substrate (*295*). The enzyme is not reported to exhibit acyl chain length selection at the *sn*-1 position of the phospholipid donor, suggesting that the substrate-binding pocket lacks capacity to function as a hydrocarbon ruler.

6.3.3 Relevance to Human Immunology:

Current immunological research has focused upon the effects of lipid A analogues on the activation of NFκB via Toll-like receptor 4 pathway (*296, 297*). It was shown that monosaccharide lipid A analogues were similar to disaccharide lipid As in that they activated both murine and human cells through TLR4. The different analogues were synthetically generated by adding acyl chains of various lengths to the C14 acyl chains present at position 2 and 3 of the sugar moiety. Amongst these, the presence of a C14-O-C14 bond at the 2 position of the sugar showed the highest degree of NFκB activation and TNF-α and IL-8 production. Recently it has been shown that hepta-acylated lipid A up regulates the surface expression of CD86 and CD40 molecules, key players of T cell activation and cell mediated immunity (*298*). Synthetic generation of such lipids is an expensive procedure. The hydrocarbon ruler mutants of PagP are engineered to generate a variety of lipid A analogues with different acylation patterns. We hope the lipid A analogues can function as candidates for vaccine adjuvants and/or endotoxin antagonists. The possibility of targeting PagP with inhibitors to block infection can also be considered once details of the enzymatic reaction mechanism are known.

Table 1 Materials and sources:

Material	Source
Ampicillin	Fischer Scientific
Tetracycline	Fischer Scientific
Granulated agar	Difco
Bacto-tryptone	Difco
NaCl	EMD
Bacto-yeast extract	Difco
Bacterial culture plates	Sarstedt
Polypropylene tubes	Falcon
Sodium dodecyl sulfate	Caledon
2-mercaptoethanol	PMD
Glycerol	Caledon
Tris	GibcoBRL
Proteinase K	EM Science
Phenol	EMD
Ethyl ether	Caledon
Chloroform	Caledon
Methanol	Caledon
Na_2HPO_4	Sigma Aldrich
KH_2PO_4	BDH
KCl	BDH
Tricine	Sigma Aldrich
Bromophenol Blue	Sigma Aldrich
Acetic Acid	EMD
Periodic Acid	Sigma Aldrich
NH_4OH	EMD
$AgNO_3$	Sigma Aldrich
Citric acid	Sigma Aldrich
Formaldehyde	Sigma Aldrich
Phosphorus-32 Radionucleotide	Perkin Elmer
Sodium Acetate	EM Science
N_2 gas	VitalAir
88% Formic acid	Sigma Aldrich
Pyridine	Sigma Aldrich
$MgSO_4$	BioShop
Ammonium persulfate (APS)	Caledon
TEMED	BioShop

Iso-butanol	Fischer
30% acrylamide 0.8% bis-acrylamide	National Diagnostics
Dithiothreitol (DTT)	Caledon
Glycine	BioShop
EDTA	BDH
IPTG	EMD
QIAprep Spin Miniprep Kit	Qiagen
QuickChange Site-Directed Mutagenesis Kit	Stratagene
Phospholipids	Avanti polar lipids
Ethidium Bromide	Sigma Aldrich
NdeI	Fermentas
XhoI	Fermentas
QiaKit Gel Extraction Kit	Qiagen
T4 DNA ligase	Fermentas
DpnI	Stratagene
GenePulser cuvette	BioRad
Guanidine	Sigma Aldrich
BCA protein assay kit	Pierce
His·bind resin	Novagen
$NiSO_4$	Sigma Aldrich
Imidazole	Sigma Aldrich
DE-52 diethylaminoethyl cellulose resin	Whatman
Acetonitrile	EMD Science
16% Tricine gel	Invitrogen
Glass-backed silica gel 60 TLC plates	Caledon
Kdo_2-lipid A	Avanti Polar Lipids
n-Dodecyl-β-D-maltoside (DDM)	Anatrace
TX-100	Anatrace
Lauroyldimethylamine N-oxide (LDAO)	Anatrace

Table 2 Oligonucleotide Sequences:

Tyr147Ala F	5'CCTACATTCCGGGGTACC***GCC***AACAATGGCCAATGTG3'
Tyr147Ala R	5'CACATTGGCCATTGTT***GGC***GGTACCCCGGAATGTAGG3'
Tyr147Phe F	5'CCTACATTCCGGGGTACC***TTC***AACAATGGCCAATGTG3'
Tyr147Phe R	5'CACATTGGCCATTGTT***GAA***GGTACCCCGGAATGTAGG3'
Pro28Cys F	5'GATTTATATATT***TGT***GCCATCACCTGGCATGC3'
Pro28Cys R	5'GCATGCCAGGTGATGGC***ACA***AATATATAAATC3'
Pro50Cys F	5'CGCTATAACGAGCGA***TGC***TGGGGTGGCGGTTTTGGC3'
Pro50Cys R	5'GCCAAAACCGCCACCCCA***GCA***TCGCTCGTTATAGCG3'
Pro127Cys F	5'CCCTCTCCCGGTTCTACTG***TGC***TTGGCCTCCGTGGG3'
Pro127Cys R	5'CCCACGGAGGCCAA***GCA***CAGTAGAACCGGGAGAGGG3'
Pro144Cys F	5'CAGATGACCTACATT***TGC***GGTACCTACAACAATGGC3'
Pro144Cys R	5'GCCATTGTTGTAGGTACC***GCA***AATGTAGGTCATCTG3'
Gly88Cys F	5'CCGATTGCCGGATAC***TGC***TGGGAAAGTACCTGGCG3'
Gly88Cys R	5'CGCCAGGTACTTTCCCA***GCA***GTATCCGGCAATCGG3'
Gly88Ser F	5'CCGATTGCCGGATAC***TCA***TGGGAAAGTACCTGGCG3'
Gly88Ser R	5'CGCCAGGTACTTTCCCA***TGA***GTATCCGGCAATCGG3'
jjjjGly88Thr F	5'CCGATTGCCGGATAC***ACA***TGGGAAAGTACCTGGCG3'
Gly88Thr R	5'CGCCAGGTACTTTCCCA***TGT***GTATCCGGCAATCGG3'
Lys42Ala F	5'CTTACGACAAAGAA***GCA***ACCGATCGCTATAACGAG3'
Lys42Ala R	5'CTCGTTATAGCGATCGGT***TGC***TTCTTTGTCGTAAG3'
Gly88Asp F	5'CCGATTGCCGGATAC***GAT***TGGGAAAGTACCTGGCG3'
Gly88Asp R	5'CGCCAGGTACTTTCCCA***ATC***GTATCCGGCAATCGG3'

Gly88Asn F	5'CCGATTGCCGGATAC***AAT***TGGGAAAGTACCTGGCG3'
Gly88Asn R	5'CGCCAGGTACTTTCCCA***ATT***GTATCCGGCAATCGG3'
Gly88Glu F	5'CCGATTGCCGGATAC***GAA***TGGGAAAGTACCTGGCG3'
Gly88Glu R	5'CGCCAGGTACTTTCCCA***TTC***GTATCCGGCAATCGG3'
Gly88Gln F	5'CCGATTGCCGGATAC***CAA***TGGGAAAGTACCTGGCG3'
Gly88Gln R	5'CGCCAGGTACTTTCCCAT***TGG***TATCCGGCAATCGG3'
Gly88His F	5'CCGATTGCCGGATAC***CAT***TGGGAAAGTACCTGGCG3'
Gly88His R	5'CGCCAGGTACTTTCCCA***ATG***GTATCCGGCAATCGG3'

Table 3: Bacterial plasmids and strains:

Plasmid/Strain	Description/Genotype	Source
pETCrcAHΔS	510 bp NdeI-XhoI PCR product carrying crcA cloned into an IPTG-inducible T7 RNA polymerase-promoter expression vector (ApR) with a C-terminal His6 tag. A N-terminal stretch of 71 bp constituting the signal sequence is replaced by a single Met.	Hwang *et al.* 2002
pBadPagP	615 bp *Sal*I and *Hind*III PCR product carrying crcA cloned into the pBad18 vector carrying an Amp cassette.	Guzman *et al.*, 1995
E. coli		
XL1-Blue	F-, recA1, gyrA96, thi-1,hsdR17 SupE44, relA1, lac, [F' proAB, lacIqZΔM15, Tn10(Tetr)]c	Stratagene
BL21(DE3)	F-, ompT, hsdSB(rB-mb-), gal, dcm, λ(DE3)	Novagen
WBB06	W3110 mtl, Δ(rfaC-rfaF)::tet6	Brabetz *et al.*, (1997)
NR754	MC4100 (F^- *araD139* Δ(*argF-lac*) *U169 rpsL150 relA1 flb5301 deoC1 ptsF25 thi*) ara^+Δ*pagP*	Ruiz *et al.*, (2008)

Table 4 ESI-MS Results:

