

MmpL3 is a lipid transporter that binds trehalose monomycolate and phosphatidylethanolamine

Chih-Chia Su^{1,Ψ}, Philip A. Klenotic^{1,Ψ}, Jani Reddy Bolla^{2,Ψ}, Georgiana E. Purdy³, Carol V. Robinson², and Edward W. Yu^{1,*}

¹Department of Pharmacology, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA.

²Department of Chemistry, University of Oxford, South Parks Road, Oxford, OX1 3QZ, UK.

³Department of Molecular Microbiology and Immunology, Oregon Health and Science University, Portland, OR 97239, USA.

ΨC.S., P.A.K. and J.R.B. contributed equally to this work.

*To whom correspondence should be addressed. Email: edward.w.yu@case.edu

Abstract

The cell envelope of *Mycobacterium tuberculosis* is notable for the abundance of mycolic acids (MAs) essential to mycobacterial viability, and other species-specific lipids. The mycobacterial cell envelope is extremely hydrophobic which contributes to virulence and antibiotic resistance. Yet, exactly how fatty acids and lipidic elements are transported across the cell envelope for cell wall biosynthesis is unclear. Mycobacterial membrane protein Large 3 (MmpL3) is essential and required for transport of trehalose monomycolates (TMMs), precursors of MA-containing trehalose dimycolates (TDM) and mycolyl arabinogalactan peptidoglycan (mAGP), but the exact function of MmpL3 remains elusive. Here, we report a crystal structure of *M. smegmatis* MmpL3 at a resolution of 2.59 Å, revealing a monomeric molecule that is structurally distinct from all known bacterial membrane proteins. A previously unknown MmpL3 ligand, phosphatidylethanolamine (PE), was discovered inside this transporter. We also show, via native mass spectrometry, that MmpL3 specifically binds both TMM and PE, but not TDM, in the micromolar range. These observations provide insight into the function of MmpL3 and suggest a possible role for this protein in shuttling a variety of lipids to strengthen the mycobacterial cell wall.

Significance Statement

The unique architecture of the mycobacterial cell wall plays a predominant role in *Mycobacterium tuberculosis* (Mtb) pathogenesis. Mycobacterial membrane protein Large 3 (MmpL3) is essential and required for transport of trehalose monomycolates (TMMs) across the cell membrane for cell wall biosynthesis. Here we report a crystal structure of *M. smegmatis* MmpL3 that indicates a plausible pathway for TMM translocation. Native mass spectrometry suggests that the monomeric MmpL3 transporter is able to specifically bind TMM and phosphatidylethanolamine (PE) lipids. Overall our data suggest a novel mechanism for TMM transport involved in mycobacterial cell wall biogenesis.

\body

Introduction

Tuberculosis (TB) is the leading cause of mortality as a result of an infectious agent, exceeding both malaria and HIV (1, 2). In 2016, there were an estimated 10.4 million new cases of active TB and 1.7 million patients died from active TB infection (2). The causative agent, *Mycobacterium tuberculosis* (Mtb), has now infected more than one third of the world's population. The emergence of multidrug resistant TB (MDR-TB) presents an increasingly difficult therapeutic challenge, so much so that MDR-TB is now the main cause of death due to antimicrobial resistance. Unfortunately, the lethality of TB combined with its multidrug-resistant capacity has now transformed this long-neglected disease into a global health priority.

The unique architecture of the mycobacterial cell wall plays a predominant role in Mtb pathogenesis. This complex cell wall structure supports membrane stability and provides a barrier against the host environment and antibiotics, thereby contributing to the intrinsic

resistance of the organism. The outer membrane of Mtb is characterized by the abundance of very long chain mycolic acids (MAs). MAs are transported across the inner membrane as trehalose monomycolates (TMMs), then are either covalently linked to the arabinogalactan-peptidoglycan layer as mycolyl arabinogalactan peptidoglycan (mAGP) or incorporated into trehalose dimycolates (TDM) which comprises the majority of the outer leaflet of this membrane. The outer leaflet also contains other non-covalently associated lipids, such as phthiocerol dimycocerosates and sulfolipids (3). This unique construction makes the outer membrane of Mtb very rigid and extremely impermeable to a wide range of compounds, including many antibiotics (3). Additionally, the surface-exposed lipids are immunomodulatory and play a crucial role in host-pathogen interactions (4-9).

TMM biosynthesis and its transport are essential for Mtb, as it is the precursor of both mAGP and TDM in the outer leaflet of the cell envelope. TMMs are synthesized in the cytoplasm via a highly conserved and well-characterized pathway (10), targeted by the first-line anti-TB drug isoniazid (11). However, exactly how these TMMs are transported across the membrane from the cytosol to participate in mycobacterial cell wall remodeling is not well understood. Recent work demonstrated that mycobacterial membrane protein large (MmpL) transporters are critical for mycobacterial physiology and pathogenesis by shuttling fatty acids and lipid components to the mycobacterial cell wall. Mtb possesses 13 MmpL transporters (12). These membrane proteins share sequence similarity with the resistance-nodulation-cell division (RND) superfamily of transporters (13). Although MmpL3 is absolutely essential *in vivo*, there is strong evidence that MmpL4, MmpL5, MmpL7, MmpL8, MmpL10 and MmpL11 are all required for full virulence as well as the growth of mycobacteria in mouse lungs (14-18). Interestingly, of the 13 MmpLs, only MmpL3 is shown to export TMM (18, 19). Using a

spheroplast system, a recent study strongly indicated that MmpL3 is a TMM flippase (20). This biochemical analysis is in line with genetic studies that determined that MmpL3 expression is necessary for Mtb survival and that depletion of MmpL3 in *M. smegmatis* results in the accumulation of TMMs concomitant with a reduction in levels of mAGP and TDM (18, 19). Thus, MmpL3 has now been shown to be essential for the biosynthesis of mAGP and TDM, as well as their incorporation into the mycobacterial cell wall, with both processes critical for mycobacterial replication and viability (21, 22). It is worth noting that Mtb *mmpL3* can rescue the viability of the *M. smegmatis mmpL3* null mutant (18). This result indeed suggests that the two *mmpL3* orthologs can substitute each other to function. Collectively, studies from both Mtb and *M. smegmatis* have established the importance of MmpL3 in membrane biosynthesis and TB viability as well as being a target of several potent anti-TB agents.

