

**Structural and functional characterization of a MATE family multidrug
resistance transporter from *Pyrococcus furiosus***

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich 14

Biochemie, Chemie und Pharmazie

der Johann Wolfgang Goethe-Universität

in Frankfurt am Main

von

Sandra Zakrzewska

aus Eschweiler

Frankfurt am Main (2019)

(D 30)

vom Fachbereich 14 - Biochemie, Chemie und Pharmazie der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

1. Gutachter: Prof. Dr. Klaas Martinus Pos

2. Gutachter: Prof. Dr. Dr. h.c. Hartmut Michel

Datum der Disputation:

Diese Doktorarbeit wurde vom 28. Mai 2014 bis zum 20. Mai 2019 unter Leitung von Prof. Dr. Dr. h.c. Hartmut Michel in der Abteilung für Molekulare Membranbiologie am Max Planck Institut für Biophysik in Frankfurt am Main durchgeführt.

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig angefertigt habe und keine weiteren Hilfsmittel und Quellen als die hier aufgeführten verwendet habe.

(Sandra Zakrzewska)

Frankfurt am Main, den

Publication:

Zakrzewska S., Mehdipour A.R., Malviya V.N., Nonaka T., Koepke J., Muenke C., Hausner W., Hummer G., Safarian S., Michel H. Inward-facing conformation of a multidrug resistance MATE family transporter. *Proc. Natl. Acad. Sci.* (2019)

I. Zusammenfassung	I
II. Summary	VII
III. List of figures	XII
IV. List of tables	XVI
V. Abbreviations	XVII
1. Introduction	- 1 -
1.1. Multidrug transporters	- 1 -
1.2. MATE family transporters	- 4 -
1.2.1. Substrate spectrum of the MATE transporters	- 5 -
1.2.2. Ion coupling in the MATE transporters	- 7 -
1.2.3. Crystal structures of the MATE transporters	- 8 -
1.2.3.1. NorM subfamily	- 11 -
1.2.3.2. DinF subfamily	- 14 -
1.2.3.2.1. The PfMATE transporter from <i>Pyrococcus furiosus</i>	- 18 -
1.2.3.3. Eukaryotic MATE subfamily	- 20 -
1.3. <i>Pyrococcus furiosus</i> - an anaerobic hyperthermophile	- 22 -
1.4. Archaeal lipids	- 24 -
1.5. Structural biology of membrane proteins	- 27 -
1.6. Motivation and aim of this work	- 30 -
2. Materials and methods	- 33 -
2.1. Prokaryotic strains and plasmids	- 33 -
2.2. Extraction of genomic DNA from <i>P. furiosus</i>	- 33 -
2.3. Determination of DNA concentrations	- 34 -
2.4. Polymerase chain reaction (PCR)	- 34 -
2.5. Agarose gel electrophoresis	- 36 -
2.6. Gel extraction of DNA	- 36 -
2.7. Gibson Assembly cloning	- 36 -

2.8. Transformation of <i>E. coli</i> cells	- 38 -
2.9. Isolation of plasmid DNA	- 38 -
2.10. Restriction enzyme reaction	- 39 -
2.11. DNA sequencing	- 39 -
2.12. Homologous gene expression in <i>P. furiosus</i>	- 40 -
2.12.1. ½ SME medium preparation	- 40 -
2.12.2. Transformation of plasmid DNA into <i>P. furiosus</i> cells	- 42 -
2.12.3. Cultivation of <i>P. furiosus</i> cells	- 43 -
2.13. Heterologous gene expression in <i>E. coli</i>	- 44 -
2.14. Membrane preparation and solubilization of membrane proteins	- 44 -
2.15. Protein purification	- 45 -
2.15.1. Affinity chromatography	- 45 -
2.15.2. Size exclusion chromatography	- 46 -
2.16. Protein concentration measurement	- 46 -
2.17. Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)	- 47 -
2.18. Non-denaturing polyacrylamide gel electrophoresis (BN-PAGE)	- 48 -
2.19. Immunoblotting (Western Blot)	- 48 -
2.20. Lipid extraction	- 49 -
2.21. Mass spectrometry-based lipidomics	- 50 -
2.22. Crystallization	- 50 -
2.23. Data collection and structure refinement	- 52 -
2.24. Molecular dynamics simulations	- 54 -
2.25. Differential scanning fluorimetry (DSF)	- 55 -
2.26. Native mass spectrometry	- 56 -
3. Results	- 58 -
3.1. PfMATE produced homologously	- 59 -
3.1.1. Construction of the shuttle vector for the homologous expression of <i>pf0708</i>	- 59 -
3.1.2. Gene expression in <i>P. furiosus</i>	- 63 -

