nature portfolio

Peer Review File

Structures of the Mycobacterium tuberculosis efflux pump EfpA reveal the mechanisms of transport and inhibition

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Wang et al. report the cryo-EM structures of essential efflux pumpA with and without the inhibitor molecule BRD-8000.3. All structures are in outward-open conformation. The authors observe strong non-protein density at three cavities and model lipid molecules (two phosphatidylethanolamine and one cardiolipin per monomer) into these densities. They also observe that BRD-8000.3 displaces one of the lipid molecules. Although no structural data are available, the authors propose that another inhibitor molecule (BRD-9327) binds to another lipid binding pocket based on the mutagenesis data and molecular simulation analysis. Comparing the models predicted through AlphaFold, the authors propose a mechanism for how the transport is achieved. They also propose a novel physiological role for these transporters in lipid transport based on the positioning of the lipid molecules within the core of the proteins and similarities to other lipid-transporting proteins. Overall, the manuscript presents interesting structural data that would be of interest to the field. However, the conclusions derived from these structures appear very speculative, and the reviewer is concerned if the data are sufficient to support these.

1) A cardiolipin molecule is modeled within the interface between the two protomers arranged in opposite directions relative to the membrane plane. The density appears to be located at the symmetry axis. Based on the analysis of the figures (and not having access to the maps), it seems very plausible that this density is of two phospholipid molecules interacting with each protomer. The connectivity between the densities could result from the two-fold symmetric arrangement. The source of cardiolipin is also not clear. The primary cardiolipin content would be in mitochondrial membranes. Do the authors think cardiolipin molecules are solubilized during detergent extraction and bind to this pocket, or the pocket is already occupied prior to the solubilization? Does the author observe any density at the same binding pocket in monomeric proteins (purified with DDM)? The lipidomic analysis seems very indirect validation. G377D and G502D mutations may result in more drastic changes at the interface affecting dimerization rather than solely affecting lipid binding. Moreover, the SEC profile for the G502D mutant suggests more of an equilibrium between the monomeric and dimeric stages. The identity of the lipid at this site does not seem critical for the conclusions of the manuscript, but the presence of the ambiguity should be more clearly explained.

2) In line 274, the authors state, "In this model, when a proton is released from the central pocket to the intracellular milieu, it triggers the transition of EfpA from an inward-open conformation to an outward-open conformation, expanding lipid-binding site A and facilitating the binding of substrate lipids from the inner-leaf of the membrane". It is unclear how authors can reach a conclusion where a proton's release is coupled to the conformational changes with the presented data.

3) The proposed role of lipid transport is very interesting but speculative without any experimental validation.

4) There is also no validation for the proposed transport mechanism based on the structures predicted through AlphaFold.

Reviewer #2 (Remarks to the Author):

The work of Wang et al describes the cryo-EM structure and functional relationships of the Mycobacterial EfpA major facilitator superfamily transporter. Based on homology and the presence of endogenous lipids, the conclusion is reached that EfpA probably functions to flop glycerophospholipids from the inner leaflet to the outer leaflet of the cytoplasmic membrane. The expected alternating access mechanism is reinforced by a complementary AlphFold model. Additional structures with bound inhibitors show displacement of lipids at the lateral gate in the inner leaflet, consistent with the proposed model. The structural analysis is thorough, but the analysis of function is largely preliminary. Nevertheless, the work lays the groundwork for future functional studies.

My main criticism of the work lies in the fact that the endogenous bound lipids are derived from the expression system, which is based on human embryonic kidney cells. As such, the bound lipids are distinctly different from those found in the native mycobacterial membranes. How are mycobacterial cytoplasmic membrane glycerophospholipids similar or different to those found in HEK293 cells? What lipids might we expect to be bound to EfpA in its native membrane environment?

I would also like to see some commentary on the general significance of cardiolipin in mediating intersubunit interfaces in the structures of membrane proteins like the mitochondrial ADP-ATP exchanger, for example. Is cardiolipin required for EfpA activity? Would other lipids be able to replace cardiolipin at the EfpA subunit interface? Are the homologous transporters monomers or dimers? It is still unclear to me whether EfpA is functional as a dimer or as a monomer.

Several additional minor concerns are listed below:

1) On line 127, please spell out the names of the detergents to define the abbreviations: Lauryl Maltose Neopentyl Glycol (LMNG) and Cholesteryl Hemisuccinate (CHS), dodecyl-β-D-maltoside (DDM) etc.

2) On line 169-170, the statement "Among these residues, Y97, which is solvent exposed and can undergo protonation under neutral and alkaline conditions while accepting protons under acidic

conditions..." is awkward. The pKa of Tyr is typically ~ 10. Your statement is worded to imply that it can be protonated under all pH conditions???

3) In some places, you refer to the proposed mechanism as being "Z"-shaped and in other places as being like a "staircase". Pick one and stick with it to avoid confusing your readers.

4) In the legend to Figure 1, panel b is a "cartoon model" not a "carton model". The clause "locates on the cytoplasm surface of membrane." should read "is located on the cytoplasmic surface of the membrane."

5) In the legend to Figure 2 "molecular of Phosphatidylethanolamine" should read "molecules of phosphatidylethanolamine".

6) In extended data Figure 1 panel D, the abbreviations are unintelligible without being defined in the legend. Also, in Panel D, what do the ovals represent with respect to the permeabilization procedure discussed?