As MmpL3 is absolutely essential for mycobacterial cell wall biogenesis, we decided to elucidate the molecular mechanism of lipid translocation across the membrane via this transporter. We here present a crystal structure of *M. smegmatis* MmpL3 to a resolution of 2.59 Å. Unexpectedly, a phosphatidylethanolamine (PE) molecule was found to be associated within a partially buried region of this protein. Using native mass spectrometry, we observed that MmpL3 is capable of binding PE and TMM, but not TDM, with dissociation constants within the micromolar range. Our work provides molecular insights into the mechanism of lipid transport via the MmpL3 membrane protein and suggests a possible role for MmpL3 in shuttling different lipids across the membrane for cell wall biogenesis.

Results

Overall structure of the *M. smegmatis* MmpL3 transporter

To obtain the structural information of MmpL3, we cloned the full-length *M. smegmatis* MmpL3 transporter, which contains 1,013 amino acids, into the *E. coli* expression vector pET15b, with a 6xHis tag at the C-terminus to generate pET15b Ω mmpL3. This MmpL3 protein was overproduced in *E. coli* BL21(DE3) Δ acrB cells and purified using Ni²⁺-affinity and Superdex 200 size exclusion columns. Native mass spectrometry (native MS) shows that the MmpL3 transporter is monomeric in detergent solution (Fig. 1A). To further confirm the oligomerization state of this protein in its native environment, we cloned and overproduced MmpL3 in *M. smegmatis* mc²155 cells. Again, the native mass spectra indicate that this purified MmpL3 protein is a monomer in detergent solution (Fig. 1B).

Crystals of MmpL3 were obtained using vapor diffusion. The best crystal diffracted X-rays to a resolution of 3.30 Å (*SI Appendix*, Table S1 and Fig. S1). However, these crystals were difficult to reproduce and most of the protein samples could not be crystallized. We then harvested several of these protein crystals and washed them with the crystallization precipitant solution. SDS-PAGE analysis indicated that the MmpL3 protein within these crystals was ~25 kDa smaller than the full-length protein. Since the C-terminal region (residues 733-1013) of MmpL3 contains a large number of prolines (52 proline residues), it is likely that this C-terminal sequence is proteolytically clipped and quite possibly unstable in solution.

We then used native MS to determine the size of the *M. smegmatis* full-length MmpL3 protein isolated from *E. coli*. Surprisingly, the data indicates that the protein sample contains three major fragments (Fig. 1C). These three fragments are shorter than the full-length protein, which is composed of 1013 residues. The observation of these three protein fragments suggests that the full-length protein may be clipped by proteases. A detailed analysis indicates that these three fragments contain residues 1-806, 1-776 and 1-763 of MmpL3. However, only a small

quantity of the fragment 1-763 was detected. These protein fragments possess the full transmembrane region, but the C-terminal end was absent. Since it has been reported that this C-terminal domain is not essential for the function of MmpL3 (23), we decided to remove this C-terminal sequence and produce the truncated MmpL3 transporter, designated to be MmpL3₇₇₃ (residues 1-773), to improve crystal quality for structural determination. Our native MS data indicate that the MmpL3₇₇₃ protein also exists as a monomer, which presents as a single protein fragment suitable for crystallization (Fig. 1D).

The best MmpL3₇₇₃ crystal diffracted X-ray to 2.59 Å resolution (Fig. 2). We determined its structure using single anomalous dispersion (SAD) (*SI Appendix*, Table S1 and Fig. S2). We then used the structure of MmpL3₇₇₃ as the template to resolve the structure of full-length MmpL3, where we found that the C-terminal residues 753-1013 were missing. The conformation of this MmpL3 structure is identical to that of MmpL3₇₇₃. Superimposition of the two structures gives a root-mean-square-deviation (r.m.s.d.) of 0.6 Å (for 722 C α atoms) (*SI Appendix*, Fig. S3). Although MmpL3 belongs to a subclass of the RND superfamily, its three-dimensional topology is unique and very different from the existing structures of RND transporters, including AcrB (24, 25), MexB (26), CusA (27, 28), MtrD (29), CmeB (30) and HpnN (31). Thus, the available structural information of RND proteins cannot be used to understand the function of this transporter.

The MmpL3₇₇₃ molecule consists of 12 transmembrane helices (TMs 1-12) and two periplasmic loops (loops 1 and 2), which create the periplasmic domain (Fig. 2A and B). Loop 1 is located between TMs 1 and 2, and loop 2 is found between TMs 7 and 8. These two loops contribute to generate subdomains PD1 and PD2 in the periplasm. Based on our structural

information, the proline-rich C-terminal residues 733-1013 should form a cytoplasmic domain (CD) of MmpL3.

The N-terminal and C-terminal halves of MmpL3₇₇₃ are assembled in a twofold pseudo-symmetrical fashion. These two halves can be superimposed with a r.m.s.d. of 2.6 Å (for 294 Cα atoms) (*SI Appendix*, Fig. S4). PD1 is composed of four α-helices and three β-strands. The majority of the PD1 amino acids come from loop 1. However, residues 437-448 of loop 2 also contribute to form helix α5 of PD1. PD2 constitutes three α-helices and three β-strands. Like PD1, the PD2 amino acids mainly arise from loop 2, but residues 49-60 of loop 1 participate in the formation of helix α1 of this periplasmic subdomain. The crossover of these two periplasmic loops allows for the two subdomains, PD1 and PD2, to be spatially adjoined within the periplasm. This structural feature is in good agreement with the finding that the two periplasmic loops closely interact with each other (32). There is a long flexible linker of 12 residues connecting the C-terminal end of TM1 and α1 of PD2. Similarly, a long, 15 residue flexible loop is found to link the C-terminal end of TM7 and α5 of PD1 together. The presence of these long linkers suggests that the periplasmic domains of MmpL3 are quite flexible in nature. The TMs are membrane embedded, but both TM2 and TM8 are significantly longer and protrude into the periplasmic region. These two TMs directly tether PD1 and PD2, respectively, and form part of the periplasmic structure of the protein.

The MmpL3₇₇₃-phosphatidylethanolamine complex

Surprisingly, the MmpL3₇₇₃ molecule forms a channel-like cavity originating from the outer leaflet of the inner membrane up to the periplasmic domain (Fig. 2C). The beginning of this cavity is generated by a hydrophobic pocket created by TMs 7-10. This pocket opens to the

outer leaflet of the inner membrane and periplasmic space. However, the majority of this cavity is found at the center of the periplasmic domain surrounded by the secondary structures of PD1 and PD2. This large space potentially constitutes a binding site for TMM.