Name	Theoretical	Experimental
WtPagP	20175.49	20176.97±1.96
G88A	20189.52	20189.48±0.55
G88M	20249.63	20250.52±0.99
G88C	20221.58	20221.46±0.84
G88S	20205.52	20205.69±0.17
G88T	20219.54	20220.00±1.22
G88N	20232.54	20233.26±1.47
G88D	20233.57	20234.12±0.97
G88Q	20246.57	20245.96±0.65
G88E	20247.55	20248.10±0.89
G88H	20255.58	20254.96±1.04
G88K	20246.61	20246.35±1.45
Y147A	20159.49	20159.52±0.67
Y147F	20125.47	20126.23±0.88
P28C	20181.51	20182.70±1.31
P50C	20181.51	20182.67±1.46
P127C	20181.51	20182.60±1.21
P144C	20181.51	20182.52±1.11
P28C P50C	20187.54	20187.60±1.53
P28C P50C + $Cu(OP)_3$	20185.52	20186.43±1.90
P127C P144C	20187.49	20187.83±1.67
P127C P144C + bBBr	20375.75	20374.90±2.10

K42A	20164.48	20163.92±1.57

Table 5-1 Data Collection and Refinement Statistics:

	PagP / SDS / MPD PDB ID XXXX
Data collection	
Space group	$P6_222$
Unit cell (Å)	a,b=113.23, c=55.06
R_{sym} [a]	0.068 (0.531)
$I / \sigma I$)	28.1 (1.3)
Completeness (%)	97.7 (80.4)
Redundancy	8.2 (4.1)
Refinement	
Resolution (Å)	20.0-1.40
No. reflections	40392
R_{work} / R_{free}	0.171 (0.303) / 0.208 (0.290)
Number of atoms	
Protein	1321
SDS	97 (6 molecules)
MPD	80 (10 molecules)
Ions	32
Water	56
B-factors (Å^2)	
Protein	27.1

SDS	47.6
MPD	48.3
Ions	60.5
Water	38.1
Rmsd bond length (Å)	0.021
Rmsd bond angles (°)	2.12

[a]Values in parentheses are for the high resolution shell from 1.42 – 1.40 Å.

References:

(1) Raetz, C. R., and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem 71*, 635-700.

(2) Holtje, J. V. (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*. *Microbiol Mol Biol Rev 62*, 181-203.

(3) Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev 67*, 593-656.

(4) Delcour, A. H. (2009) Outer membrane permeability and antibiotic resistance. *Biochim Biophys Acta 1794*, 808-16.

(5) Rosner, M. R., Tang, J., Barzilay, I., and Khorana, H. G. (1979) Structure of the lipopolysaccharide from an *Escherichia coli* heptose-less mutant. I. Chemical degradations and identification of products. *J Biol Chem 254*, 5906-17.

(6) Vaara, M. (1993) Antibiotic-supersusceptible mutants of *Escherichia coli* and *Salmonella typhimurium*. *Antimicrob Agents Chemother 37*, 2255-60.

(7) Miller, S. I., Ernst, R. K., and Bader, M. W. (2005) LPS, TLR4 and infectious disease diversity. *Nat Rev Microbiol 3*, 36-46.

(8) Iwasaki, A., and Medzhitov, R. (2004) Toll-like receptor control of the adaptive immune responses. *Nat Immunol 5*, 987-95.

(9) Salomao, R., Martins, P. S., Brunialti, M. K., Fernandes Mda, L., Martos, L. S., Mendes, M. E., Gomes, N. E., and Rigato, O. (2008) TLR signaling pathway in patients with sepsis. *Shock 30 Suppl 1*, 73-7.

(10) Schlichting, D., and McCollam, J. S. (2007) Recognizing and managing severe sepsis: a common and deadly threat. *South Med J 100*, 594-600.

(11) Yang, K. H., and Lee, M. G. (2008) Effects of endotoxin derived from *Escherichia coli* lipopolysaccharide on the pharmacokinetics of drugs. *Arch Pharm Res 31*, 1073-86.

(12) Raetz, C. R., Garrett, T. A., Reynolds, C. M., Shaw, W. A., Moore, J. D., Smith, D. C., Jr., Ribeiro, A. A., Murphy, R. C., Ulevitch, R. J., Fearns, C., Reichart, D., Glass, C. K., Benner, C., Subramaniam, S., Harkewicz, R., Bowers-Gentry, R. C.,

Buczynski, M. W., Cooper, J. A., Deems, R. A., and Dennis, E. A. (2006) Kdo_2-Lipid A of *Escherichia coli*, a defined endotoxin that activates macrophages via TLR-4. *J Lipid Res 47*, 1097-111.

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

(14) Williams, A. H., Immormino, R. M., Gewirth, D. T., and Raetz, C. R. (2006) Structure of UDP-*N*-acetylglucosamine acyltransferase with a bound antibacterial pentadecapeptide. *Proc Natl Acad Sci U S A 103*, 10877-82.

(15) Williams, A. H., and Raetz, C. R. (2007) Structural basis for the acyl chain selectivity and mechanism of UDP-*N*-acetylglucosamine acyltransferase. *Proc Natl Acad Sci U S A 104*, 13543-50.

(16) Anderson, M. S., Bull, H. G., Galloway, S. M., Kelly, T. M., Mohan, S., Radika, K., and Raetz, C. R. (1993) UDP-*N*-acetylglucosamine acyltransferase of *Escherichia coli*. The first step of endotoxin biosynthesis is thermodynamically unfavorable. *J Biol Chem 268*, 19858-65.

(17) Sorensen, P. G., Lutkenhaus, J., Young, K., Eveland, S. S., Anderson, M. S., and Raetz, C. R. (1996) Regulation of UDP-3-*O*-[*R*-3-hydroxymyristoyl]-*N*-acetylglucosamine deacetylase in *Escherichia coli*. The second enzymatic step of lipid a biosynthesis. *J Biol Chem 271*, 25898-905.

(18) Shin, H., Gennadios, H. A., Whittington, D. A., and Christianson, D. W. (2007) Amphipathic benzoic acid derivatives: synthesis and binding in the hydrophobic tunnel of the zinc deacetylase LpxC. *Bioorg Med Chem 15*, 2617-23.

(19) Whittington, D. A., Rusche, K. M., Shin, H., Fierke, C. A., and Christianson, D. W. (2003) Crystal structure of LpxC, a zinc-dependent deacetylase essential for endotoxin biosynthesis. *Proc Natl Acad Sci U S A 100*, 8146-50.

(20) Mochalkin, I., Knafels, J. D., and Lightle, S. (2008) Crystal structure of LpxC from *Pseudomonas aeruginosa* complexed with the potent BB-78485 inhibitor. *Protein Sci 17*, 450-7.

(21) Coggins, B. E., McClerren, A. L., Jiang, L., Li, X., Rudolph, J., Hindsgaul, O., Raetz, C. R., and Zhou, P. (2005) Refined solution structure of the LpxC-TU-514 complex and pKa analysis of an active site histidine: insights into the mechanism and inhibitor design. *Biochemistry 44*, 1114-26.

(22) Kelly, T. M., Stachula, S. A., Raetz, C. R., and Anderson, M. S. (1993) The firA gene of *Escherichia coli* encodes UDP-3-*O*-(*R*-3-hydroxymyristoyl)-glucosamine *N*-acyltransferase. The third step of endotoxin biosynthesis. *J Biol Chem 268*, 19866-74.

(23) Buetow, L., Smith, T. K., Dawson, A., Fyffe, S., and Hunter, W. N. (2007) Structure and reactivity of LpxD, the *N*-acyltransferase of lipid A biosynthesis. *Proc Natl Acad Sci U S A 104*, 4321-6.

(24) Babinski, K. J., Ribeiro, A. A., and Raetz, C. R. (2002) The *Escherichia coli* gene encoding the UDP-2,3-diacylglucosamine pyrophosphatase of lipid A biosynthesis. *J Biol Chem 277*, 25937-46.

(25) Crowell, D. N., Anderson, M. S., and Raetz, C. R. (1986) Molecular cloning of the genes for lipid A disaccharide synthase and UDP-*N*-acetylglucosamine acyltransferase in *Escherichia coli*. *J Bacteriol 168*, 152-9.

(26) Garrett, T. A., Kadrmas, J. L., and Raetz, C. R. (1997) Identification of the gene encoding the *Escherichia coli* lipid A 4'-kinase. Facile phosphorylation of endotoxin analogs with recombinant LpxK. *J Biol Chem 272*, 21855-64.

(27) Clementz, T., and Raetz, C. R. (1991) A gene coding for 3-deoxy-D-*manno*-octulosonic-acid transferase in *Escherichia coli*. Identification, mapping, cloning, and sequencing. *J Biol Chem 266*, 9687-96.

(28) Raetz, C. R., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) Lipid A modification systems in Gram-negative bacteria. *Annu Rev Biochem 76*, 295-329.

(29) Clementz, T., Zhou, Z., and Raetz, C. R. (1997) Function of the *Escherichia coli* *msbB* gene, a multicopy suppressor of *htrB* knockouts, in the acylation of lipid A. Acylation by MsbB follows laurate incorporation by HtrB. *J Biol Chem 272*, 10353-60.

(30) Ellis, J., Oyston, P. C., Green, M., and Titball, R. W. (2002) Tularemia. *Clin Microbiol Rev 15*, 631-46.

(31) Wang, X., Karbarz, M. J., McGrath, S. C., Cotter, R. J., and Raetz, C. R. (2004) MsbA transporter-dependent lipid A 1-dephosphorylation on the periplasmic surface of the inner membrane: topography of francisella novicida LpxE expressed in *Escherichia coli*. *J Biol Chem 279*, 49470-8.