7) In extended data Figure 5 panel A, "Monoer" should read "Monomer" in 4 places. In the legend line 687, "interacted" should be "interacting".

8) In extended data Figure 10 line 727 "homologous" should read "homologues."

Reviewer #3 (Remarks to the Author):

The study by Wang and colleagues characterized with help of cryo-EM and molecular dynamics simulations the efflux pump EfpA of Mycobacterium tuberculosis that could be a potential drug target. In this process, they identified three lipids that are bound to the transporter and that novel inhibitors can block the lipid binding. Within that publication, cells expressing dimeric and monomeric MtEfpA underwent a lipid analysis by using a shotgun lipidomics approach. This review is based on my expertise in the lipidomics field and thus does not critically evaluate the other parts. In my opinion, there are some more points that should be addressed:

1. Data sharing: In my opinion, it is important to share the lipidomics data in a repository for other colleagues that work in the same field as well as for reviewers to check your data.

2. Lipidomics method description:

a. In general, the lipidomics method is very short. Some parameters are missing, e.g. mass range, acquisition rate, collision energy for fragmentation, how many injections per polarity etc.

b. Information on the lipid standard used is lacking. The source file shows that a SPLASH mix was used, please add to method. In addition, a quantification would have been better than the reporting of percentages.

c. Analysis of MS data: Which parameters and mfql files were used for LipidXplorer? Was the data normalized? in general, I would have specified the data evaluation part a bit more in detail.

d. For what specific reason is XCalibur named in the method? Did you use Freestyle or the Quan/Qual Browser to extract some data?

3. Extended Data Fig 1e:

a. I am missing information on replicates, there are no error bars shown + then a statistical test would underpin the analysis.

b. In the bar chart, PC is missing. It is however shown in the supplementary excel file with all lipids identified.

Minor comments:

• Extended Data Fig. 1e: y-axis description, percentage or [%]; Scaling of the y-axis, the first lipid classes are not noticeable in the bar chart (even if they might show no differences between the two cell types)

• Line 381: FWHM at 200 m/z

Reviewer #4 (Remarks to the Author):

Brief summary:

In this manuscript, the authors elucidate a novel and therapeutically interesting protein, EfpA from Mycobacterium tuberculosis, identifying and characterizing an unexpected putative lipid-flippase mechanism. This work utilized a combination of CryoEM and computational structure prediction to characterize Mycobacterium tuberculosis (Mtb) efflux pump (EfpA) – a critical component in multi-drug resistant Mtb. In addition to elucidation of the inward, outward, and inhibited conformations, the authors discovered an unexpected lipid-transport path, reminiscent of the MFS-lysophospholipid transporters. The discovery of the lipid-transporting moieties helps to shine a light on the recently observed role of EfpA in cell-growth, and furthermore provide a mechanistic framework for which novel therapeutics could be designed to target specific functions of EfpA.

Overall the structural biology experiments appear to be sound, the manuscript reads well and is interesting, though there are critical aspects of the text/figures which require addressing.

Major points:

1. The authors acknowledge the difficulty of identifying the endogenous substrate to MtbEfpA, however what was not mentioned was that lipid-binding-site-B coincides with the putative substrate-binding-site. While the authors did mention that the mutual interplay between the lipid-transport and substrate-transport is currently unknown, it would undoubtedly alter the "comprehensive" lipid-transport model.

2. Regarding the lipid-transport model proposed, there are two complete Inward-Open to Outward-Open transitions, i.e., cycles, which are needed to transport a single lipid. What evidence supports this ?

3. The authors data suggest that the anti-parallel MtbEfpA-dimer conformation may be an artifact of sample-solubilization (i.e., utilizing allophycocyanin), and demonstrate that the overall monomer structure is unchanged when compared to MsEfpA monomer. Though I fail to see an explanation as to why MsEfpA doesn't form a dimer. Could there be an important structural feature in Mtb which could impart this strange dimerization?

4. In line 171, Y97 is proposed to be the proton acceptor-donor, an exchange which is critical for the proposed lipid-transport model. Tyrosine is not typically considered protonatable as the side-chain pKa at STP is >> 7 (not unlike Arg & Lys). Given that lipid-binding site B (and substrate binding pocket?) is perpetually solvated, it would seem as though it would seldom experience a local pH high enough to deprotonate. Is there any evidence that suggest Y97 could exchange protons ?

Minor points:

1. There was no mention of the specific parameterization of BRD8000.3 and BRD9327. Parameterization is a non-trivial process and could have substantial consequences to the observed interactions between the inhibitors and the protein.

2. It is unclear to me whether the MD simulation were performed for a monomeric or dimeric protein ? In any case, I believe it would be crucial to show a comparative (i.e., monomeric vs

dimeric) analysis of lipid-protein interactions. Furthermore, a simulation analysis concerning the intersubunit interactions that stabilize the dimer needs to be included (i.e. regarding the contribution of protein-protein and protein-lipid interactions in stabilizing the dimer)

3. Extended Figure 5a should say “monomer” as opposed to “monoer”

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Wang et al. report the cryo-EM structures of essential efflux pumpA with and without the inhibitor molecule BRD-8000.3. All structures are in outward-open conformation. The authors observe strong non-protein density at three cavities and model lipid molecules (two phosphatidylethanolamine and one cardiolipin per monomer) into these densities. They also observe that BRD-8000.3 displaces one of the lipid molecules. Although no structural data are available, the authors propose that another inhibitor molecule (BRD-9327) binds to another lipid binding pocket based on the mutagenesis data and molecular simulation analysis. Comparing the models predicted through AlphaFold, the authors propose a mechanism for how the transport is achieved. They also propose a novel physiological role for these transporters in lipid transport based on the positioning of the lipid molecules within the core of the proteins and similarities to other lipid-transporting proteins. Overall, the manuscript presents interesting structural data that would be of interest to the field. However, the conclusions derived from these structures appear very speculative, and the reviewer is concerned if the data are sufficient to support these.