Unexpectedly, two large extra electron densities were found within the structure of MmpL3₇₇₃. One of these densities was observed in the pocket surrounded by TMs 7-10. The shape of this extra density is compatible with a DDM detergent molecule (Fig. 3A). This is not surprising as we solubilized, purified and crystallized the MmpL3₇₇₃ transporter in solutions containing DDM detergent. Within 4.5 Å of the bound DDM, there are 13 residues, including L422, S423, L424, N524, A527, Q554, I557, F561, L564, P565, A568, L600 and I636 that provide hydrophobic and electrostatic interactions for anchoring this detergent molecule (Fig. 3A). Among these residues, the conserved leucine L564 has been observed to be important for the function of MmpL3 (33). Interestingly, within the vicinity of this binding site, we also find the conserved amino acids Q309, D555 and P630. These three residues are located at the beginning of the channel formed by the MmpL3 transporter. It has been shown that these three amino acids are important for the activity of the transporter in *M. smegmatis* (23).

The second extra density is observed within the large space between PD1 and PD2. The shape of this electron density resembles a large lipid molecule with two elongated hydrocarbon backbones (Fig. 3B). Presumably, this extra electron density belongs to a fortuitous ligand that we copurified and cocrystallized with the MmpL3₇₇₃ protein. Consistent with this observation, our MS data of purified MmpL3₇₇₃ indicate a charge state series corresponding to an adduct mass of ~700 Da (Fig. 1D). We then used liquid chromatography coupled with mass spectrometry (LC-MS) to identify this unknown bound ligand. The result indicates that this fortuitous ligand is phosphatidylethanolamine (16:0-17:1(9Z)) (PE) (*SI Appendix*, Fig. S5A). Interestingly, these PE

lipids are also observed to bind in the full-length MmpL3 protein expressed in both *E. coli* (*SI Appendix*, Fig. S5B) and *M. smegmatis* (*SI Appendix*, Fig. S5C). Because of the presence of these DDM and PE molecules, the conformation of our structure of MmpL3 should represent its ligand-bound form.

PE binds at the central cavity formed by PD1 and PD2 of the periplasmic domain. Two extended loops (residues 61-68, and residues 449-456) that both run across subdomains PD1 and PD2, are found to sandwich the bound PE. To secure the binding, residues 40-44 of the elongated loop connecting TM1 and PD2, as well as residues 426-429 that belong to the other elongated loop connecting TM7 and PD1 also participate to form the bottom of this large cavity. Several conserved amino acids are found to surround the wall of this central cavity. The binding of PE is extensive. Within 4.5 Å of the bound PE molecule, there are at least 32 amino acids involved in the binding. These residues are Q40, S41, F43, Y44, D64, T66, S67, V70, V109, T121, M125, F134, S136, D144, L171, L174, A175, Q421, I427, S428, E429, F445, F452, R453, T454, P456, R501, P502, A503, N504, Q517 and T549, which provide electrostatic and hydrophobic interactions to bind PE (Fig. 3B). Among them, the conserved residue Q40 has been reported to be crucial to the function of this transporter (23).

The proton-relay network

MmpL3 is a proton-motive-force (PMF)-dependent transporter that functions via an antiport mechanism. Coupled with the movement of substrates towards the periplasm, protons have to flow into the cytoplasm to energize this translocation process. Within the transmembrane region, a hydrogen bond is formed between the conserved residues D256 and Y646. Likewise, the nearby conserved residues Y257 and D645 also form a hydrogen-bonded pair. Residue S293,

which is also conserved, is within 3.7 Å away from D645 and interacts with this aspartate through electrostatic interaction (*SI Appendix*, Fig. S6). Within the vicinity of these residues, the crystal structure suggests that the conserved charged residues E647 and K591 also form another hydrogen bond between them. Taken together, these data suggest that these residues may form an important network for proton transfer within MmpL3. This is in good agreement with previous studies that have emphasized the importance of many of these residues to the function of this transporter (23), perhaps by creating a proton-relay network for energy coupling.

PE binding to MmpL3

As our structure of MmpL3 contained a bound PE, we wanted to further characterize this MmpL3-PE interaction. To achieve this, we performed a lipid binding experiment using native MS. A solution containing 5 µM delipidated MmpL3₇₇₃, 200 mM ammonium acetate (pH 8.0) and 0.05% (w/v) lauryldimethylamine N-oxide (LDAO) was used for the experiment. Equimolar amounts of MmpL3₇₇₃ and PE were mixed together and then the mass spectra under our optimized conditions were recorded (Fig. 4A). The experimental result depicts a charge state series for a second higher mass species at ~15% intensity of the peak corresponding to MmpL3₇₇₃. The 700 Da mass difference is consistent with the binding of one molecule PE to MmpL3 (Fig. 4A).

Subsequently, we incubated the MmpL3₇₇₃ protein with increasing concentrations of PE. At concentrations above 20 µM, we observed an increase in the intensity of the one PE-bound species, as well as an additional charge state series that corresponds binding of a second molecule of PE to MmpL3₇₇₃. At 80 µM of PE, the predominant species is the lipid-bound form of MmpL3₇₇₃. Based on our mass spectral data, we extracted and plotted the relative intensity of

PE-bound forms as a function of lipid concentration. The data suggests a dissociation constant (K_d) of $19.5 \pm 6.3 \mu\text{M}$ for the first PE binding (Fig. 4B).

TMM binding to MmpL3

Since there is strong evidence that MmpL3 is a TMM transporter, we decided to perform TMM binding experiments to probe the affinity of MmpL3 to TMM. Similar to the binding experiments described above, equimolar concentrations of MmpL3₇₇₃ and TMM were mixed together to acquire mass spectra under the conditions similar to those for PE binding. As expected, the data suggest a charge state series for a second higher mass species at ~25% intensity of the peak corresponding to MmpL3₇₇₃. The mass difference of 1,425 Da is consistent with the binding of one molecule of TMM to MmpL3₇₇₃, indicating that MmpL3 is capable of recognizing TMM (Fig. 4C).