(32) Wang, X., McGrath, S. C., Cotter, R. J., and Raetz, C. R. (2006) Expression cloning and periplasmic orientation of the *Francisella novicida* lipid A 4'-phosphatase LpxF. *J Biol Chem 281*, 9321-30.

(33) Doerrler, W. T., Gibbons, H. S., and Raetz, C. R. (2004) MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. *J Biol Chem 279*, 45102-9.

(34) Marino, P. A., McGrath, B. C., and Osborn, M. J. (1991) Energy dependence of O-antigen synthesis in *Salmonella typhimurium*. *J Bacteriol 173*, 3128-33.

(35) Valvano, M. A. (2003) Export of O-specific lipopolysaccharide. *Front Biosci 8*, s452-71.

(36) Ruiz, N., Kahne, D., and Silhavy, T. J. (2009) Transport of lipopolysaccharide across the cell envelope: the long road of discovery. *Nat Rev Microbiol 7*, 677-83.

(37) Sperandeo, P., Cescutti, R., Villa, R., Di Benedetto, C., Candia, D., Deho, G., and Polissi, A. (2007) Characterization of *lptA* and *lptB*, two essential genes implicated in lipopolysaccharide transport to the outer membrane of *Escherichia coli*. *J Bacteriol 189*, 244-53.

(38) Stenberg, F., Chovanec, P., Maslen, S. L., Robinson, C. V., Ilag, L. L., von Heijne, G., and Daley, D. O. (2005) Protein complexes of the *Escherichia coli* cell envelope. *J Biol Chem 280*, 34409-19.

(39) Sperandeo, P., Lau, F. K., Carpentieri, A., De Castro, C., Molinaro, A., Deho, G., Silhavy, T. J., and Polissi, A. (2008) Functional analysis of the protein machinery required for transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. *J Bacteriol 190*, 4460-9.

(40) Ruiz, N., Gronenberg, L. S., Kahne, D., and Silhavy, T. J. (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A 105*, 5537-42.

(41) Bos, M. P., Tefsen, B., Geurtsen, J., and Tommassen, J. (2004) Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *Proc Natl Acad Sci U S A 101*, 9417-22.

(42) Braun, M., and Silhavy, T. J. (2002) Imp/OstA is required for cell envelope biogenesis in *Escherichia coli*. *Mol Microbiol 45*, 1289-302.

(43) Steeghs, L., den Hartog, R., den Boer, A., Zomer, B., Roholl, P., and van der Ley, P. (1998) Meningitis bacterium is viable without endotoxin. *Nature 392*, 449-50.

(44) Wu, T., McCandlish, A. C., Gronenberg, L. S., Chng, S. S., Silhavy, T. J., and Kahne, D. (2006) Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A 103*, 11754-9.

(45) Coughlin, R. T., Tonsager, S., and McGroarty, E. J. (1983) Quantitation of metal cations bound to membranes and extracted lipopolysaccharide of *Escherichia coli*. *Biochemistry 22*, 2002-7.

(46) Epand, R. M., Shai, Y., Segrest, J. P., and Anantharamaiah, G. M. (1995) Mechanisms for the modulation of membrane bilayer properties by amphipathic helical peptides. *Biopolymers 37*, 319-38.

(47) Hancock, R. E., Falla, T., and Brown, M. (1995) Cationic bactericidal peptides. *Adv Microb Physiol 37*, 135-75.

(48) Zasloff, M. (2002) Antimicrobial peptides of multicellular organisms. *Nature 415*, 389-95.

(49) Guo, L., Lim, K. B., Gunn, J. S., Bainbridge, B., Darveau, R. P., Hackett, M., and Miller, S. I. (1997) Regulation of lipid A modifications by *Salmonella typhimurium* virulence genes *phoP-phoQ*. *Science 276*, 250-3.

(50) Trent, M. S., Pabich, W., Raetz, C. R., and Miller, S. I. (2001) A PhoP/PhoQ-induced Lipase (PagL) that catalyzes 3-*O*-deacylation of lipid A precursors in membranes of *Salmonella typhimurium*. *J Biol Chem 276*, 9083-92.

(51) Gibbons, H. S., Kalb, S. R., Cotter, R. J., and Raetz, C. R. (2005) Role of Mg^{2+} and pH in the modification of Salmonella lipid A after endocytosis by macrophage tumour cells. *Mol Microbiol 55*, 425-40.

(52) Kim, S. H., Jia, W., Parreira, V. R., Bishop, R. E., and Gyles, C. L. (2006) Phosphoethanolamine substitution in the lipid A of *Escherichia coli* O157 : H7 and its association with PmrC. *Microbiology 152*, 657-66.

(53) Williams, G. J., Breazeale, S. D., Raetz, C. R., and Naismith, J. H. (2005) Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. *J Biol Chem 280*, 23000-8.

(54) Breazeale, S. D., Ribeiro, A. A., and Raetz, C. R. (2003) Origin of lipid A species modified with 4-amino-4-deoxy-L-arabinose in polymyxin-resistant mutants of *Escherichia coli*. An aminotransferase (ArnB) that generates UDP-4-deoxyl-$_L$-arabinose. *J Biol Chem 278*, 24731-9.

(55) Breazeale, S. D., Ribeiro, A. A., McClerren, A. L., and Raetz, C. R. (2005) A formyltransferase required for polymyxin resistance in *Escherichia coli* and the modification of lipid A with 4-Amino-4-deoxy-$_L$-arabinose. Identification and function oF UDP-4-deoxy-4-formamido-$_L$-arabinose. *J Biol Chem 280*, 14154-67.

(56) Yan, A., Guan, Z., and Raetz, C. R. (2007) An undecaprenyl phosphate-aminoarabinose flippase required for polymyxin resistance in *Escherichia coli*. *J Biol Chem 282*, 36077-89.

(57) Trent, M. S. (2004) Biosynthesis, transport, and modification of lipid A. *Biochem Cell Biol 82*, 71-86.

(58) Pristovsek, P., and Kidric, J. (1999) Solution structure of polymyxins B and E and effect of binding to lipopolysaccharide: an NMR and molecular modeling study. *J Med Chem 42*, 4604-13.

(59) Tran, A. X., Lester, M. E., Stead, C. M., Raetz, C. R., Maskell, D. J., McGrath, S. C., Cotter, R. J., and Trent, M. S. (2005) Resistance to the antimicrobial peptide polymyxin requires myristoylation of *Escherichia coli* and *Salmonella typhimurium* lipid A. *J Biol Chem 280*, 28186-94.

(60) Bishop, R. E., Kim, S. H., and El Zoeiby, A. (2005) Role of lipid A palmitoylation in bacterial pathogenesis. *J Endotoxin Res 11*, 174-80.

(61) Gibbons, H. S., Lin, S., Cotter, R. J., and Raetz, C. R. (2000) Oxygen requirement for the biosynthesis of the *S*-2-hydroxymyristate moiety in *Salmonella typhimurium* lipid A. Function of LpxO, A new Fe^{2+}/alpha-ketoglutarate-dependent dioxygenase homologue. *J Biol Chem 275*, 32940-9.

(62) Gibbons, H. S., Reynolds, C. M., Guan, Z., and Raetz, C. R. (2008) An inner membrane dioxygenase that generates the 2-hydroxymyristate moiety of Salmonella lipid A. *Biochemistry 47*, 2814-25.

(63) Somerville, J. E., Jr., Cassiano, L., Bainbridge, B., Cunningham, M. D., and Darveau, R. P. (1996) A novel *Escherichia coli* lipid A mutant that produces an antiinflammatory lipopolysaccharide. *J Clin Invest 97*, 359-65.

(64) Hajjar, A. M., Ernst, R. K., Tsai, J. H., Wilson, C. B., and Miller, S. I. (2002) Human Toll-like receptor 4 recognizes host-specific LPS modifications. *Nat Immunol 3*, 354-9.

(65) Kawasaki, K., Ernst, R. K., and Miller, S. I. (2004) 3-O-deacylation of lipid A by PagL, a PhoP/PhoQ-regulated deacylase of *Salmonella typhimurium*, modulates signaling through toll-like receptor 4. *J Biol Chem 279*, 20044-20048.

(66) Bader, M. W., Sanowar, S., Daley, M. E., Schneider, A. R., Cho, U., Xu, W., Klevit, R. E., Le Moual, H., and Miller, S. I. (2005) Recognition of antimicrobial peptides by a bacterial sensor kinase. *Cell 122*, 461-72.

(67) Groisman, E. A. (2001) The pleiotropic two-component regulatory system PhoP-PhoQ. *J Bacteriol 183*, 1835-42.

(68) Garcia Vescovi, E., Soncini, F. C., and Groisman, E. A. (1996) Mg^{2+} as an extracellular signal: environmental regulation of Salmonella virulence. *Cell 84*, 165-74.

(69) Prost, L. R., and Miller, S. I. (2008) The Salmonellae PhoQ sensor: mechanisms of detection of phagosome signals. *Cell Microbiol 10*, 576-82.

(70) Bishop, R. E., Gibbons, H. S., Guina, T., Trent, M. S., Miller, S. I., and Raetz, C. R. (2000) Transfer of palmitate from phospholipids to lipid A in outer membranes of Gram-negative bacteria. *Embo J 19*, 5071-80.

(71) Gunn, J. S., Belden, W. J., and Miller, S. I. (1998) Identification of PhoP-PhoQ activated genes within a duplicated region of the *Salmonella typhimurium* chromosome. *Microb Pathog 25*, 77-90.