3.1.3. Purification of PfMATE produced homologously	- 64 -
3.1.4. Crystallization of PfMATE produced homologously	- 70 -
3.2. PfMATE produced heterologously	- 73 -
3.2.1. Heterologous expression of <i>pf0708</i> in <i>E. coli</i>	- 73 -
3.2.2. Purification of PfMATE produced heterologously	- 74 -
3.2.3. Blue native PAGE results	- 76 -
3.2.4. Mass spectrometry based lipidomics of <i>P. furiosus</i> lipids	- 77 -
3.2.5. Crystallization of PfMATE produced heterologously	- 78 -
3.2.6. Data collection and structure determination	- 83 -
3.2.7. Crystal structure of PfMATE in the outward facing conformation	- 88 -
3.2.8. Crystal structure of PfMATE in the inward-facing conformation	- 90 -
3.3. Functional studies on PfMATE	- 93 -
3.3.1. Drug susceptibility test in <i>E. coli</i>	- 93 -
3.3.2. Drug susceptibility test in <i>P. furiosus</i>	- 98 -
4. Discussion	- 102 -
4.1. Homologous and heterologous production of PfMATE	- 102 -
4.2. Structural and functional implications of the TM1 bending	- 104 -
4.3. Structural basis of the extracellular barrier formation	- 110 -
4.4. Structural basis of the intracellular barrier formation	- 112 -
4.5. Ion binding sites	- 115 -
4.6. Interaction of PfMATE with the archaeal lipids	- 120 -
4.2. Functional studies on PfMATE	- 129 -
5. Concluding remarks	- 131 -
6. Collaborative work	- 135 -
7. References	- 137 -
8. Appendix	- 150 -
Acknowledgments	- 171 -
Curriculum vitae	- 173 -

I. ZUSAMMENFASSUNG

Die Multidrug/Oligosaccharidyl-Lipid/Polysaccharid (MOP)-Transporter-Superfamilie vermittelt den Export einer Vielzahl von Molekülen von physiologischer und pharmakologischer Bedeutung durch biologische Membranen. Gemäß der Transporter-Klassifikationsdatenbank (TCDB) ist die MOP-Superfamilie hauptsächlich in sechs weit voneinander entfernte Familien unterteilt, die funktionell charakterisiert sind: die Multidrug und Toxizitätsextrusion (MATE) Familie, die Polysaccharid-Transporter (PST) Familie, die Oligosaccharidyl-Lipid-Flippase Familie (OLF), die Mausvirulenzfaktor (MVF) Familie, die Agrocin 84 Antibiotika-Exporteur (AgnG) Familie und die progressive Ankylose (Ank) Familie. Von diesen sind die Transporter der MATE-Familie mit Multidrug-Resistenz am weitesten verbreitet und in allen Domänen des Lebens präsent: Archaeen, Bakterien und Eukaryonten. Als sekundär aktive Transporter nutzen sie elektrochemische Ionengradienten von Na^+ und/oder H^+ über die Membran, um den Export von Xenobiotika oder zytotoxischen Stoffwechselabfällen mit einer Spezifität hauptsächlich für polyaromatische und kationische Substrate anzutreiben. Der aktive Export von Medikamenten und toxischen Verbindungen durch Multidrug-Transporter ist eine der Strategien bakterieller Krankheitserreger um Multidrug-Resistenz zu erlangen. Die Anwesenheit von MATE-Proteinen führt zu Resistenzen z.B. gegen Fluorchinolone und Aminoglykosid-Antibiotika bei bakteriellen Infektionen und gegen Chemotherapeutika bei der Tumorbehandlung. Die Entwicklung von Inhibitoren von MATE-Transporter ist damit ein vielversprechendes Ziel zur Bekämpfung von Multidrug-Resistenzen und damit zur Bewältigung eines schwerwiegenden globalen Gesundheitsproblems.