To further define the affinity of TMM for MmpL3, we incubated the protein with increasing concentrations of TMM. As with PE binding, at concentrations higher than 20 μM TMM, we observed evolution of an additional charge state series, corresponding to binding of a second molecule of TMM to MmpL3₇₇₃. Again, we extracted and plotted the relative intensity of TMM-bound forms with respect to TMM concentration, suggesting a K_d of $3.7 \pm 1.3 \mu\text{M}$ for the first TMM binding (Fig. 4D).

MmpL3 does not bind TDM

As MmpL3 is able to specially bind TMM, we decided to elucidate if this transporter can bind TDM as well. We purified glycolipids from *M. smegmatis* mc²155 cells. MS analysis indicates that the two main species of this glycolipid sample are TDM and TMM (Fig. 5). We

then added this sample to the purified MmpL3₇₇₃ membrane protein and recorded the mass spectra. In addition to the MmpL3₇₇₃ and MmpL3₇₇₃-PE peaks, the MS data show a new charge state series corresponding to an adduct mass of ~1,428 Da, which is the mass of TMM. Surprisingly, our data show that only short chain TMM lipids (~1,428 Da) bind specifically to MmpL3₇₇₃ even though different longer chain lengths of TMM lipids are present in the solution (red inlet in Fig. 5). The result is consistent with the binding data that MmpL3₇₇₃ selectively interacts with TMM. Of specific interest, the spectra did not present any other adduct peaks that correspond to the TDM bound species, demonstrating that MmpL3₇₇₃ does not bind TDM.

To further test the interaction between MmpL3 and TDM, we performed binding experiments with the purified TDM lipid (Sigma-Aldrich, St. Louis). Again, we did not observe the event of TDM binding to the MmpL3 protein (*SI Appendix*, Fig. S7). Collectively, our data indicate that TMM, but not TDM, is preferentially bound by the MmpL3 transporter.

Other potential MmpL3 substrates

In addition to PE, our lipidomic analysis indicates that diacylglycerol (DAG) and other phospholipids, including phosphatidylglycerol (PG), phosphatidylinositol (PI) and cardiolipin (CDL), are bound by full-length MmpL3 purified from *M. smegmatis* (*SI Appendix*, Figs. S5 and S8). To elucidate if these lipids are potential substrates for MmpL3, we used native mass spectrometry to detect the direct binding of PG or CDL to MmpL3. We observed that both PG and CDL specifically interact with MmpL3 (*SI Appendix*, Fig. S9), indicating that these lipids are potential substrates for the MmpL3 transporter.

Discussion

MmpL3 functions as a monomer

The translocation of TMM across the cytoplasmic membrane is an essential step in mycobacterial cell wall biogenesis. This process is absolutely required for the biosyntheses of mAGPs and TDMs that subsequently allow the cell envelope of Mtb to form a very rigid and extremely impermeable layer to a wide range of antimicrobial compounds. However, exactly how TMMs are shuttled across the membrane for cell wall biosynthesis remains largely unknown. This void significantly hampers progress in understanding the molecular mechanisms of Mtb cell wall biogenesis.

Our crystallographic and native MS data strongly indicate that the *M. smegmatis* MmpL3 transporter functions as a monomer. Its oligomerization state and three-dimensional structure are very distinct from all other known structures belonging to the RND superfamily of membrane proteins. For example, the well-studied hydrophobic and amphiphilic efflux (HAE) subfamily of RND transporters, including AcrB (24, 25), MexB (26), MtrD (29) and CmeB (30), exist as trimers within the cytoplasmic membrane of Gram-negative bacteria. They have large, periplasmic domains that confer substrate specificity and interact with periplasmic adaptors and outer membrane channels to form tripartite efflux complexes. Likewise, crystal structures of the heavy metal efflux (HME) subfamily of RND transporters, such as CusA (27, 28, 34) and ZneA (35), also suggest a trimeric oligomerization, where each protomer was crystallized around a crystallographic pseudo threefold axis. Recently, another subfamily of the RND-superfamily transporters, termed hopanoid biosynthesis-associated RND (HpnN), from *Burkholderia multivorans* was crystallized (31). Surprisingly, this HpnN transporter presents as a dimer, which shuttles hopanoid lipids from the cytoplasmic membrane to outer membrane of *B. multivorans* for cell wall remodeling. In the present study, we observed that MmpL3, which is affiliated with

the MmpL subfamily of RND-superfamily transporters, operates as a monomer. Based on the MmpL3 structure, it is likely that members of the MmpL-family transporters may constitute as a monomer to translocate substrates. Our data indeed poses a possibility that different subfamilies of these RND membrane proteins may assemble into various oligomerization states to function.

Proposed mechanism for TMM translocation

We propose that the monomeric MmpL3 transporter takes up TMM from the outer leaflet of the cytoplasmic membrane via a channel constituted by the transporter. A cleft surrounded by TMs 7-10 forms the entrance of this channel, which spans the outer leaflet of the cytoplasmic membrane and up to the central cavity of the periplasmic domain between PD1 and PD2. The TMM molecule could be shuttled through the channel to reach the periplasmic domain of the protein. This bound lipid could then be released to the inner leaflet of the outer membrane, where biosyntheses of mAGP and TDM take place (Fig. 6). Proton transfer via the proton-relay network mediates the energy needed for substrate translocation. The processes of proton import and substrate export may be coupled within the transport cycle.

MmpL3 specifically binds a variety of lipids

The unique composition of lipids that make up the cell envelope is a predominant feature of Mtb and other mycobacteria. These lipids constitute ~40% of the dry cell mass (36), with the organization of these lipids between the inner and outer membranes significantly contributing to its impermeability to biocides and antimicrobials. Besides being integral components of the mycobacterial cell wall, Mtb is capable of releasing PE, PI and CDL into the infected organism (37, 38). These mycobacterial lipids may play a role in the infectivity and persistence of

mycobacteria within the host (38). Therefore, mycobacterial lipid biogenesis and cell wall biosynthesis are regarded as important targets for many first-line antimycobacterial agents (39). Using native MS we observed that MmpL3 binds to a variety of lipids, including TMM, PE, PG, PI, DAG and CDL, indicating that this transporter is probably a promiscuous lipid binding protein.