(72) Kox, L. F., Wosten, M. M., and Groisman, E. A. (2000) A small protein that mediates the activation of a two-component system by another two-component system. *Embo J 19*, 1861-72.

(73) Kato, A., Mitrophanov, A. Y., and Groisman, E. A. (2007) A connector of two-component regulatory systems promotes signal amplification and persistence of expression. *Proc Natl Acad Sci U S A 104*, 12063-8.

(74) Kato, A., and Groisman, E. A. (2004) Connecting two-component regulatory systems by a protein that protects a response regulator from dephosphorylation by its cognate sensor. *Genes Dev 18*, 2302-13.

(75) Wosten, M. M., Kox, L. F., Chamnongpol, S., Soncini, F. C., and Groisman, E. A. (2000) A signal transduction system that responds to extracellular iron. *Cell 103*, 113-25.

(76) Kato, A., Latifi, T., and Groisman, E. A. (2003) Closing the loop: the PmrA/PmrB two-component system negatively controls expression of its posttranscriptional activator PmrD. *Proc Natl Acad Sci U S A 100*, 4706-11.

(77) Schulz, G. E. (2003) Transmembrane beta-barrel proteins. *Adv Protein Chem 63*, 47-70.

(78) Schulz, G. E. (2002) The structure of bacterial outer membrane proteins. *Biochim Biophys Acta 1565*, 308-17.

(79) Schnaitman, C. A. (1973) Outer membrane proteins of *Escherichia coli*. I. Effect of preparative conditions on the migration of protein in polyacrylamide gels. *Arch Biochem Biophys 157*, 541-52.

(80) Bernstein, H. D. (2000) The biogenesis and assembly of bacterial membrane proteins. *Curr Opin Microbiol 3*, 203-9.

(81) Driessen, A. J., and Nouwen, N. (2008) Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem 77*, 643-67.

(82) Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (2000) The structure and mechanism of bacterial type I signal peptidases. A novel antibiotic target. *Pharmacol Ther 87*, 27-49.

(83) Thome, B. M., and Muller, M. (1991) Skp is a periplasmic *Escherichia coli* protein requiring SecA and SecY for export. *Mol Microbiol 5*, 2815-21.

(84) Behrens, S., Maier, R., de Cock, H., Schmid, F. X., and Gross, C. A. (2001) The SurA periplasmic PPIase lacking its parvulin domains functions *in vivo* and has chaperone activity. *Embo J 20*, 285-94.

(85) Chen, R., and Henning, U. (1996) A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol Microbiol 19*, 1287-94.

(86) Schafer, U., Beck, K., and Muller, M. (1999) Skp, a molecular chaperone of Gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. *J Biol Chem 274*, 24567-74.

(87) Bitto, E., and McKay, D. B. (2003) The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J Biol Chem 278*, 49316-22.

(88) Saul, F. A., Arie, J. P., Vulliez-le Normand, B., Kahn, R., Betton, J. M., and Bentley, G. A. (2004) Structural and functional studies of FkpA from *Escherichia coli*, a cis/trans peptidyl-prolyl isomerase with chaperone activity. *J Mol Biol 335*, 595-608.

(89) Knowles, T. J., Scott-Tucker, A., Overduin, M., and Henderson, I. R. (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol 7*, 206-14.

(90) Doerrler, W. T., and Raetz, C. R. (2005) Loss of outer membrane proteins without inhibition of lipid export in an *Escherichia coli* YaeT mutant. *J Biol Chem 280*, 27679-87.

(91) de Cock, H., Brandenburg, K., Wiese, A., Holst, O., and Seydel, U. (1999) Non-lamellar structure and negative charges of lipopolysaccharides required for efficient folding of outer membrane protein PhoE of *Escherichia coli*. *J Biol Chem 274*, 5114-9.

(92) Scandella, C. J., and Kornberg, A. (1971) A membrane-bound phospholipase A1 purified from *Escherichia coli*. *Biochemistry 10*, 4447-56.

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

(94) Cronan, J. E., Jr., and Wulff, D. L. (1969) A role for phospholipid hydrolysis in the lysis of *Escherichia coli* infected with bacteriophage T4. *Virology 38*, 241-6.

(95) Okuyama, H. (1969) Phospholipid metabolism in *Escherichia coli* after a shift in temperature. *Biochim Biophys Acta 176*, 125-34.

(96) Hardaway, K. L., and Buller, C. S. (1979) Effect of ethylenediaminetetraacetate on phospholipids and outer membrane function in *Escherichia coli*. *J Bacteriol 137*, 62-8.

(97) Pugsley, A. P., and Schwartz, M. (1984) Colicin E2 release: lysis, leakage or secretion? Possible role of a phospholipase. *Embo J 3*, 2393-7.

(98) Snijder, H. J., Ubarretxena-Belandia, I., Blaauw, M., Kalk, K. H., Verheij, H. M., Egmond, M. R., Dekker, N., and Dijkstra, B. W. (1999) Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. *Nature 401*, 717-21.

(99) Kingma, R. L., Fragiathaki, M., Snijder, H. J., Dijkstra, B. W., Verheij, H. M., Dekker, N., and Egmond, M. R. (2000) Unusual catalytic triad of *Escherichia coli* outer membrane phospholipase A. *Biochemistry 39*, 10017-22.

(100) Stanley, A. M., Treubrodt, A. M., Chuawong, P., Hendrickson, T. L., and Fleming, K. G. (2007) Lipid chain selectivity by outer membrane phospholipase A. *J Mol Biol 366*, 461-8.

(101) Dekker, N., Cox, R. C., Kramer, R. A., and Egmond, M. R. (2001) Substrate specificity of the integral membrane protease OmpT determined by spatially addressed peptide libraries. *Biochemistry 40*, 1694-701.

(102) Kramer, R. A., Dekker, N., and Egmond, M. R. (2000) Identification of active site serine and histidine residues in *Escherichia coli* outer membrane protease OmpT. *FEBS Lett 468*, 220-4.

(103) Vandeputte-Rutten, L., Kramer, R. A., Kroon, J., Dekker, N., Egmond, M. R., and Gros, P. (2001) Crystal structure of the outer membrane protease OmpT from *Escherichia coli* suggests a novel catalytic site. *Embo J 20*, 5033-9.

(104) Rutten, L., Mannie, J. P., Stead, C. M., Raetz, C. R., Reynolds, C. M., Bonvin, A. M., Tommassen, J. P., Egmond, M. R., Trent, M. S., and Gros, P. (2009) Active-site architecture and catalytic mechanism of the lipid A deacylase LpxR of *Salmonella typhimurium*. *Proc Natl Acad Sci U S A 106*, 1960-4.

(105) Manabe, T., and Kawasaki, K. (2008) Extracellular loops of lipid A 3-*O*-deacylase PagL are involved in recognition of aminoarabinose-based membrane modifications in Salmonella enterica serovar typhimurium. *J Bacteriol 190*, 5597-606.

(106) Kawasaki, K., China, K., and Nishijima, M. (2007) Release of the lipopolysaccharide deacylase PagL from latency compensates for a lack of lipopolysaccharide aminoarabinose modification-dependent resistance to the antimicrobial peptide polymyxin B in Salmonella enterica. *J Bacteriol 189*, 4911-9.

(107) Geurtsen, J., Steeghs, L., Hamstra, H. J., Ten Hove, J., de Haan, A., Kuipers, B., Tommassen, J., and van der Ley, P. (2006) Expression of the lipopolysaccharide-modifying enzymes PagP and PagL modulates the endotoxic activity of *Bordetella pertussis*. *Infect Immun 74*, 5574-85.

(108) Ernst, R. K., Adams, K. N., Moskowitz, S. M., Kraig, G. M., Kawasaki, K., Stead, C. M., Trent, M. S., and Miller, S. I. (2006) The *Pseudomonas aeruginosa* lipid A deacylase: selection for expression and loss within the cystic fibrosis airway. *J Bacteriol 188*, 191-201.

(109) Rutten, L., Geurtsen, J., Lambert, W., Smolenaers, J. J., Bonvin, A. M., de Haan, A., van der Ley, P., Egmond, M. R., Gros, P., and Tommassen, J. (2006) Crystal structure and catalytic mechanism of the LPS 3-*O*-deacylase PagL from *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A 103*, 7071-6.

(110) Hwang, P. M., Choy, W. Y., Lo, E. I., Chen, L., Forman-Kay, J. D., Raetz, C. R., Prive, G. G., Bishop, R. E., and Kay, L. E. (2002) Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc Natl Acad Sci U S A 99*, 13560-5.

(111) Ahn, V. E., Lo, E. I., Engel, C. K., Chen, L., Hwang, P. M., Kay, L. E., Bishop, R. E., and Prive, G. G. (2004) A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. *Embo J 23*, 2931-41.

(112) Gouaux, E. (1998) Roll out the barrel. *Nat Struct Biol 5*, 931-2.

(113) Guo, L., Lim, K. B., Poduje, C. M., Daniel, M., Gunn, J. S., Hackett, M., and Miller, S. I. (1998) Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides. *Cell 95*, 189-98.

(114) Preston, A., Maxim, E., Toland, E., Pishko, E. J., Harvill, E. T., Caroff, M., and Maskell, D. J. (2003) *Bordetella bronchiseptica* PagP is a Bvg-regulated lipid A palmitoyl transferase that is required for persistent colonization of the mouse respiratory tract. *Mol Microbiol 48*, 725-36.