Aufgrund von Ähnlichkeiten ihrer Aminosäuresequenzen werden die Mitglieder der MATE-Familie in die NorM-, die DNA-damage-inducible Protein F (DinF)- und die eukaryotische

Unterfamilie eingeteilt. Strukturelle Informationen über die alternativen Konformationszustände und die Kenntnis des detaillierten Mechanismus des MATE-Transports sind für das strukturbasierte Wirkstoffdesign von großer Bedeutung. In den letzten zehn Jahren wurden die Kristallstrukturen repräsentativer Mitglieder der Unterfamilien NorM, DinF und eukaryotisch veröffentlicht. Sie alle weisen eine ähnliche Gesamtarchitektur mit 12 Transmembran-Helices (TMs) auf, die in zwei Domänen unterteilt sind, die N-terminale Domäne (TMs 1-6) und die C-terminalen Domäne (TMs 7-12). Beide sind durch eine zytoplasmatische Schleife zwischen TM6 und TM7 verbunden (Abb. II.1).

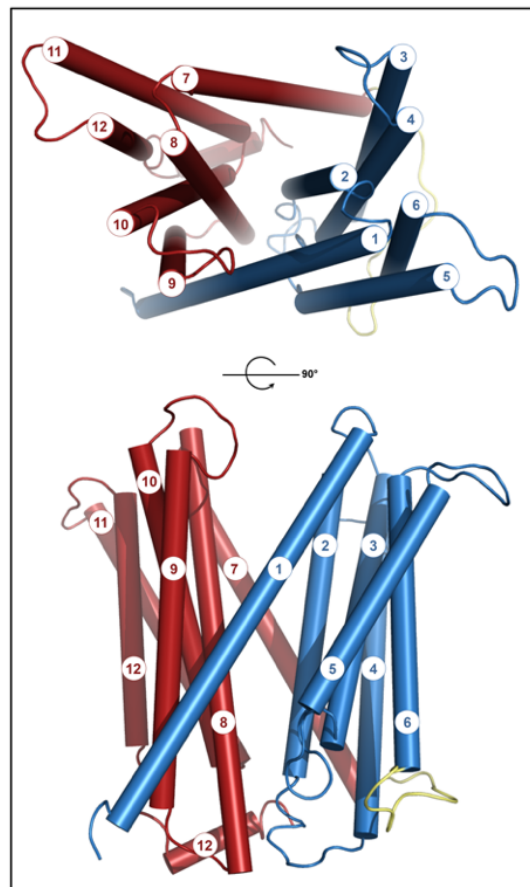


Abb. II.1. Gesamtarchitektur der MATE-Transporter. Schematische Darstellung der NorM_VC-Kristallstruktur, welche die Existenz von 12 TMs zeigt, die in zwei Domänen unterteilt sind, die N-Domäne (TMs 1-6, blau) und die C-Domäne (TMs 7-12, rot). Die zytoplasmatische Schleife, die diese beiden Domänen zwischen TM6 und TM7 verbindet, ist gelb dargestellt.

Da alle verfügbaren Strukturen der MATE-Familie nur in V-förmigen Zuständen bekannt sind, bei denen der zentrale Bindungshohlraum von der extrazellulären Seite aus zugänglich ist, fehlt ein detailliertes Verständnis des gesamten Transportzyklus. Um die zugrunde liegenden Schritte des Transports durch MATE-Transporter zu verstehen, ist die Kenntnis ihrer Strukturen insbesondere in nach innen gerichtete Konformationen, erforderlich.

Während meiner Doktorarbeit wurden strukturelle und funktionelle Studien an einem Transporter der MATE-Familie (DinF-Unterfamilie), PfMATE, des hyperthermophilen und anaeroben Archäons *Pyrococcus furiosus* durchgeführt. Dieses Protein wurde rekombinant homolog in *Pyrococcus furiosus*, sowie heterolog in *Escherichia coli* produziert, und für die nachfolgenden Reinigungs- und Kristallisationsversuche nach der Dampfdiffusions- (VD) und der lipidischen kubischen Phase (LCP) Methode verwendet. PfMATE ist nach meinem Wissen das erste Beispiel für eine erfolgreiche homologe Produktion eines Membranproteins in *P. furiosus*. Aufgrund der sehr geringen Menge des gereinigten Proteins aus der nativen Quelle wurden die heterolog in *E. coli* produzierten PfMATE-Proben typischerweise für die umfangreichen Strukturstudien verwendet.