Our data also suggest that MmpL3 likely contributes to mycobacterial cell envelope biogenesis by active translocation of several important lipid components to the cell wall. The ability of several potential therapeutics to block the binding and transport of these lipids via MmpL3 implicates this membrane protein as an important TB drug target. Our study is a significant advance in the field as it defines the structure of MmpL3 in the context of lipid substrates. While our paper was under review, Zhang and colleagues reported the T4 lysozyme-MmpL3 fusion protein structure as well as structures of this protein with several presumed inhibitors of MmpL3 (40) (*SI Appendix*, Fig. S10). Combined, our studies will ultimately inform a new era in structure-guided drug design for the critical cell envelope transporter MmpL3 and other virulence-associated MmpL proteins.

Methods

All protein crystals were grown at 25°C using vapor diffusion. A 2 µl protein solution containing 20 mg/ml MmpL3 or MmpL3₇₇₃ in 20 mM Na-HEPES (pH 7.5) and 0.05% (w/v) DDM was mixed with 2 µl of reservoir solution containing 25% PEG 400, 0.1 M sodium acetate (pH 5.4) and 0.05 M magnesium acetate. The resultant mixture was equilibrated against 500 µl of the reservoir solution. The crystallization conditions for SeMet-MmpL3₇₇₃ were the same as those for MmpL3₇₇₃. Cryoprotection of these crystals was achieved by raising the PEG 400

concentration to 30%. A full description of methods for protein expression and purification, X-ray data collection and analysis, native mass spectrometry, lipidomic analysis, and lipid binding are described in *SI Appendix, SI Materials and Methods*.

Acknowledgements. This work was supported by NIH Grants R01AI123148 (G.E.P. and E.W.Y.), R01AI145069 (E.W.Y.), and an MRC Grant MR/N020413/1 (C.V.R.).

Supplementary information

Supplementary Materials

Supplementary Figs. 1 to 10

Supplementary Table 1

References

1. Wallis RS, *et al.* (2016) Tuberculosis--advances in development of new drugs, treatment regimens, host-directed therapies, and biomarkers. *Lancet Infect Dis* 16(4):e34-46.
2. WHO (2017) Global tuberculosis report 2017.
3. Brennan PJ & Nikaido H (1995) The envelope of mycobacteria. *Annu Rev Biochem* 64:29-63.
4. Camacho LR, *et al.* (2001) Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier. *J Biol Chem* 276(23):19845-19854.
5. Gilmore SA, *et al.* (2012) Sulfolipid-1 biosynthesis restricts *Mycobacterium tuberculosis* growth in human macrophages. *ACS Chem Biol* 7(5):863-870.
6. Perez RL, *et al.* (2000) Cytokine message and protein expression during lung granuloma formation and resolution induced by the mycobacterial cord factor trehalose-6,6'-dimycolate. *J Interferon Cytokine Res* 20(9):795-804.
7. Bekierkunst A, *et al.* (1969) Granuloma formation induced in mice by chemically defined mycobacterial fractions. *J Bacteriol* 100(1):95-102.
8. Indrigo J, Hunter RL, Jr., & Actor JK (2003) Cord factor trehalose 6,6'-dimycolate (TDM) mediates trafficking events during mycobacterial infection of murine macrophages. *Microbiology* 149(Pt 8):2049-2059.
9. Geisel RE, Sakamoto K, Russell DG, & Rhoades ER (2005) In vivo activity of released cell wall lipids of *Mycobacterium bovis* bacillus Calmette-Guerin is due principally to trehalose mycolates. *J Immunol* 174(8):5007-5015.
10. Takayama K, Wang C, & Besra GS (2005) Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin Microbiol Rev* 18(1):81-101.
11. Banerjee A, *et al.* (1994) inhA, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science* 263(5144):227-230.
12. Cole ST, *et al.* (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393(6685):537-544.
13. Tseng TT, *et al.* (1999) The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J Mol Microbiol Biotechnol* 1(1):107-125.
14. Lamichhane G, Tyagi S, & Bishai WR (2005) Designer arrays for defined mutant analysis to detect genes essential for survival of *Mycobacterium tuberculosis* in mouse lungs. *Infect Immun* 73(4):2533-2540.
15. Domenech P, Reed MB, & Barry CE, 3rd (2005) Contribution of the *Mycobacterium tuberculosis* MmpL protein family to virulence and drug resistance. *Infect Immun* 73(6):3492-3501.
16. Domenech P, *et al.* (2004) The role of MmpL8 in sulfatide biogenesis and virulence of *Mycobacterium tuberculosis*. *J Biol Chem* 279(20):21257-21265.
17. Cox JS, Chen B, McNeil M, & Jacobs WR, Jr. (1999) Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* 402(6757):79-83.
18. Grzegorzewicz AE, *et al.* (2012) Inhibition of mycolic acid transport across the *Mycobacterium tuberculosis* plasma membrane. *Nat Chem Biol* 8(4):334-341.