(115) Robey, M., O'Connell, W., and Cianciotto, N. P. (2001) Identification of *Legionella pneumophila rcp*, a *pagP*-like gene that confers resistance to cationic antimicrobial peptides and promotes intracellular infection. *Infect Immun 69*, 4276-86.

(116) Fukuoka, S., Brandenburg, K., Muller, M., Lindner, B., Koch, M. H., and Seydel, U. (2001) Physico-chemical analysis of lipid A fractions of lipopolysaccharide

from *Erwinia carotovora* in relation to bioactivity. *Biochim Biophys Acta 1510*, 185-97.

(117) Derzelle, S., Turlin, E., Duchaud, E., Pages, S., Kunst, F., Givaudan, A., and Danchin, A. (2004) The PhoP-PhoQ two-component regulatory system of *Photorhabdus luminescens* is essential for virulence in insects. *J Bacteriol 186*, 1270-9.

(118) Pilione, M. R., Pishko, E. J., Preston, A., Maskell, D. J., and Harvill, E. T. (2004) *pagP* is required for resistance to antibody-mediated complement lysis during *Bordetella bronchiseptica* respiratory infection. *Infect Immun 72*, 2837-42.

(119) Preston, A., Parkhill, J., and Maskell, D. J. (2004) The bordetellae: lessons from genomics. *Nat Rev Microbiol 2*, 379-90.

(120) Rebeil, R., Ernst, R. K., Gowen, B. B., Miller, S. I., and Hinnebusch, B. J. (2004) Variation in lipid A structure in the pathogenic yersiniae. *Mol Microbiol 52*, 1363-73.

(121) Leive, L. (1965) Release of lipopolysaccharide by EDTA treatment of *E. coli*. *Biochem Biophys Res Commun 21*, 290-6.

(122) Nikaido, H., and Nakae, T. (1979) The outer membrane of Gram-negative bacteria. *Adv Microb Physiol 20*, 163-250.

(123) Jia, W., Zoeiby, A. E., Petruzziello, T. N., Jayabalasingham, B., Seyedirashti, S., and Bishop, R. E. (2004) Lipid trafficking controls endotoxin acylation in outer membranes of *Escherichia coli*. *J Biol Chem 279*, 44966-75.

(124) Whitelegge, J. P., le Coutre, J., Lee, J. C., Engel, C. K., Prive, G. G., Faull, K. F., and Kaback, H. R. (1999) Toward the bilayer proteome, electrospray ionization-mass spectrometry of large, intact transmembrane proteins. *Proc Natl Acad Sci U S A 96*, 10695-8.

(125) Hwang, P. M., Bishop, R. E., and Kay, L. E. (2004) The integral membrane enzyme PagP alternates between two dynamically distinct states. *Proc Natl Acad Sci U S A 101*, 9618-23.

(126) Khan, M. A., Neale, C., Michaux, C., Pomes, R., Prive, G. G., Woody, R. W., and Bishop, R. E. (2007) Gauging a hydrocarbon ruler by an intrinsic exciton probe. *Biochemistry 46*, 4565-79.

(127) Loo, T. W., Bartlett, M. C., and Clarke, D. M. (2004) Val133 and Cys137 in transmembrane segment 2 are close to Arg935 and Gly939 in transmembrane segment 11 of human P-glycoprotein. *J Biol Chem 279*, 18232-8.

(128) Kim, J. S., and Raines, R. T. (1995) Dibromobimane as a fluorescent crosslinking reagent. *Anal Biochem 225*, 174-6.

(129) Loo, T. W., and Clarke, D. M. (2000) Identification of residues within the drug-binding domain of the human multidrug resistance P-glycoprotein by cysteine-scanning mutagenesis and reaction with dibromobimane. *J Biol Chem 275*, 39272-8.

(130) Wu, J., Voss, J., Hubbell, W. L., and Kaback, H. R. (1996) Site-directed spin labeling and chemical crosslinking demonstrate that helix V is close to helices VII and VIII in the lactose permease of *Escherichia coli*. *Proc Natl Acad Sci U S A 93*, 10123-7.

(131) Bhattacharjee, H., and Rosen, B. P. (1996) Spatial proximity of Cys113, Cys172, and Cys422 in the metalloactivation domain of the ArsA ATPase. *J Biol Chem 271*, 24465-70.

(132) Vance, D. E., and Vance, J. E. (2002) *Biochemistry of Lipids, Lipoproteins, and Membranes.*, Amsterdam: Elsevier.

(133) Dowhan, W. (1997) Molecular basis for membrane phospholipid diversity: why are there so many lipids? *Annu Rev Biochem 66*, 199-232.

(134) Zhang, Y. M., White, S. W., and Rock, C. O. (2006) Inhibiting bacterial fatty acid synthesis. *J Biol Chem 281*, 17541-4.

(135) Bracey, M. H., Cravatt, B. F., and Stevens, R. C. (2004) Structural commonalities among integral membrane enzymes. *FEBS Lett 567*, 159-65.

(136) Long, S. B., Casey, P. J., and Beese, L. S. (1998) Cocrystal structure of protein farnesyltransferase complexed with a farnesyl diphosphate substrate. *Biochemistry 37*, 9612-8.

(137) Wyckoff, T. J., Lin, S., Cotter, R. J., Dotson, G. D., and Raetz, C. R. (1998) Hydrocarbon rulers in UDP-*N*-acetylglucosamine acyltransferases. *J Biol Chem 273*, 32369-72.

(138) White, S. H., Ladokhin, A. S., Jayasinghe, S., and Hristova, K. (2001) How membranes shape protein structure. *J Biol Chem 276*, 32395-8.

(139) Reynolds, C. M., Kalb, S. R., Cotter, R. J., and Raetz, C. R. (2005) A phosphoethanolamine transferase specific for the outer 3-deoxy-D-*manno*-octulosonic acid residue of *Escherichia coli* lipopolysaccharide. Identification of the eptB gene and Ca^{2+} hypersensitivity of an eptB deletion mutant. *J Biol Chem 280*, 21202-11.

(140) Harada, N., and Nakanishi, K. (1983) *Circular dichroic spectroscopy : exciton coupling in organic stereochemistry*, University Science Books, [Mill Valley, Calif.].

(141) Schellman, J. A. (1968) Symmetry rules for optical rotation. *Acc Chem Res 1*, 144-51.

(142) Davydov, A. S. (1971) *Theory of molecular excitons. Translated from Russian by Stephen B. Dresner*, New American ed., Plenum Press, New York,.

(143) Woody, R. W. (1995) Circular dichroism. *Methods Enzymol 246*, 34-71.

(144) Woody, R. W. (1994) Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins. *Eur Biophys J 23*, 253-62.

(145) Sreerama, N., and Woody, R. W. (2004) Computation and analysis of protein circular dichroism spectra. *Methods Enzymol 383*, 318-51.

(146) Sreerama, N., and Woody, R. W. (2004) On the analysis of membrane protein circular dichroism spectra. *Protein Sci 13*, 100-12.

(147) Grishina, I. B., and Woody, R. W. (1994) Contributions of tryptophan side chains to the circular dichroism of globular proteins: exciton couplets and coupled oscillators. *Faraday Discuss*, 245-62.

(148) Bishop, R. E. (2000) The bacterial lipocalins. *Biochim Biophys Acta 1482*, 73-83.

(149) Kamio, Y., and Nikaido, H. (1976) Outer membrane of *Salmonella typhimurium*: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry 15*, 2561-70.

(150) Smith, A. E., Kim, S. H., Liu, F., Jia, W., Vinogradov, E., Gyles, C. L., and Bishop, R. E. (2008) PagP activation in the outer membrane triggers R3 core

oligosaccharide truncation in the cytoplasm of *Escherichia coli* O157:H7. *J Biol Chem 283*, 4332-43.

(151) Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid. *Anal Biochem 150*, 76-85.

(152) Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry 6*, 1948-54.

(153) Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci 4*, 2411-23.

(154) Woody, R. W., and Sreerama, N. (1999) Comment on 'Improving protein circular dichroism calculations in the far-ultraviolet through reparameterizing the amide chromophore' [J. Chem. Phys. 109, 782 (1998)]. *J Chem Phys 111*, 2844-5.

(155) Kanipes, M. I., Lin, S., Cotter, R. J., and Raetz, C. R. (2001) Ca^{2+}-induced phosphoethanolamine transfer to the outer 3-deoxy-D-*manno*-octulosonic acid moiety of *Escherichia coli* lipopolysaccharide. A novel membrane enzyme dependent upon phosphatidylethanolamine. *J Biol Chem 276*, 1156-63.

(156) Lee, C. H., and Tsai, C. M. (1999) Quantification of bacterial lipopolysaccharides by the purpald assay: measuring formaldehyde generated from 2-keto-3-deoxyoctonate and heptose at the inner core by periodate oxidation. *Anal Biochem 267*, 161-8.

(157) Feller, S. E., and MacKerell, A. D. (2000) An improved empirical potential energy function for molecular simulations of phospholipids. *J Phys Chem B 104*, 7510-15.

(158) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys 79*, 926-35.

(159) Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry 4*, 187-217.

(160) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph 14*, 33-8, 27-8.