We thank the reviewer for summarizing the key findings of our work. We will address the comments and concerns point-by-point below.

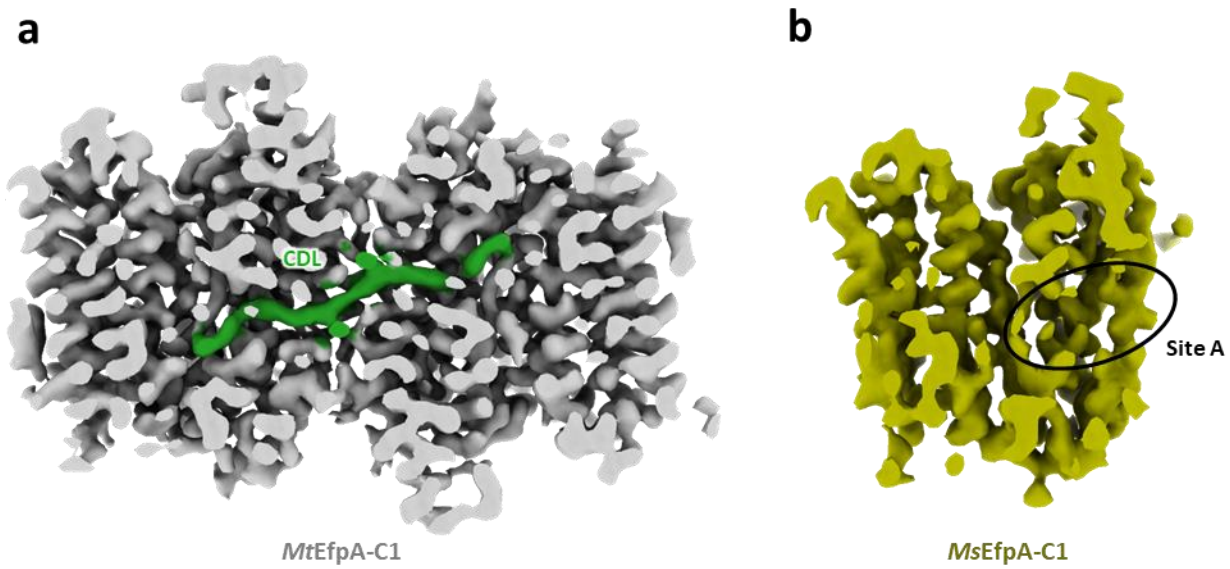
1) A cardiolipin molecule is modeled within the interface between the two protomers arranged in opposite directions relative to the membrane plane. The density appears to be located at the symmetry axis. Based on the analysis of the figures (and not having access to the maps), it seems very plausible that this density is of two phospholipid molecules interacting with each protomer. The connectivity between the densities could result from the two-fold symmetric arrangement. The source of cardiolipin is also not clear. The primary cardiolipin content would be in mitochondrial membranes. Do the authors think cardiolipin molecules are solubilized during detergent extraction and bind to this pocket, or the pocket is already occupied prior to the solubilization? Does the author observe any density at the same binding pocket in monomeric proteins (purified with DDM)? The lipidomic analysis seems very indirect validation. G377D and G502D mutations may result in more drastic changes at the interface affecting dimerization rather than solely affecting lipid binding. Moreover, the SEC profile for the G502D mutant suggests more of an equilibrium between the monomeric and dimeric stages. The identity of the lipid at this site does not seem critical for the conclusions of the manuscript, but the presence of the ambiguity should be more clearly explained.

To test if the connectivity between the lipid densities from two EfpA protomers is an artifact due to applying C2 symmetry, we reprocessed the dataset without symmetry throughout global refinement and local refinement. The resulting lipid densities displayed similar connectivity at the interface (Comments Fig. 1a). In addition, there is no observed density in the same pocket in monomeric EfpA-DDMCHS (Comments Fig. 1b). Considering the upside-down EfpA dimer likely formed after membrane solubilization, our observations suggest that the lipid at the dimer interface is not tightly bound to EfpA initially in the membrane.

It is an interesting question whether cardiolipin was bound before or after EfpA solubilization from the membrane. While there is no definitive answer based on available data, we favor a

model in which cardiolipin binds to EfpA after membrane solubilization and stabilizes EfpA dimer by connecting two EfpA proteins. In consistency with this hypothesis, the upside-down EfpA dimer is unlikely the original conformation in the membrane and probably forms during solubilization and purification in detergents (lines 112-118), and EfpA dimer appears more stable and homogeneous in our biochemical and structural analyses. We have revised the manuscript to improve clarity (lines 143-145).

Thanks for pointing this out. The G377D and G502D mutations are located close to the negatively charged head group of cardiolipin, the entrance of the lateral gate, as well as the dimeric interface. This suggests they may affect lipid binding and directly contribute to disrupting dimerization. The manuscript has been revised to mention both aspects. (lines 160-162).