19. Varela C, *et al.* (2012) MmpL genes are associated with mycolic acid metabolism in mycobacteria and corynebacteria. *Chem Biol* 19(4):498-506.
20. Xu Z, Meshcheryakov VA, Poce G, & Chng SS (2017) MmpL3 is the flippase for mycolic acids in mycobacteria. *Proc Natl Acad Sci U S A* 114(30):7993-7998.
21. Belisle JT, *et al.* (1997) Role of the major antigen of Mycobacterium tuberculosis in cell wall biogenesis. *Science* 276(5317):1420-1422.
22. Bhatt A, Kremer L, Dai AZ, Sacchettini JC, & Jacobs WR, Jr. (2005) Conditional depletion of KasA, a key enzyme of mycolic acid biosynthesis, leads to mycobacterial cell lysis. *J Bacteriol* 187(22):7596-7606.
23. Belardinelli JM, *et al.* (2016) Structure-Function Profile of MmpL3, the Essential Mycolic Acid Transporter from Mycobacterium tuberculosis. *ACS Infect Dis* 2(10):702-713.
24. Murakami S, Nakashima R, Yamashita E, & Yamaguchi A (2002) Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* 419(6907):587-593.
25. Murakami S, Nakashima R, Yamashita E, Matsumoto T, & Yamaguchi A (2006) Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* 443(7108):173-179.
26. Sennhauser G, Bukowska MA, Briand C, & Grutter MG (2009) Crystal structure of the multidrug exporter MexB from Pseudomonas aeruginosa. *J Mol Biol* 389(1):134-145.
27. Long F, *et al.* (2010) Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport. *Nature* 467(7314):484-488.
28. Su CC, *et al.* (2011) Crystal structure of the CusBA heavy-metal efflux complex of Escherichia coli. *Nature* 470(7335):558-562.
29. Bolla JR, *et al.* (2014) Crystal structure of the Neisseria gonorrhoeae MtrD inner membrane multidrug efflux pump. *PLoS One* 9(6):e97903.
30. Su CC, *et al.* (2017) Structures and transport dynamics of a Campylobacter jejuni multidrug efflux pump. *Nat Commun* 8(1):171.
31. Kumar N, *et al.* (2017) Crystal structures of the Burkholderia multivorans hopanoid transporter HpnN. *Proc Natl Acad Sci U S A* 114(25):6557-6562.
32. Chim N, *et al.* (2015) The Structure and Interactions of Periplasmic Domains of Crucial MmpL Membrane Proteins from Mycobacterium tuberculosis. *Chem Biol* 22(8):1098-1107.
33. Li W, *et al.* (2014) Novel insights into the mechanism of inhibition of MmpL3, a target of multiple pharmacophores in Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 58(11):6413-6423.
34. Su CC, *et al.* (2012) Charged amino acids (R83, E567, D617, E625, R669, and K678) of CusA are required for metal ion transport in the Cus efflux system. *J Mol Biol* 422(3):429-441.
35. Pak JE, *et al.* (2013) Structures of intermediate transport states of ZneA, a Zn(II)/proton antiporter. *Proc Natl Acad Sci U S A* 110(46):18484-18489.
36. Jackson M (2014) The mycobacterial cell envelope-lipids. *Cold Spring Harb Perspect Med* 4(10):a021105.
37. Wade AA & Rabson AR (1983) Binding of phosphatidylethanolamine and phosphatidylinositol to OKT8+ lymphocytes activates suppressor cell activity. *J Immunol* 130(5):2271-2276.

38. Fischer K, *et al.* (2001) Mycobacterial lysocardiolipin is exported from phagosomes upon cleavage of cardiolipin by a macrophage-derived lysosomal phospholipase A2. *J Immunol* 167(4):2187-2192.
39. Jackson M, Crick DC, & Brennan PJ (2000) Phosphatidylinositol is an essential phospholipid of mycobacteria. *J Biol Chem* 275(39):30092-30099.
40. Zhang B, *et al.* (2019) Crystal Structures of Membrane Transporter MmpL3, an Anti-TB Drug Target. *Cell* 176(3):636-648.

Legends of Figures

Fig. 1. Mass spectra of purified MmpL3 proteins. (A) Mass spectrum of full-length MmpL3 expressed in *E. coli* indicates that the protein exists as a monomer. Calculated and observed masses, including the C-terminal 6×His tag, are 110,222 Da and 110,246 Da. (B) Mass spectrum of full-length MmpL3 expressed in *M. smegmatis*. In addition to the monomeric protein in solution, the presence of degraded protein bands indicates that the purified protein is unstable. The two observed masses are 110,371 Da and 91,990 Da. (C) Mass spectrum of a five-day old sample of full-length MmpL3 expressed and purified from *E. coli*. The spectrum depicts three major degraded species in the solution sample. The observed masses of these species are 87,998 Da, 84,842 Da and 83,359 Da, corresponding to residues 1-806, 1-776 and 1-763 of the protein, respectively. (D) Mass spectrum of the MmpL3₇₇₃ protein. The spectrum indicates that this protein exists as a monomer in solution. Observed and calculated masses in this case are 85,950 Da and 85,925 Da. The mass observed from the adduct peaks is 86,660 Da, which corresponds to the MmpL3₇₇₃-PE complex.

Fig. 2. Structure of the *M. smegmatis* MmpL3 transporter. (A) Secondary structural topology of the MmpL3₇₇₃ monomer. The topology was constructed based on the crystal structure of MmpL3₇₇₃. The transmembrane helices (TMs) are colored blue. The periplasmic loops 1 and 2 are in cyan and green, respectively. The cytoplasmic domain (CD) of MmpL3 is colored red. (B) Ribbon diagram of a monomer of MmpL3₇₇₃ viewed in the membrane plane. The TMs, periplasmic loops 1 and 2, and CD are colored slate, cyan, green and red, respectively. The bound DDM is depicted as yellow spheres, whereas the bound PE is depicted as pink spheres. The periplasmic loops 1 and 2 cross over each other to form the periplasmic domains 1 and 2

(PD1 and PD2). (C) The MmpL3₇₇₃ monomer forms a channel spanning the outer leaflet of the inner membrane and up to the periplasmic domain. The orientation of this MmpL3 molecule has been rotated by 110° (as shown in the figure) when compared with the orientation of (B). The channel (colored gray) was calculated using the program CAVER (<http://loschmidt.chemi.muni.cz/caver>). The TMs, periplasmic loops 1 and 2, and CD are colored the same as in (B). The bound DDM is depicted as yellow sticks, whereas the bound PE is depicted as pink sticks.