Kristallstrukturen von PfMATE wurden zuvor in einer nach außen gerichteten Konformation in zwei verschiedenen Zuständen bestimmt, die auf der Struktur der TM1 definiert ist (gebogen und gerade). Ein pH-abhängiger Konformationsübergang von TM1, der durch den Protonierungszustand des konservierten Asp41 reguliert wird, wurde vorgeschlagen. Dieser Befund wurde jedoch kontrovers diskutiert. Dies führte zu der Hypothese, dass die TM1-Biegung durch Wechselwirkungen mit exogenen Lipiden (Monoolein), die unter den Kristallisationsbedingungen vorliegen, beeinflusst wird und nicht durch den niedrigen pH-Wert.

Basierend auf diesen offenen Fragen wurde im Rahmen meines Promotionsvorhabens ein experimenteller Ansatz zur Untersuchung der Rolle von Lipiden als strukturelle und funktionelle Modulatoren von PfMATE entwickelt. Die Wechselwirkung zwischen Membranproteinen und Lipiden kann die Topologie, Struktur und Funktion von Membranproteinen beeinflussen. Um die Unterschiede zwischen den Lipiden von Archäen und Bakterien zu berücksichtigen, folgte auf die Kultivierung von *P. furiosus*-Zellen und die Extraktion ihrer Lipide und eine auf Massenspektrometrie (MS) basierende Lipidomik zur Identifizierung der einzelnen Lipide von *P. furiosus*. Um den Effekt von Lipiden auf die Struktur von PfMATE zu ermitteln, wurden verschiedene Lipidmoleküle für Co-Reinigungs- und Co-Kristallisationsversuche eingesetzt. Diese Dissertation beschreibt, wie die Struktur eines MATE-Transporter im lang gesuchten nach innen offenen Zustand ermittelt werden konnte. Entscheidend für diesen Erfolg war, dass der heterolog produzierten PfMATE-Transporter in Gegenwart von Lipiden aus seiner nativen Quelle *P. furiosus* kristallisiert wurde. Zusem wurden neue Strukturen in der nach außen offenen Konformation von PfMATE bestimmt. Die dazu notwendigen Kristalle wurden unter den relativ sauren pH-Bedingungen gezüchtet. Diese Ergebnisse lassen die zuvor vorgeschlagenen pH-abhängigen Strukturveränderungen innerhalb der TM1 in einem neuen Licht erscheinen.

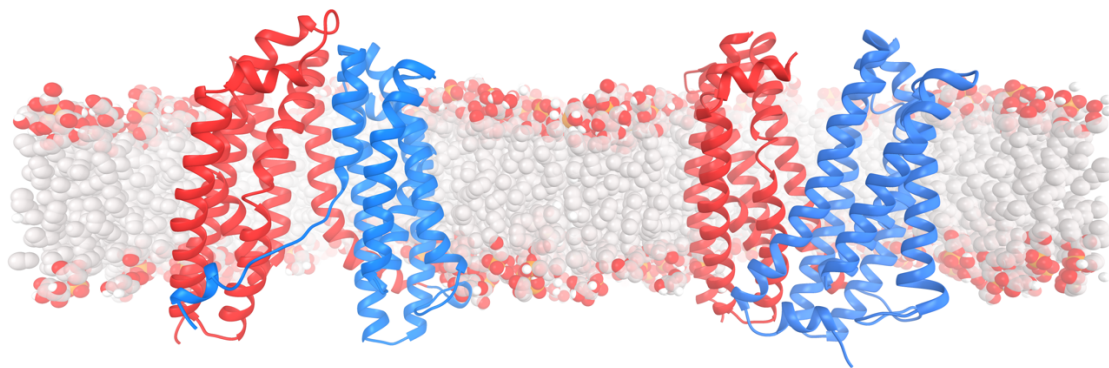


Abb. II.2. Schematische Darstellung von PfMATE-Strukturen. PfMATE in der nach außen offenen (recht) und nach innen offenen Konformation (links). Die N-terminale Domäne ist blau dargestellt, die C-terminale Domäne rot.