Comments Fig. 1 Density at lipid binding site A of *Mt*EfpA dimer (a) and *Ms*EfpA monomer (b), processed with no symmetry. a, Cross-sectional view of the density map of *MtEfpA* dimer showing lipid-binding site A. The density within 3 Å of the cardiolipin (CDL) model is colored green. **b**, Cross-sectional view of the density map of *Ms*EfpA showing the empty lipid-binding site A.

2) In line 274, the authors state, "In this model, when a proton is released from the central pocket to the intracellular milieu, it triggers the transition of EfpA from an inward-open conformation to an outward-open conformation, expanding lipid-binding site A and facilitating the binding of substrate lipids from the inner-leaf of the membrane". It is unclear how authors can reach a conclusion where a proton's release is coupled to the conformational changes with the presented data.

Thanks for pointing this out. With available data, we cannot draw conclusion about the coupling between proton release and conformational changes of EfpA. Figure 5 and related discussion were to summarize our key findings and relate those to the current knowledge in the field. We have carefully revised this section to separate the findings directly supported by our data and the current hypotheses/models that require future studies (lines 281-283).

3) The proposed role of lipid transport is very interesting but speculative without any experimental validation.

Regarding the potential role of lipid transport by EfpA, our work has made several interesting findings: 1) a lipid pathway within EfpA spanning from the inner leaflet to outer leaflet, 2) inhibitor occupying lipid-binding site, and 3) structures and lipid-binding sites highly similar to well established lipid transporter from the same family. Together, these new findings support a model in which EfpA functions as a lipid transporter to flip lipids between membrane leaflets. In our opinion, this is an important contribution to understanding the function and mechanism of EfpA. However, the endogenous substrate and exact transport cycle of EfpA require future studies. We have made the necessary revision in the text to clarify this (lines 152-155, 242-243).

4) There is also no validation for the proposed transport mechanism based on the structures predicted through AlphaFold.

Our proposed conformational transition model for EfpA transport is based on a combination of cryo-EM analysis, AlphaFold prediction, and current knowledge in the field. Cryo-EM structures of EfpA represent outward-facing conformation, and AlphaFold-predicated structure of EfpA without bound substrates is in inward-facing conformation, suggesting that a conformational transition is correlated with substrate binding and transport. This is also consistent with a previously reported related transporter MFSD2A, in which multiple conformations are present (reference PMID: 34135507, 34349262 and 37156797).

We agree with the reviewer that full experimental validation, which is beyond the scope of current study, is essential to prove this model. In fact, we've tried various ways, including supplementing with compounds during cell culture or purifying with different pH buffers, to capture the inward conformation of EfpA in our cryo-EM studies, as well as performed molecular dynamic simulation between the outward- and inward-facing conformations. Unfortunately, these attempts did not produce clear results.

Reviewer #2 (Remarks to the Author):

The work of Wang et al describes the cryo-EM structure and functional relationships of the Mycobacterial EfpA major facilitator superfamily transporter. Based on homology and the presence of endogenous lipids, the conclusion is reached that EfpA probably functions to flop glycerophospholipids from the inner leaflet to the outer leaflet of the cytoplasmic membrane. The expected alternating access mechanism is reinforced by a complementary AlphFold model. Additional structures with bound inhibitors show displacement of lipids at the lateral gate in the inner leaflet, consistent with the proposed model. The structural analysis is thorough, but the analysis of function is largely preliminary. Nevertheless, the work lays the groundwork for future functional studies.

We thank the reviewer for positive comments and acknowledging our key findings.

My main criticism of the work lies in the fact that the endogenous bound lipids are derived from the expression system, which is based on human embryonic kidney cells. As such, the bound lipids are distinctly different from those found in the native mycobacterial membranes. How are mycobacterial cytoplasmic membrane glycerophospholipids similar or different to those found in HEK293 cells? What lipids might we expect to be bound to EfpA in its native membrane environment?

In the plasma membrane of HEK293 cells, the predominant glycerophospholipids include phosphatidylcholine (PC), phosphatidylethanolamine (PE) and phosphatidylserine (PS) (reference PMID: 31052427), but, during membrane solubilization and protein purification, more lipids can be released from intracellular organelles, e.g., cardiolipin from mitochondrial membrane. In contrast, the plasma membrane of mycobacteria is characterized by the presence of phosphatidylinositol (PI), PS, PE and cardiolipin (reference PMID: 25104772). The endogenous lipids bound to EfpA in the mycobacterial membrane are currently unknown and require future investigation. We have made the necessary revision in the text to clarify this (lines 152-155, 242-243).

I would also like to see some commentary on the general significance of cardiolipin in mediating intersubunit interfaces in the structures of membrane proteins like the mitochondrial ADP-ATP exchanger, for example. Is cardiolipin required for EfpA activity? Would other lipids be able to replace cardiolipin at the EfpA subunit interface? Are the homologous transporters monomers or dimers? It is still unclear to me whether EfpA is functional as a dimer or as a monomer.

Indeed, cardiolipin has been found mediating protein interactions, particularly in many cryo-EM structures of membrane proteins. For example, respiratory supercomplex (PDB ID: 5XTE) and mycobacterial arabinosyltransferase (PDB ID: 7BVC). Cardiolipin is considered to play a structural role in stabilizing protein complex, as well as a functional role (reference PMID: 28077870, 34417182).