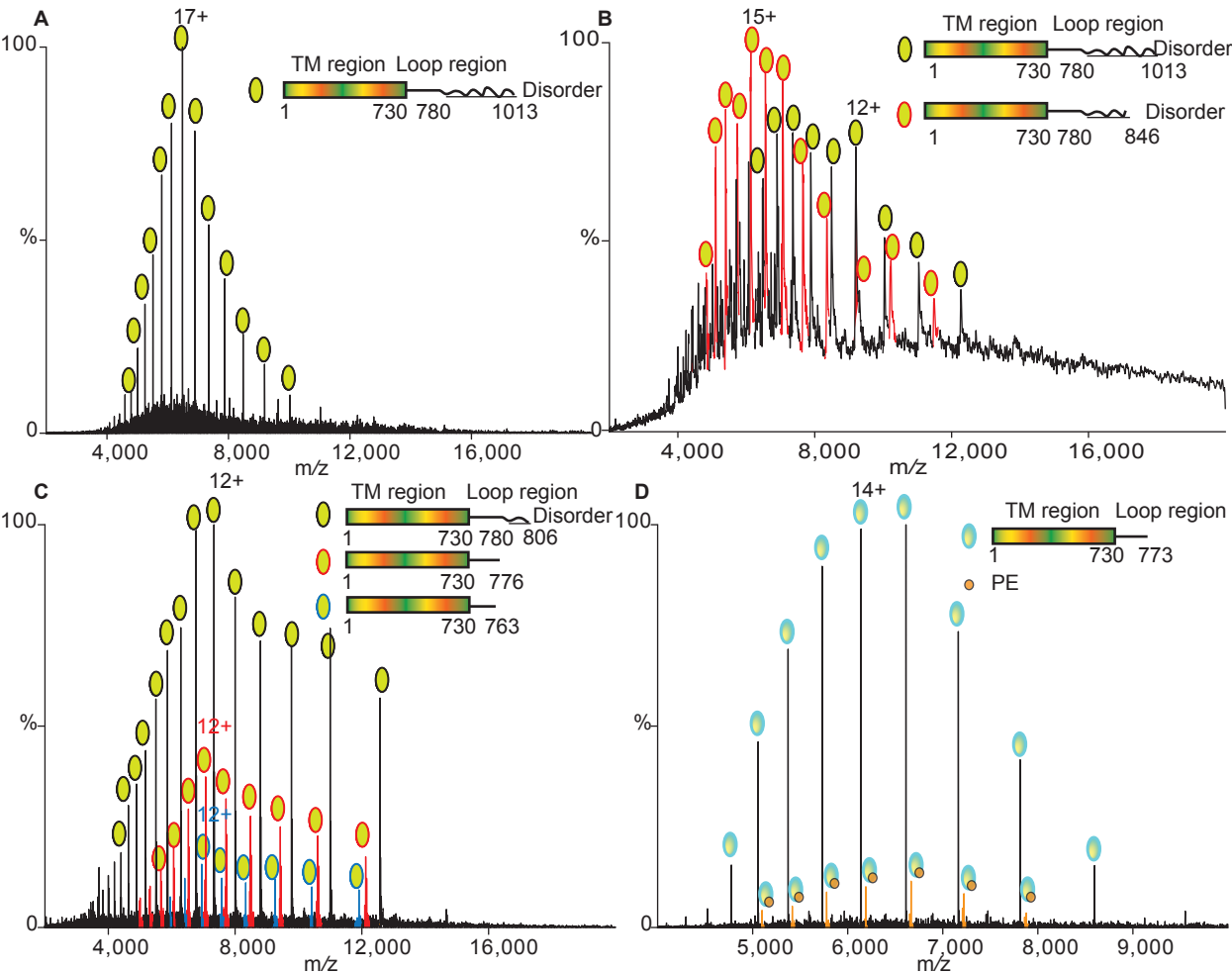
Fig. 3. The DDM and PE binding sites of the MmpL3-PE complex. (A) The F_o - F_c electron density map of bound DDM in MmpL3. The bound DDM is shown as a stick model (yellow, carbon; red, oxygen). The F_o - F_c map is contoured at 3σ. Residues involved in DDM binding are shown as green sticks. The secondary structural elements of MmpL3₇₇₃ are colored light brown. (B) The F_o - F_c electron density map of bound PE in MmpL3. The bound PE is shown as a stick model (magenta, carbon; red, oxygen; blue, nitrogen). The F_o - F_c map is contoured at 3σ. Residues involved in PE binding are shown as green sticks. The secondary structural elements of MmpL3₇₇₃ are colored light brown.

Fig. 4. Determination of dissociation constants for the binding of PE and TMM to MmpL3₇₇₃. (A) Mass spectra recorded for solutions of MmpL3₇₇₃ with increasing concentrations of PE. At 5 μM PE a charge state series is observed (light orange), corresponding to bound PE, which increases in intensity as the PE concentration is increased to 80 μM. A second PE binding peak (gray) emerges at concentrations above 20 μM. (B) Plot of relative fractional intensity of lipid binding peaks over the total peak intensity versus PE concentration (see Experimental Methods),

yielding a curve for the first binding event and linear-like fit for the second, consistent with non-specific PE binding. Each data point and standard deviation are calculated from the average of five observed charge states in three independent experiments. Error bars represent standard deviations ($n = 3$). (C) Mass spectra recorded for solutions of MmpL3₇₇₃ with increasing concentrations of TMM. At 2.5 μ M TMM a charge state series is observed (orange), corresponding to bound TMM, which increases in intensity as the TMM concentration is increased to 40 μ M. A second TMM binding peak (purple) emerges at concentrations above 20 μ M. (D) Plot of the relative fractional intensity of lipid binding peaks over total peak intensity versus TMM concentration (see Experimental Methods), yielding a curve for the first binding event and linear-like fit for the second, consistent with non-specific TMM binding. Each data point and standard deviation are calculated from the average of five observed charge states in three independent experiments. Error bars represent standard deviations ($n = 3$).

Fig. 5. Analysis of different components of purified *M. smegmatis* glycolipids and their binding to MmpL3₇₇₃. Mass spectrum of MmpL3₇₇₃ with purified *M. smegmatis* glycolipids shows a preferential binding to TMM (orange charge state series) but not to TDM. Purified lipid fraction has both TMM (red inlet) and TDM (blue inlet) at the lower m/z region.

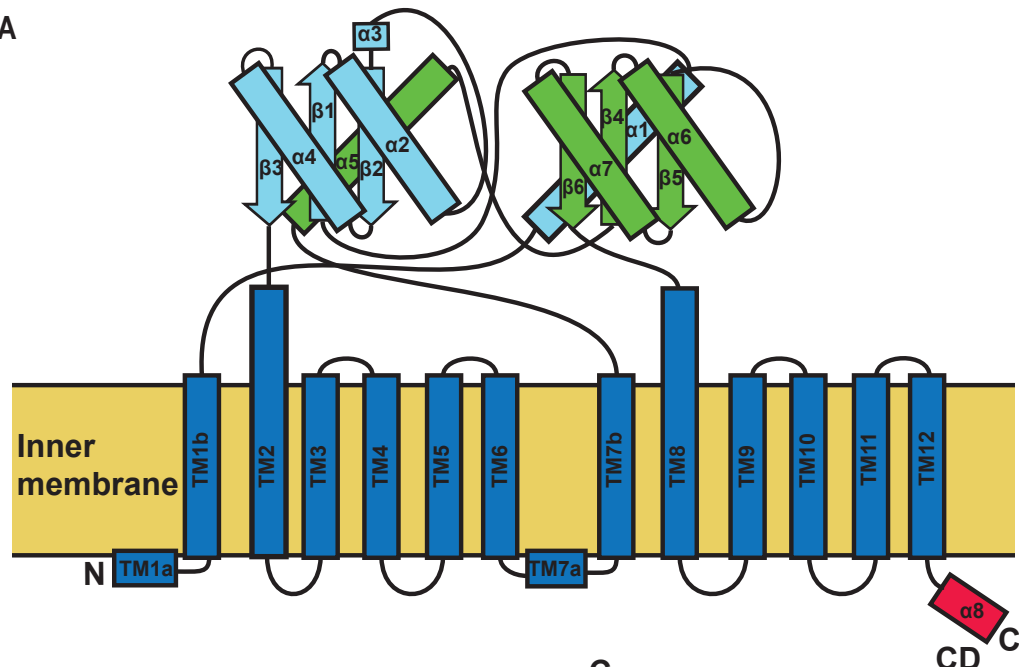
Fig. 6. Proposed mechanism for TMM translocation via MmpL3. This schematic diagram indicates that the MmpL3 transporter is capable of picking up a TMM molecule from the outer leaflet of the cytoplasmic membrane. This TMM molecule will pass through the channel formed by MmpL3 and arrive at the periplasmic lipid binding site. The TMM moiety will then be exported to the inner leaflet of the outer membrane for the biosynthesis of TDMs and mAGPs.