(161) Smart, O. S., Goodfellow, J. M., and Wallace, B. A. (1993) The pore dimensions of gramicidin A. *Biophys J 65*, 2455-60.

(162) Heinrikson, R. L. (1971) The selective S-methylation of sulfhydryl groups in proteins and peptides with methyl-*p*-nitrobenzenesulfonate. *J Biol Chem 246*, 4090-6.

(163) Saxena, V. P., and Wetlaufer, D. B. (1971) A new basis for interpreting the circular dichroic spectra of proteins. *Proc Natl Acad Sci U S A 68*, 969-72.

(164) Gellman, S. H. (1991) On the role of methionine residues in the sequence-independent recognition of nonpolar protein surfaces. *Biochemistry 30*, 6633-6.

(165) Gasymov, O. K., Abduragimov, A. R., Yusifov, T. N., and Glasgow, B. J. (2002) RET and anisotropy measurements establish the proximity of the conserved Trp17 to Ile98 and Phe99 of tear lipocalin. *Biochemistry 41*, 8837-48.

(166) Nakao, M., Maki, K., Arai, M., Koshiba, T., Nitta, K., and Kuwajima, K. (2005) Characterization of kinetic folding intermediates of recombinant canine milk lysozyme by stopped-flow circular dichroism. *Biochemistry 44*, 6685-92.

(167) McConn, J., Fasman, G. D., and Hess, G. P. (1969) Conformation of the high pH form of chymotrypsin. *J Mol Biol 39*, 551-62.

(168) Kuwajima, K., Garvey, E. P., Finn, B. E., Matthews, C. R., and Sugai, S. (1991) Transient intermediates in the folding of dihydrofolate reductase as detected by far-ultraviolet circular dichroism spectroscopy. *Biochemistry 30*, 7693-703.

(169) Svensson, A. K., O'Neill, J. C., Jr., and Matthews, C. R. (2003) The coordination of the isomerization of a conserved non-prolyl cis peptide bond with the rate-limiting steps in the folding of dihydrofolate reductase. *J Mol Biol 326*, 569-83.

(170) Bowie, J. U. (2005) Solving the membrane protein folding problem. *Nature 438*, 581-9.

(171) Gray, T. M., and Matthews, B. W. (1984) Intrahelical hydrogen bonding of serine, threonine and cysteine residues within alpha-helices and its relevance to membrane-bound proteins. *J Mol Biol 175*, 75-81.

(172) Duan, G., Smith Jr., V. H., and Weaver, D. F. (2001) Characterization of aromatic-thiol pi-type hydrogen bonding and phenylalanine-cysteine side chain interactions through ab initio calculations and protein database analyses. *Mol Phys 99*, 1689-1699.

(173) Smirnova, I. N., and Kaback, H. R. (2003) A mutation in the lactose permease of *Escherichia coli* that decreases conformational flexibility and increases protein stability. *Biochemistry 42*, 3025-31.

(174) Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H. R., and Iwata, S. (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science 301*, 610-5.

(175) Ermolova, N. V., Smirnova, I. N., Kasho, V. N., and Kaback, H. R. (2005) Interhelical packing modulates conformational flexibility in the lactose permease of *Escherichia coli*. *Biochemistry 44*, 7669-77.

(176) Bishop, R. E. (2005) The lipid A palmitoyltransferase PagP: molecular mechanisms and role in bacterial pathogenesis. *Mol Microbiol 57*, 900-12.

(177) Muroi, M., Ohnishi, T., and Tanamoto, K. (2002) MD-2, a novel accessory molecule, is involved in species-specific actions of Salmonella lipid A. *Infect Immun 70*, 3546-50.

(178) Tanamoto, K., and Azumi, S. (2000) Salmonella-type heptaacylated lipid A is inactive and acts as an antagonist of lipopolysaccharide action on human line cells. *J Immunol 164*, 3149-56.

(179) van Meer, G. (2005) Cellular lipidomics. *Embo J 24*, 3159-65.

(180) Forneris, F., and Mattevi, A. (2008) Enzymes without borders: mobilizing substrates, delivering products. *Science 321*, 213-6.

(181) Bishop, R. E. (2008) Structural biology of membrane-intrinsic beta-barrel enzymes: sentinels of the bacterial outer membrane. *Biochim Biophys Acta 1778*, 1881-96.

(182) Khan, M. A., and Bishop, R. E. (2009) Molecular mechanism for lateral lipid diffusion between the outer membrane external leaflet and a beta-barrel hydrocarbon ruler. *Biochemistry 48*, 9745-56.

(183) Harris, T. K., and Turner, G. J. (2002) Structural basis of perturbed pKa values of catalytic groups in enzyme active sites. *IUBMB Life 53*, 85-98.

(184) Pace, C. N., Grimsley, G. R., and Scholtz, J. M. (2009) Protein Ionizable Groups: pK Values and Their Contribution to Protein Stability and Solubility. *J Biol Chem 284*, 13285-9.

(185) Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol 177*, 4121-30.

(186) Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning : a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

(187) Casadaban, M. J. (1976) Transposition and fusion of the lac genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J Mol Biol 104*, 541-55.

(188) Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol 2*, 2006 0008.

(189) Silhavy, T. J., Berman, M. L., and Enquist, L. W. (1984) *Experiments with gene fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

(190) Cherepanov, P. P., and Wackernagel, W. (1995) Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene 158*, 9-14.

(191) Zhou, Z., Lin, S., Cotter, R. J., and Raetz, C. R. (1999) Lipid A modifications characteristic of *Salmonella typhimurium* are induced by NH_4VO_3 in *Escherichia coli* K12. Detection of 4-amino-4-deoxy-$_L$-arabinose, phosphoethanolamine and palmitate. *J Biol Chem 274*, 18503-14.

(192) Odegaard, T. J., Kaltashov, I. A., Cotter, R. J., Steeghs, L., van der Ley, P., Khan, S., Maskell, D. J., and Raetz, C. R. (1997) Shortened hydroxyacyl chains on lipid A of *Escherichia coli* cells expressing a foreign UDP-*N*-acetylglucosamine O-acyltransferase. *J Biol Chem 272*, 19688-96.

(193) Oursel, D., Loutelier-Bourhis, C., Orange, N., Chevalier, S., Norris, V., and Lange, C. M. (2007) Lipid composition of membranes of *Escherichia coli* by liquid chromatography/tandem mass spectrometry using negative electrospray ionization. *Rapid Commun Mass Spectrom 21*, 1721-8.

(194) Cronan, J. E., Jr. (1968) Phospholipid alterations during growth of *Escherichia coli*. *J Bacteriol 95*, 2054-61.

(195) Ho, M. C., Menetret, J. F., Tsuruta, H., and Allen, K. N. (2009) The origin of the electrostatic perturbation in acetoacetate decarboxylase. *Nature 459*, 393-7.

(196) Lo Bello, M., Parker, M. W., Desideri, A., Polticelli, F., Falconi, M., Del Boccio, G., Pennelli, A., Federici, G., and Ricci, G. (1993) Peculiar spectroscopic and kinetic properties of Cys-47 in human placental glutathione transferase. Evidence for an atypical thiolate ion pair near the active site. *J Biol Chem 268*, 19033-8.

(197) Voss, J., Sun, J., Venkatesan, P., and Kaback, H. R. (1998) Sulfhydryl oxidation of mutants with cysteine in place of acidic residues in the lactose permease. *Biochemistry 37*, 8191-6.

(198) Fahy, E., Subramaniam, S., Brown, H. A., Glass, C. K., Merrill, A. H., Jr., Murphy, R. C., Raetz, C. R., Russell, D. W., Seyama, Y., Shaw, W., Shimizu, T., Spener, F., van Meer, G., VanNieuwenhze, M. S., White, S. H., Witztum, J. L., and Dennis, E. A. (2005) A comprehensive classification system for lipids. *J Lipid Res 46*, 839-61.

(199) White, S. H. (2009) Biophysical dissection of membrane proteins. *Nature 459*, 344-6.

(200) Martinez Molina, D., Wetterholm, A., Kohl, A., McCarthy, A. A., Niegowski, D., Ohlson, E., Hammarberg, T., Eshaghi, S., Haeggstrom, J. Z., and Nordlund, P. (2007) Structural basis for synthesis of inflammatory mediators by human leukotriene C4 synthase. *Nature 448*, 613-6.

(201) Ferguson, A. D., McKeever, B. M., Xu, S., Wisniewski, D., Miller, D. K., Yamin, T. T., Spencer, R. H., Chu, L., Ujjainwalla, F., Cunningham, B. R., Evans, J. F., and Becker, J. W. (2007) Crystal structure of inhibitor-bound human 5-lipoxygenase-activating protein. *Science 317*, 510-2.

(202) Ago, H., Kanaoka, Y., Irikura, D., Lam, B. K., Shimamura, T., Austen, K. F., and Miyano, M. (2007) Crystal structure of a human membrane protein involved in cysteinyl leukotriene biosynthesis. *Nature 448*, 609-12.

(203) van Horn, W. D., Kim, H. J., Ellis, C. D., Hadziselimovic, A., Sulistijo, E. S., Karra, M. D., Tian, C., Sonnichsen, F. D., and Sanders, C. R. (2009) Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science 324*, 1726-9.

(204) Ruiz, N., Wu, T., Kahne, D., and Silhavy, T. J. (2006) Probing the barrier function of the outer membrane with chemical conditionality. *ACS Chem Biol 1*, 385-95.