We are not aware of published work on the functional effects of cardiolipin on EfpA. As shown in our cryo-EM maps of BRD-bound EfpA, the inhibitor replaced the cardiolipin at the interface of EfpA dimer, thus various hydrophobic molecules can potentially occupy this site. Furthermore, we do not observe specific interactions between EfpA and the head group of cardiolipin, supporting the notion that other lipids may replace cardiolipin in binding to EfpA.

The lipid transporters MFSD2A (reference PMID: 34135507, 34349262) and LtaA (reference PMID: 35197476) in MFS family operate as monomers. EfpA may function as a monomer as well. The antiparallel dimer as observed in our cryo-EM maps is probably an artifact during sample preparation and not relevant to its physiological function.

Several additional minor concerns are listed below:

1) On line 127, please spell out the names of the detergents to define the abbreviations: Lauryl Maltose Neopentyl Glycol (LMNG) and Cholesteryl Hemisuccinate (CHS), dodecyl-β-D-maltoside (DDM) etc.

Updated according to the suggestion (line 127).

2) On line 169-170, the statement "Among these residues, Y97, which is solvent exposed and can undergo protonation under neutral and alkaline conditions while accepting protons under acidic conditions..." is awkward. The pKa of Tyr is typically ~ 10. Your statement is worded to imply that it can be protonated under all pH conditions???

Thanks for pointing this out. Tyrosine tends to be deprotonated at alkaline pH and protonated at neutral and acidic condition. Previous study reported that a tyrosine (Y55) is located adjacent to Na+-coupled residues and may stabilize Na+ in the lipid transporter MFSD2A (reference PMID: 34349262). However, the Y97 in *Mt*EfpA does not align to the Y55 in MFSD2A, therefore, we have revised the statement regarding Y97 and removed the statement of Y97's potential role in protonation/deprotonation (lines 176-178).

3) In some places, you refer to the proposed mechanism as being "Z"-shaped and in other places as being like a "staircase". Pick one and stick with it to avoid confusing your readers.
Thanks for bringing this to our attention. We've decided to use the term "staircase" to describe the pathway, and the relevant text has been updated.

4) In the legend to Figure 1, panel b is a "cartoon model" not a "carton model". The clause "locates on the cytoplasm surface of membrane." should read "is located on the cytoplasmic surface of the membrane."

Corrected.

5) In the legend to Figure 2 "molecular of Phosphatidylethanolamine" should read "molecules of phosphatidylethanolamine".

Corrected.

6) In extended data Figure 1 panel D, the abbreviations are unintelligible without being defined in the legend. Also, in Panel D, what do the ovals represent with respect to the permeabilization procedure discussed?

Thanks for the comments. We have revised the legend of Extended Data Figure 1D to enhance clarity.

7) In extended data Figure 5 panel A, "Monoer" should read "Monomer" in 4 places. In the legend line 687, "interacted" should be "interacting".

Corrected.

8) In extended data Figure 10 line 727 "homologous" should read "homologues."
Corrected (line 747).

Reviewer #3 (Remarks to the Author):

The study by Wang and colleagues characterized with help of cryo-EM and molecular dynamics simulations the efflux pump EfpA of Mycobacterium tuberculosis that could be a potential drug target. In this process, they identified three lipids that are bound to the transporter and that novel inhibitors can block the lipid binding. Within that publication, cells expressing dimeric and monomeric MtEfpA underwent a lipid analysis by using a shotgun lipidomics approach. This review is based on my expertise in the lipidomics field and thus does not critically evaluate the other parts.

We thank the reviewer for the constructive comments particularly on the aspect of lipidomics. Below we have addressed these comments point-by-point.

In my opinion, there are some more points that should be addressed:
1. Data sharing: In my opinion, it is important to share the lipidomics data in a repository for other colleagues that work in the same field as well as for reviewers to check your data.

We have now included the complete raw data, mzmL files, mfql files and search results as supplementary files.

2. Lipidomics method description:
a. In general, the lipidomics method is very short. Some parameters are missing, e.g. mass range, acquisition rate, collision energy for fragmentation, how many injections per polarity etc.

Thank you for bringing this to our attention. We have added the parameters and other details in the lipidomics method (lines 380-395) and provided the source data files for the reviewers' reference.

b. Information on the lipid standard used is lacking. The source file shows that a SPLASH mix was used, please add to method. In addition, a quantification would have been better than the reporting of percentages.

Thanks for your suggestions. We have included the lipid standard information and vendors in the methodology. Since our lipidomic analysis aims to provide a qualitative assessment of the extra density in dimeric EfpA, we believe that using % of total ion counts provides a better qualitative representation.

c. Analysis of MS data: Which parameters and mfql files were used for LipidXplorer? Was the data normalized? ◊ in general, I would have specified the data evaluation part a bit more in detail.

Analysis parameters and mfql files are now included in supplementary files. We performed two stages of data evaluation: manual curation of spectra against S/N and removing duplicate lipids in LipidXplorer results files.

d. For what specific reason is XCalibur named in the method? Did you use Freestyle or the Quan/Qual Browser to extract some data?

Neither Xcalibur, Freestyle nor Quan/Qual Browsers were used for data extractions. We have updated the method of shotgun lipidomic to reflect data extractions (lines 392-395).

3. Extended Data Fig 1e:
a. I am missing information on replicates, there are no error bars shown + then a statistical test would underpin the analysis.