PD1

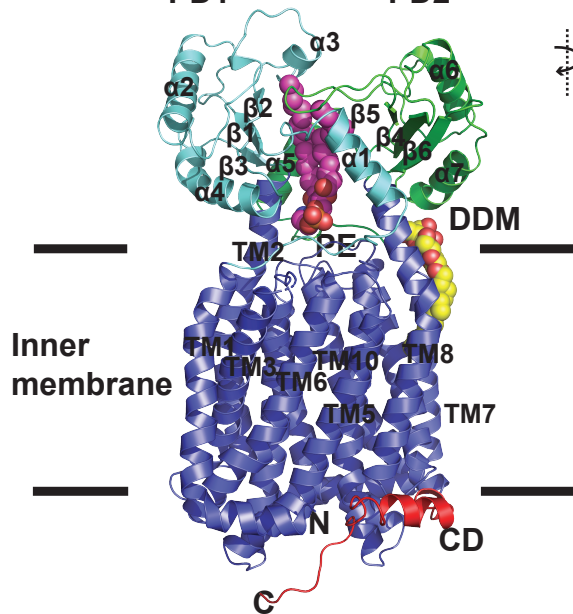
PD2

A



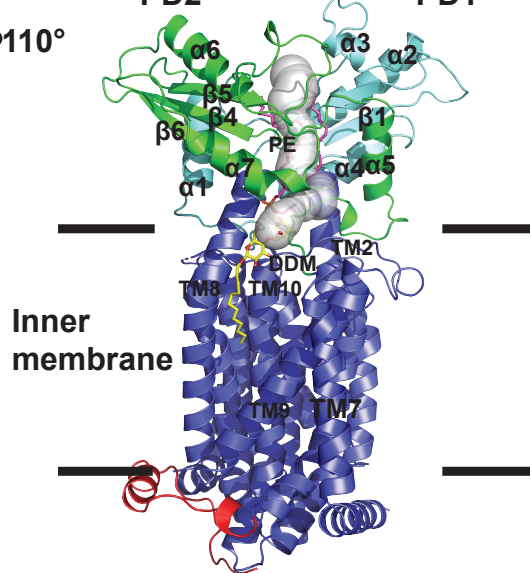
B

PD1 PD2

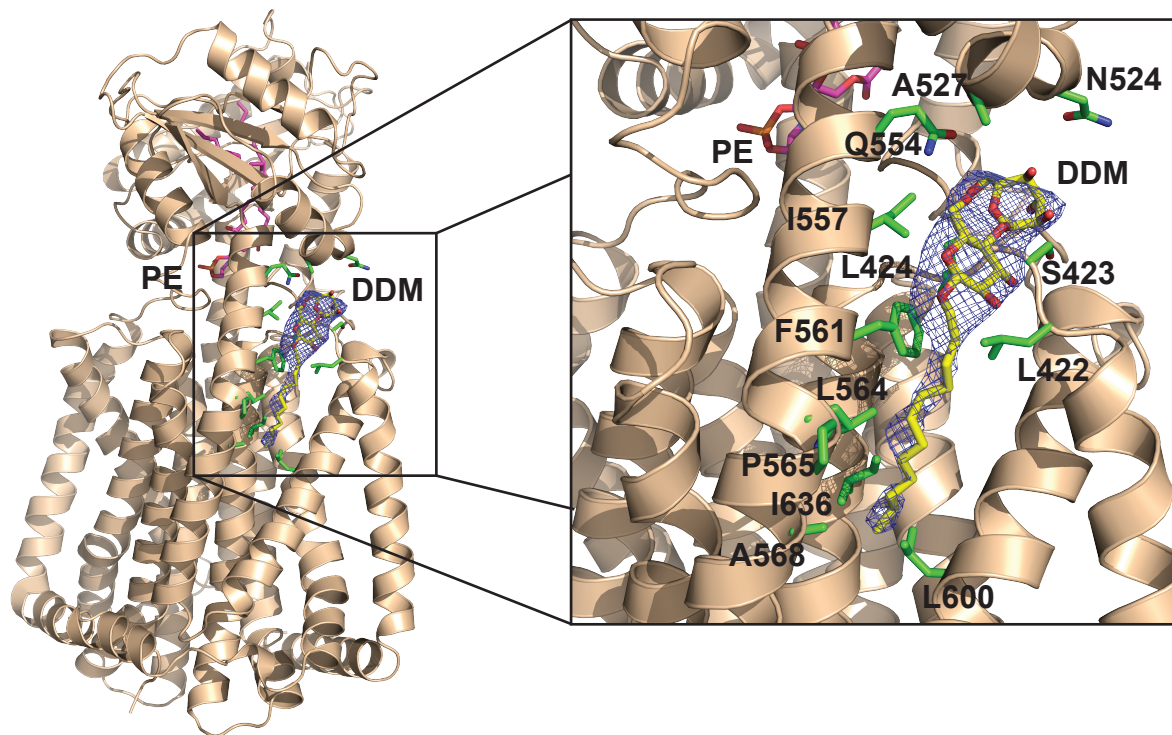


C

PD2 PD1



A



B