(205) Murata, T., Tseng, W., Guina, T., Miller, S. I., and Nikaido, H. (2007) PhoPQ-mediated regulation produces a more robust permeability barrier in the outer membrane of Salmonella enterica serovar typhimurium. *J Bacteriol 189*, 7213-22.

(206) Malinverni, J. C., and Silhavy, T. J. (2009) An ABC transport system that maintains lipid asymmetry in the Gram-negative outer membrane. *Proc Natl Acad Sci U S A 106*, 8009-14.

(207) Gao, J., Bosco, D. A., Powers, E. T., and Kelly, J. W. (2009) Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat Struct Mol Biol 16*, 684-90.

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

(209) Lee, G. F., Lebert, M. R., Lilly, A. A., and Hazelbauer, G. L. (1995) Transmembrane signaling characterized in bacterial chemoreceptors by using sulfhydryl cross-linking in vivo. *Proc Natl Acad Sci U S A 92*, 3391-5.

(210) Ellman, G. L. (1959) Tissue sulfhydryl groups. *Arch Biochem Biophys 82*, 70-7.

(211) Bishop, R. E., and Weiner, J. H. (1993) Overproduction, solubilization, purification and DNA-binding properties of AmpR from Citrobacter freundii. *Eur J Biochem 213*, 405-12.

(212) Sreerama, N., Manning, M. C., Powers, M. E., Zhang, J. X., Goldenberg, D. P., and Woody, R. W. (1999) Tyrosine, phenylalanine, and disulfide contributions to the circular dichroism of proteins: circular dichroism spectra of wild-type and mutant bovine pancreatic trypsin inhibitor. *Biochemistry 38*, 10814-22.

(213) Bewley, T. A. (1977) Optical activity of disulfide bonds in proteins. 1. Studies on plasmin modified human somatotropin. *Biochemistry 16*, 209-15.

(214) Kurapkat, G., Kruger, P., Wollmer, A., Fleischhauer, J., Kramer, B., Zobel, E., Koslowski, A., Botterweck, H., and Woody, R. W. (1997) Calculations of the CD spectrum of bovine pancreatic ribonuclease. *Biopolymers 41*, 267-87.

(215) Bewley, T. A. (1977) Optical activity of disulfide bonds in proteins: studies on human choriomammotropin and bovine pituitary somatotropin. *Biochemistry 16*, 4408-14.

(216) Huysmans, G. H., Radford, S. E., Brockwell, D. J., and Baldwin, S. A. (2007) The N-terminal helix is a post-assembly clamp in the bacterial outer membrane protein PagP. *J Mol Biol 373*, 529-40.

(217) Michaux, C., Pomroy, N. C., and Prive, G. G. (2008) Refolding SDS-denatured proteins by the addition of amphipathic cosolvents. *J Mol Biol 375*, 1477-88.

(218) Bishop, R. E., and Raetz, C. R. (2000) in *Fortieth Interscience Conference on Antimicrobial Agents and Chemotherapy*, Toronto, ON, Canada.

(219) Kadokura, H., Katzen, F., and Beckwith, J. (2003) Protein disulfide bond formation in prokaryotes. *Annu Rev Biochem 72*, 111-35.

(220) Mogensen, J. E., and Otzen, D. E. (2005) Interactions between folding factors and bacterial outer membrane proteins. *Mol Microbiol 57*, 326-46.

(221) Bos, M. P., Robert, V., and Tommassen, J. (2007) Biogenesis of the Gram-negative bacterial outer membrane. *Annu Rev Microbiol 61*, 191-214.

(222) Ruiz, N., Kahne, D., and Silhavy, T. J. (2006) Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol 4*, 57-66.

(223) Sperandeo, P., Deho, G., and Polissi, A. (2009) The lipopolysaccharide transport system of Gram-negative bacteria. *Biochim Biophys Acta 1791*, 594-602.

(224) Segel, I. H. (1975) *Enzyme kinetics : behavior and analysis of rapid equilibrium and steady state enzyme systems*, Wiley, New York.

(225) Koshland, D. E. (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci U S A 44*, 98-104.

(226) Hearn, E. M., Patel, D. R., Lepore, B. W., Indic, M., and van den Berg, B. (2009) Transmembrane passage of hydrophobic compounds through a protein channel wall. *Nature 458*, 367-70.

(227) Walther, D. M., Rapaport, D., and Tommassen, J. (2009) Biogenesis of beta-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell Mol Life Sci.*

(228) Helenius, A., and Simons, K. (1975) Solubilization of membranes by detergents. *Biochim Biophys Acta 415*, 29-79.

(229) Guo, X. H., Zhao, N. M., Chen, S. H., and Teixeira, J. (1990) Small-angle neutron scattering study of the structure of protein/detergent complexes. *Biopolymers 29*, 335-46.

(230) Manning, M., and Colon, W. (2004) Structural basis of protein kinetic stability: resistance to sodium dodecyl sulfate suggests a central role for rigidity and a bias toward beta-sheet structure. *Biochemistry 43*, 11248-54.

(231) Mattice, W. L., Riser, J. M., and Clark, D. S. (1976) Conformational properties of the complexes formed by proteins and sodium dodecyl sulfate. *Biochemistry 15*, 4264-72.

(232) Nielsen, M. M., Andersen, K. K., Westh, P., and Otzen, D. E. (2007) Unfolding of beta-sheet proteins in SDS. *Biophys J 92*, 3674-85.

(233) Otzen, D. E., and Oliveberg, M. (2002) Burst-phase expansion of native protein prior to global unfolding in SDS. *J Mol Biol 315*, 1231-40.

(234) Otzen, D. E. (2002) Protein unfolding in detergents: effect of micelle structure, ionic strength, pH, and temperature. *Biophys J 83*, 2219-30.

(235) Reynolds, J. A., and Tanford, C. (1970) The gross conformation of protein-sodium dodecyl sulfate complexes. *J Biol Chem 245*, 5161-5.

(236) Samso, M., Daban, J. R., Hansen, S., and Jones, G. R. (1995) Evidence for sodium dodecyl sulfate/protein complexes adopting a necklace structure. *Eur J Biochem 232*, 818-24.

(237) Smith, B. J. (1994) SDS polyacrylamide gel electrophoresis of proteins. *Methods Mol Biol 32*, 23-34.

(238) Anand, K., Pal, D., and Hilgenfeld, R. (2002) An overview on 2-methyl-2,4-pentanediol in crystallization and in crystals of biological macromolecules. *Acta Crystallogr D Biol Crystallogr 58*, 1722-8.

(239) Prive, G. G. (2007) Detergents for the stabilization and crystallization of membrane proteins. *Methods 41*, 388-97.

(240) Arand, M., Friedberg, T., and Oesch, F. (1992) Colorimetric quantitation of trace amounts of sodium lauryl sulfate in the presence of nucleic acids and proteins. *Anal Biochem 207*, 73-5.

(241) Berger, O., Edholm, O., and Jahnig, F. (1997) Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J 72*, 2002-13.

(242) Tieleman, D. P., Forrest, L. R., Sansom, M. S., and Berendsen, H. J. (1998) Lipid properties and the orientation of aromatic residues in OmpF, influenza M2, and alamethicin systems: molecular dynamics simulations. *Biochemistry 37*, 17554-61.

(243) Kaminski, G., Friesner, R., Tirado-Rives, J., and Jorgensen, W. (2001) Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *Journal of Physical Chemistry B 105*, 6474-6487.

(244) Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling 7*, 306-317.

(245) Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr 55*, 484-91.

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

(247) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr 53*, 240-55.

(248) Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr 59*, 1131-7.

(249) Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and

Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res 35*, W375-83.

(250) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography 26*, 283-291.

(251) Coiro, V., and Mazza, F. (1991) Crystal phases of dodecyl sulfates obtained from aqueous solutions: structure of the hexaaquamagnesium salt. *Crystal Structure Communications 47*, 1169-1173.

(252) Kabsch, W. (1976) A solution for the best rotation to relate two sets of vectors. *Crystal Physics, Diffraction, Theoretical and General Crystallography 32*, 922-923.

(253) Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, R., and Klein, M. (1983) Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics 79*, 926.

(254) van Gunsteren, W., and Berendsen, H. (1990) Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. *Angewandte Chemie International Edition in English 29*, 992-1023.

(255) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N+log(N) method for Ewald sums in large systems. *Journal of Chemical Physics 98*, 10089-10092.

(256) Essmann, U., Perera, L., Berkowitz, M., Darden, T., Lee, H., and Pedersen, L. (1995) A smooth particle mesh Ewald method. *The Journal of Chemical Physics 103*, 8577.

(257) Berendsen, H., Postma, J., van Gunsteren, W., DiNola, A., and Haak, J. (1984) Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics 81*, 3684.

(258) Miyamoto, S., and Kollman, P. (1992) Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry 13*, 952-962.

(259) Hess, B., Bekker, H., Berendsen, H., and Fraaije, J. (1997) *J. Comput. Chem. 18*, 1463-1472.

(260) Elder, M., Hitchcock, P., Mason, R., and Shipley, G. (1977) A Refinement Analysis of the Crystallography of the Phospholipid, 1, 2-Dilauroyl-DL-Phosphatidylethanolamine, and Some Remarks on Lipid-Lipid and Lipid-Protein Interactions. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences (1934-1990) 354*, 157-170.