Since the purpose of our lipidomic assay is to qualitatively infer which lipids contribute to the observed extra density, rather than precise quantification, we do not have replicates.

b. In the bar chart, PC is missing. It is however shown in the supplementary excel file with all lipids identified.

Thanks for pointing this out. PC was not measured in the positive mode. In the revised manuscript, we have included both positive mode and negative mode, in the latter of which PC is included (Extended Data Fig. 1e).

Minor comments:
• Extended Data Fig. 1e: y-axis description, percentage or [%]; Scaling of the y-axis, the first lipid classes are not noticeable in the bar chart (even if they might show no differences between the two cell types)

The y-axis description has been updated, and the y-axis has been rescaled in Extended Data Fig. 1e.

• Line 381: FWHM at 200 m/z

Both expressions are used in the field. The “FWHM at m/z 200” expression is more commonly used in published literature.

Reviewer #4 (Remarks to the Author):

Brief summary:

In this manuscript, the authors elucidate a novel and therapeutically interesting protein, EfpA from Mycobacterium tuberculosis, identifying and characterizing an unexpected putative lipid-flippase mechanism. This work utilized a combination of CryoEM and computational structure prediction to characterize Mycobacterium tuberculosis (Mtb) efflux pump (EfpA) – a critical component in multi-drug resistant Mtb. In addition to elucidation of the inward, outward, and inhibited conformations, the authors discovered an unexpected lipid-transport path, reminiscent of the MFS-lysophospholipid transporters. The discovery of the lipid-transporting moieties helps to shine a light on the recently observed role of EfpA in cell-growth, and furthermore provide a mechanistic framework for which novel therapeutics could be designed to target specific functions of EfpA. Overall the structural biology experiments appear to be sound, the manuscript

reads well and is interesting, though there are critical aspects of the text/figures which require addressing.

We thank the reviewer for supportive comments. Please see below for our point-by-point response.

Major points:

1. The authors acknowledge the difficulty of identifying the endogenous substrate to MtbEfpA, however what was not mentioned was that lipid-binding-site-B coincides with the putative substrate-binding-site. While the authors did mention that the mutual interplay between the lipid-transport and substrate-transport is currently unknown, it would undoubtedly alter the "comprehensive" lipid-transport model.

The hydrophilic substrates of EfpA, including ethidium bromide (EtBr) and drugs (e.g. INH), presumably bind in the central pocket, which is surrounded by TM1, TM2, TM5, TM12, as well as the head groups of lipid B, lipid C (lines 172-174, Fig. 4d and Extended Data Fig. 10b). Thus, the putative substrate binding site is distinct from the lipid-binding site B. Whether there is a synergistic relationship between drug efflux and lipid transport mediated by EfpA remains unclear. We have revised the text (lines 174-176) and Extended Data Fig. 10b for better clarity.

2. Regarding the lipid-transport model proposed, there are two complete Inward-Open to Outward-Open transitions, i.e., cycles, which are needed to transport a single lipid. What evidence supports this?

Based on available data, we cannot be certain about the number of cycles needed to transport a single lipid from the inner leaflet to the outer leaflet. We have revised the text (lines 281-283, 288-290) to make it clear that this is currently a hypothesis.

3. The authors data suggest that the anti-parallel MtbEfpA-dimer conformation may be an artifact of sample-solubilization (i.e., utilizing allophycocyanin), and demonstrate that the overall monomer structure is unchanged when compared to MsEfpA monomer. Though I fail to see an explanation as to why MsEfpA doesn't form a dimer. Could there be an important structural feature in Mtb which could impart this strange dimerization?

Shown in many structural studies of membrane proteins, the use of different detergents for solubilization and purification can result in different aggregation states of proteins. In our work, EfpA exists predominantly in dimeric form when using Lauryl Maltose Neopentyl Glycol (LMNG) with Cholesteryl Hemisuccinate (CHS). However, using dodecyl-β-D-maltoside (DDM) with CHS, both *Mt*EfpA and *Ms*EfpA mainly exist as monomers (Extended Data Fig. 1b and Extended Data Fig. 3a).

4. In line 171, Y97 is proposed to be the proton acceptor-donor, an exchange which is critical for the proposed lipid-transport model. Tyrosine is not typically considered protonatable as the side-chain pKa at STP is >> 7 (not unlike Arg & Lys). Given that lipid-binding site B (and substrate binding pocket?) is perpetually solvated, it would seem as though it would seldom experience a

local pH high enough to deprotonate. Is there any evidence that suggest Y97 could exchange protons?

Yes, the pKa value of tyrosine's phenolic group is around 10, and it tends to be deprotonated at alkaline pH and protonated at neutral and acidic condition. Tyrosine has ability to bind to cations in certain biochemical contexts. Previous study reported that a tyrosine (Y55) is located adjacent to Na+-coupled residues and may stabilize Na+ in the lipid transporter MFSD2A (reference PMID: 34349262). However, the Y97 in *Mt*EfpA does not align to the Y55 in MFSD2A, therefore, we have revised the statement regarding Y97 and removed the statement of Y97's potential role in protonation/deprotonation (lines 176-178).

Minor points:

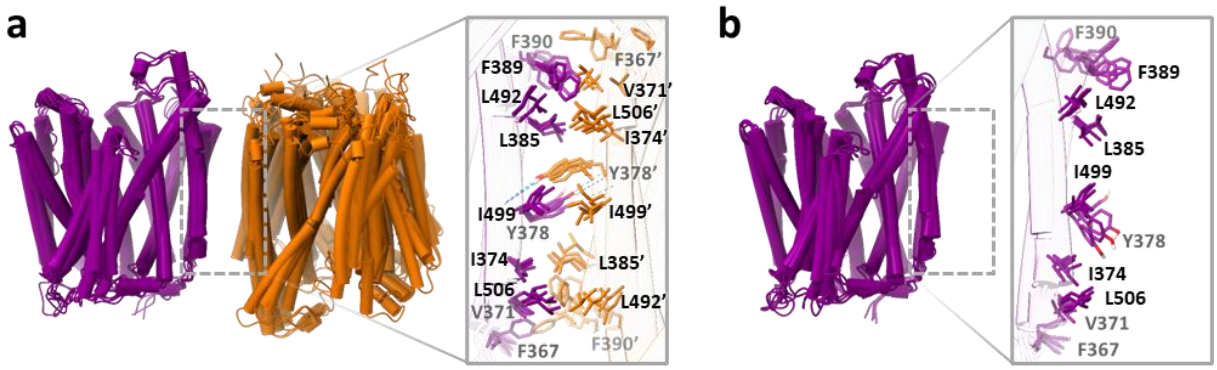
1. There was no mention of the specific parameterization of BRD8000.3 and BRD9327. Parameterization is a non-trivial process and could have substantial consequences to the observed interactions between the inhibitors and the protein.

Thanks for pointing this out. The ligand molecules, including CDL, PE, and BRD, were assigned GAFF2 parameters and AM1-BCC charge using the antechamber and parmchk2 modules in AmberTools 22 package. We have added this description in the method (lines 401-404).

2. It is unclear to me whether the MD simulation were performed for a monomeric or dimeric protein? In any case, I believe it would be crucial to show a comparative (i.e., monomeric vs dimeric) analysis of lipid-protein interactions. Furthermore, a simulation analysis concerning the intersubunit interactions that stabilize the dimer needs to be included (i.e. regarding the contribution of protein-protein and protein-lipid interactions in stabilizing the dimer).

We appreciate the suggestion regarding MD simulation. We have conducted MD simulations for both monomeric and dimeric forms of *Mt*EfpA/*Mt*EfpA-BRD, and included them in Extended Data Fig. 11. The results indicate that the RMSD of lipids at lipid-binding site B and C is approximately 4 Å, with considerable fluctuation, observed in both dimeric and monomeric *Mt*EfpA MD simulation. Specifically, the RMSD of cardiolipin in dimeric *Mt*EfpA stabilizes around 3 Å, while in the monomeric state, it exhibits significant fluctuations and shows the exit of cardiolipin from the pocket (Extended Data Fig. 11a, bottom). This discrepancy in cardiolipin behavior is rationalized by three lipid tails locating outside of lipid-binding channel in monomeric *Mt*EfpA.

The dimeric interface of *Mt*EfpA-dimer after MD simulation aligns with the cryo-EM structure, characterized by hydrophobic interactions and the presence of a delocalized pi bond and hydrogen bond between the two Tyr378 (Comments Fig. 2a). Conversely, in the monomeric *Mt*EfpA from MD simulation, this tyrosine exhibits flexible (Comments Fig. 2b). These findings suggest that Tyr378 contributes to dimer formation, which is consistent with the SEC and negative stain results of Y378A mutation (Extended Data Fig. 3 b and c).



Comments Fig. 2 Comparison of the dimeric interface in monomeric and dimeric *Mt*EfpA from molecular dynamics simulation. a, Dimeric interface of *Mt*EfpA-dimer from the final frame of four MD simulations. The hydrogen bond is represented by a blue dashed line. **b**, Residues at the dimeric interface of *Mt*EfpA-monomer from the final frame of four MD simulations.

3. Extended Figure 5a should say “monomer” as opposed to “monoer”

Corrected.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors addressed my concerns sufficiently.

Reviewer #2 (Remarks to the Author):

The author's have provided a satisfactory response to my prior criticisms.

Reviewer #3 (Remarks to the Author):

This is my second review on the paper entitled: Structures of the essential efflux pump EfpA from Mycobacterium tuberculosis reveal the mechanisms of substrate transport and small-molecule inhibition. This review is based on my expertise in the lipidomics field and thus does not critically evaluate the

other parts. All of my questions have been answered and were implemented in the manuscript which improved the comprehensibility of the used lipidomics approach. However, I have still some concerns:

Thank you so much for including the raw data. However, I have some issues with it from briefly checking the correct assignment of lipids:

1. PE 742 m/z was assigned as PE 16:1/20:1: the MSMS fragmentation pattern does not fit to the assigned lipid, it better fits to PE 18:1/18:1

--> Therefore I would suggest another round of data evaluation and to check the MSMS fragmentation patterns (especially for the cardiolipins as they are the basis for your argumentation)

2. By checking the raw files, I saw that every sample was measured once in negative and positive polarity. I understand that it is just a qualitative assessment, but in my opinion, every sample should have been injected twice per polarity (technical replicates) and there should have been at

least 3 bioreplicates to compare the data with each other. Now with just one measurement, I am unsure whether this is a good foundation for your well described hypothesis.

3. I am missing informations on normalisation of data. As you stated out, you used % of total ion counts, which is ok. However, in the data you provided, the standard you added has varying intensities in the different samples (e.g. for LPE in the monomer sample: 0), which, in my opinion, has to be considered in the calculation. It seems that ionization efficiancy was different in both samples. Please check.