(261) Takaoka, Y., Pasenkiewicz-Gierula, M., Miyagawa, H., Kitamura, K., Tamura, Y., and Kusumi, A. (2000) Molecular Dynamics Generation of Nonarbitrary Membrane Models Reveals Lipid Orientational Correlations. *Biophysical Journal 79*, 3118-3138.

(262) Xiang, Z., Soto, C. S., and Honig, B. (2002) Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. *Proc Natl Acad Sci U S A 99*, 7432-7.

(263) Canutescu, A. A., Shelenkov, A. A., and Dunbrack, R. L., Jr. (2003) A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci 12*, 2001-14.

(264) Tieleman, D., van der Spoel, D., and Berendsen, H. (2000) Molecular Dynamics Simulations of Dodecylphosphocholine Micelles at Three Different Aggregate Sizes: Micellar Structure and Chain Relaxation. *Journal of Physical Chemistry A 104*, 6380-6388.

(265) Sanner, M. F., Olson, A. J., and Spehner, J. C. (1996) Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers 38*, 305-20.

(266) Faraldo-Gomez, J. D., Smith, G. R., and Sansom, M. S. (2002) Setting up and optimization of membrane protein simulations. *Eur Biophys J 31*, 217-27.

(267) Tieleman, D., MacCallum, J., Ash, W., Kandt, C., Xu, Z., and Monticelli, L. (2006) Membrane protein simulations with a united-atom lipid and all-atom protein model: lipid-protein interactions, side chain transfer free energies and model proteins. *Journal of Physics Condensed Matter 18*, 1221.

(268) Berendsen, H., Postma, J., van Gunsteren, W., and Hermans, J. (1981) Interaction models for water in relation to protein hydration, in *Intermolecular Forces* (Pullman, B., Ed.) pp 331-342, Reidel, Dordrecht, The Netherlands.

(269) Tieleman, D., and Berendsen, H. (1996) Molecular dynamics simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer with different macroscopic boundary conditions and parameters. *The Journal of Chemical Physics 105*, 4871.

(270) Hunte, C., and Richers, S. (2008) Lipids and membrane protein structures. *Curr Opin Struct Biol 18*, 406-11.

(271) Qin, L., Hiser, C., Mulichak, A., Garavito, R. M., and Ferguson-Miller, S. (2006) Identification of conserved lipid/detergent-binding sites in a high-resolution structure of the membrane protein cytochrome c oxidase. *Proc Natl Acad Sci U S A 103*, 16117-22.

(272) Michaux, C., Pouyez, J., Wouters, J., and Privé, G. G. (2008) Protecting role of cosolvents in protein denaturation by SDS: a structural study. *BMC Struct Biol 8*, 29.

(273) Huber, G. A., and Kim, S. (1996) Weighted-ensemble Brownian dynamics simulations for protein association reactions. *Biophys J 70*, 97-110.

(274) Khan, M. A., and Bishop, R. E. (2009) Molecular Mechanism for Lateral Lipid Diffusion between the Outer Membrane External Leaflet and a beta-Barrel Hydrocarbon Ruler. *Biochemistry*.

(275) Strop, P., and Brunger, A. T. (2005) Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci 14*, 2207-11.

(276) Gelb, M. H., Min, J. H., and Jain, M. K. (2000) Do membrane-bound enzymes access their substrates from the membrane or aqueous phase: interfacial versus non-interfacial enzymes. *Biochim Biophys Acta 1488*, 20-7.

(277) Carman, G. M., Deems, R. A., and Dennis, E. A. (1995) Lipid signaling enzymes and surface dilution kinetics. *J Biol Chem 270*, 18711-4.

(278) Creer, M. H., and Gross, R. W. (1985) Separation of isomeric lysophospholipids by reverse phase HPLC. *Lipids 20*, 922-8.

(279) Pluckthun, A., and Dennis, E. A. (1982) Role of monomeric activators in cobra venom phospholipase A2 action. *Biochemistry 21*, 1750-6.

(280) Nishijima, M., Sa-Eki, T., Tamori, Y., Doi, O., and Nojima, S. (1978) Synthesis of acyl phosphatidylglycerol from phosphatidylglycerol in *Escherichia coli* K-12. Evidence for the participation of detergent-resistant phospholipase A and heat-labile membrane-bound factor(s). *Biochim Biophys Acta 528*, 107-18.

(281) Thanassi, D. G., Stathopoulos, C., Karkal, A., and Li, H. (2005) Protein secretion in the absence of ATP: the autotransporter, two-partner secretion and chaperone/usher pathways of Gram-negative bacteria (review). *Mol Membr Biol 22*, 63-72.

(282) Ito, K., and Inaba, K. (2008) The disulfide bond formation (Dsb) system. *Curr Opin Struct Biol 18*, 450-8.

(283) Ross, A. C., Zolfaghari, R., and Weisz, J. (2001) Vitamin A: recent advances in the biotransformation, transport, and metabolism of retinoids. *Curr Opin Gastroenterol 17*, 184-192.

(284) Saari, J. C. (2000) Biochemistry of visual pigment regeneration: the Friedenwald lecture. *Invest Ophthalmol Vis Sci 41*, 337-48.

(285) Rousset, X., Vaisman, B., Amar, M., Sethi, A. A., and Remaley, A. T. (2009) Lecithin: cholesterol acyltransferase--from biochemistry to role in cardiovascular disease. *Curr Opin Endocrinol Diabetes Obes 16*, 163-71.

(286) Warden, C. H., Langner, C. A., Gordon, J. I., Taylor, B. A., McLean, J. W., and Lusis, A. J. (1989) Tissue-specific expression, developmental regulation, and chromosomal mapping of the lecithin: cholesterol acyltransferase gene. Evidence for expression in brain and testes as well as liver. *J Biol Chem 264*, 21573-81.

(287) Hixson, J. E., Driscoll, D. M., Birnbaum, S., and Britten, M. L. (1993) Baboon lecithin cholesterol acyltransferase (LCAT): cDNA sequences of two alleles, evolution, and gene expression. *Gene 128*, 295-9.

(288) Smith, K. M., Lawn, R. M., and Wilcox, J. N. (1990) Cellular localization of apolipoprotein D and lecithin:cholesterol acyltransferase mRNA in rhesus monkey tissues by in situ hybridization. *J Lipid Res 31*, 995-1004.

(289) Vaisman, B. L., Klein, H. G., Rouis, M., Berard, A. M., Kindt, M. R., Talley, G. D., Meyn, S. M., Hoyt, R. F., Jr., Marcovina, S. M., Albers, J. J., and et al. (1995) Overexpression of human lecithin cholesterol acyltransferase leads to hyperalphalipoproteinemia in transgenic mice. *J Biol Chem 270*, 12269-75.

(290) Yang, C. Y., Manoogian, D., Pao, Q., Lee, F. S., Knapp, R. D., Gotto, A. M., Jr., and Pownall, H. J. (1987) Lecithin:cholesterol acyltransferase. Functional regions and a structural model of the enzyme. *J Biol Chem 262*, 3086-91.

(291) Grove, D., and Pownall, H. J. (1991) Comparative specificity of plasma lecithin:cholesterol acyltransferase from ten animal species. *Lipids 26*, 416-20.

(292) Portman, O. W., and Sugano, M. (1964) Factors Influencing the Level and Fatty Acid Specificity of the Cholesterol Esterification Activity in Human Plasma. *Arch Biochem Biophys 105*, 532-40.

(293) Buddelmeijer, N., and Young, R. The essential Escherichia coli apolipoprotein N-acyltransferase (Lnt) exists as an extracytoplasmic thioester acyl-enzyme intermediate. *Biochemistry 49*, 341-6.

(294) Fukuda, A., Matsuyama, S., Hara, T., Nakayama, J., Nagasawa, H., and Tokuda, H. (2002) Aminoacylation of the N-terminal cysteine is essential for Lol-dependent release of lipoproteins from membranes but does not depend on lipoprotein sorting signals. *J Biol Chem 277*, 43512-8.

(295) Vidal-Ingigliardi, D., Lewenza, S., and Buddelmeijer, N. (2007) Identification of essential residues in apolipoprotein N-acyl transferase, a member of the CN hydrolase family. *J Bacteriol 189*, 4456-64.

(296) Tamai, R., Asai, Y., Hashimoto, M., Fukase, K., Kusumoto, S., Ishida, H., Kiso, M., and Ogawa, T. (2003) Cell activation by monosaccharide lipid A analogues utilizing Toll-like receptor 4. *Immunology 110*, 66-72.

(297) Stover, A. G., Da Silva Correia, J., Evans, J. T., Cluff, C. W., Elliott, M. W., Jeffery, E. W., Johnson, D. A., Lacy, M. J., Baldridge, J. R., Probst, P., Ulevitch, R. J., Persing, D. H., and Hershberg, R. M. (2004) Structure-activity relationship of synthetic toll-like receptor 4 agonists. *J Biol Chem 279*, 4440-9.

(298) Hamdy, S., Elamanchili, P., Alshamsan, A., Molavi, O., Satou, T., and Samuel, J. (2007) Enhanced antigen-specific primary CD4+ and CD8+ responses by codelivery of ovalbumin and toll-like receptor ligand monophosphoryl lipid A in poly(D,L-lactic-co-glycolic acid) nanoparticles. *J Biomed Mater Res A 81*, 652-62.