4. Could you please add information on the protein concentration used before lipid extraction? In the text, you write 20 µL. I assume, you measured the protein concentration beforehand and started lipid extraction with equal concentrations in both samples.

Minor points:

1. In the monomer, PE, ether-linked PE and Hexosylceramide were more abundant. Although its not entirely the scope of your study, how do you explain those differences?

2. Maybe you could add informations on adducts: e.g for Cardiolipin [M-2H]2- was used.

Reviewer #4 (Remarks to the Author):

I feel the manuscript by Wang et al greatly improved upon revision and is now eligible for publication. The authors satisfactorily addressed all my points.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors addressed my concerns sufficiently.

Reviewer #2 (Remarks to the Author):

The author's have provided a satisfactory response to my prior criticisms.

Reviewer #3 (Remarks to the Author):

This is my second review on the paper entitled: Structures of the essential efflux pump EfpA from Mycobacterium tuberculosis reveal the mechanisms of substrate transport and small-molecule inhibition. This review is based on my expertise in the lipidomics field and thus does not critically evaluate the other parts. All of my questions have been answered and were implemented in the manuscript which improved the comprehensibility of the used lipidomics approach.

We thank the reviewer for confirming that we have addressed all questions and improved the comprehensibility of lipidomics approach.

However, I have still some concerns:
Thank you so much for including the raw data. However, I have some issues with it from briefly checking the correct assignment of lipids:

1. PE 742 m/z was assigned as PE 16:1/20:1: the MSMS fragmentation pattern does not fit to the assigned lipid, it better fits to PE 18:1/18:1

--> Therefore I would suggest another round of data evaluation and to check the MSMS fragmentation patterns (especially for the cardiolipins as they are the basis for your argumentation)

We thank the reviewer for pointing out a slightly different lipid species assignment from any other open-source bioinformatics software. We would like to note that the spreadsheet data was provided as requested in its entirety without any modification or manual re-assignment. We strongly believe that it is important to maintain full transparency of the mass spectrometry data.

While different reassignment of PE 742 m/z (16:1/20:1 -> 18:1/18:1) is certainly possible, it is worth noting that the overall PE as a lipid class is related to the main points in the manuscript and should remain a focus of discussion. Thus, we had previously illustrated the differential abundance on overall lipid class (cf. acyl chains and saturation level).

2. By checking the raw files, I saw that every sample was measured once in negative and positive polarity. I understand that it is just a qualitative assessment, but in my opinion, every

sample should have been injected twice per polarity (technical replicates) and there should have been at least 3 bioreplicates to compare the data with each other. Now with just one measurement, I am unsure whether this is a good foundation for your well described hypothesis.

We agree with the reviewer that, to draw a quantitative conclusion, it is essential to use technical replicates and bioreplicates. However, the purpose of this qualitative experiment is to provide additional support for our modeling of cardiolipin into the extra density at the dimer interface, instead of accurate assignment of cardiolipin or any lipids from lipidomics data as definitive substrate of EfpA.

As we explained with great details in the first round of revision, EfpA likely functions as a monomer and cardiolipin is probably not relevant to EfpA's physiological functions. We do acknowledge this limitation and would like to further clarify this in the manuscript. In the revised manuscript, we have added "one biological repeat was performed for this experiment" to the legend of Extended Data Fig. 1.

3. I am missing informations on normalisation of data. As you stated out, you used % of total ion counts, which is ok. However, in the data you provided, the standard you added has varying intensities in the different samples (e.g. for LPE in the monomer sample: 0), which, in my opinion, has to be considered in the calculation. It seems that ionization efficiancy was different in both samples. Please check.

We appreciate the comments on data normalization. From the experience of our lab and other lipidomcis groups in the field, differential ionization efficiencies in positive and negative polarity modes for phosphotidylcholine and phosphotidylethanolamine are expected. However, typically only single mode is analyzed (positive or negative, not both) and included in final analysis. It is important to point out that, without a standard curve titration, we would hesitate to calculate the difference using the strategy suggested by the reviewer. Therefore, we have used the % of total ion counts as a general strategy for normalization.

4. Could you please add information on the protein concentration used before lipid extraction? In the text, you write 20 µL. I assume, you measured the protein concentration beforehand and started lipid extraction with equal concentrations in both samples.

Thanks for pointing this out. The protein concentration is 0.1 mg/mL. We have included this protein concentration into the revised methods.

Minor points:

1. In the monomer, PE, ether-linked PE and Hexosylceramide were more abundant. Although its not entirely the scope of your study, how do you explain those differences?

The observed differences may not be physiologically relevant, since the lipids were extracted from the membrane protein preparation that was from mammalian overexpression system and went through extensive purification procedure using detergents. We believe that PE, ether-linked PE and hexosyleramides are non-specific lipids (non-evident from Cryo-EM) that were enriched during the protein purification process.

2. Maybe you could add informations on adducts: e.g for Cardiolipin [M-2H]2- was used.

We appreciate the reviewer's suggestion of adding information on adducts. The adduct information was provided in the corresponding MFQL files as requested. We believe that inclusion of adduct information on the final spreadsheet is unnecessary and not the focus on the manuscript.

Reviewer #4 (Remarks to the Author):

I feel the manuscript by Wang et al greatly improved upon revision and is now eligible for publication. The authors satisfactorily addressed all my points.