Interessanterweise wurden die Strukturen von PfMATE im nach innen oder außen offenen Zustand mit Hilfe von Kristallen bestimmt, die unter ähnlichen Bedingungen, aber in Anwesenheit und Abwesenheit von natürlichen Lipiden erhalten wurden. Diese Beobachtung stützt die Hypothese der Relevanz physiologischer Lipide, die als Konformationsmodulatoren oder/und eine neue Klasse von Substraten fungieren und somit das Substratspektrum der Transporter der MATE-Familie erweitern. Die vergleichende Analyse zweier PfMATE-Zustände zeigt, dass der Übergang vom nach außen offenen zum nach innen offenen Zustand starre Körperbewegungen (*rigid body movements*) der TMs 2-6 und 8-12 zur Bildung eines umgekehrten V umfasst. Der Übergang wird durch eine lose Bindung der TMs 1 und 7 an ihre jeweiligen Bündel und ihre Flexibilität erleichtert. Lokale Strukturänderungen innerhalb von TM1 in der nach innen offenen Konformation, einschließlich einer Auflösung der helikalen Struktur in der intrazellulären Hälfte, zeigen ihre hochflexible Natur, die für die Ionen- und Substrattransport geeignet ist.

Um die Interaktion zwischen PfMATE und seinen nativen Lipiden zu untersuchen, wurden auch Molekulardynamiksimulationen des in eine Archäenlipiddoppelschicht eingebetteten Transporters durchgeführt. Diese Simulationen zeigen unter nativen Bedingungen den Lipidzugangsweg in den zentralen Hohlraum von PfMATE und belegen auf dessen funktionelle Relevanz. Native Massenspektrometrie (MS) Ergebnisse an mit *P. furiosus* Lipiden rekonstituierten Proteins zeigen, dass die Archäenlipidmoleküle trotz umfangreicher Reinigungsschritte auch im Vakuum fest an PfMATE gebunden bleiben. Sie liefern damit einen Nachweis für die an diesem Transporter auftretenden Lipidbindungsereignisse. Die Daten aus SAD (single wavelength anomalous dispersion)-Experimenten mit anomalen Signalen von Cs⁺ und die Ergebnisse der MD-Simulationen sind mit dem Vorhandensein einer monovalenten Kationenbindungsstelle in der Nähe des Aspartarests Asp41 kompatibel. Eine solche wäre mit einem natriumionenabhängigen Transportmechanismus von PfMATE vereinbar.

Basierend auf den Ergebnissen meines Promotionsvorhabens konnte ein Rahmen für die strukturellen Veränderungen, die bei dem Übergang zwischen den nach innen und außen offenen Konformationen des MATE-Transporter auftreten, festgelegt werden. PfMATE von *Pyrococcus furiosus* ist heute das einzige Mitglied der MATE-Familie, dessen Strukturen in zwei verschiedenen Konformationen bekannt sind. Diese Arbeit liefert die ersten strukturellen Beweise für den alternierenden Zugangssmechanismus (*alternating access*), verbessert damit das aktuelle Verständnis des MATE-Transportzyklus und bildet eine wichtige Grundlage für das rationelle Design von Inhibitoren.

Die Verwendung nativer Lipide für die Co-Kristallisation ermöglichte das Fixieren und Visualisieren von PfMATE in einer nach innen offenen Konformation. Dieser Ansatz kann auf verschiedene MATE, MOP-Proteine oder sogar sekundäre aktive Transporter im Allgemeinen weitgehend anwendbar sein. Die neuartigen Erkenntnisse aus den oben genannten Strukturermittlungen und Simulationen zeigen die Bedeutung physiologisch relevanter Lipide, und eröffnen eine alternative Sicht auf die Funktion und den Wirkmechanismus der Transporter der MATE-Transporter Familie.

II. SUMMARY

The members of the multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) transporter superfamily mediate export of a wealth of molecules of physiological and pharmacological importance. According to the Transporter Classification Database (TCDB), the MOP superfamily is mainly categorized into six distantly related families functionally characterized families: the multidrug and toxic compound extrusion (MATE), the polysaccharide transporter (PST), the oligosaccharidyl-lipid flippase (OLF), the mouse virulence factor (MVF) the agrocin 84 antibiotic exporter (AgnG), and the progressive ankylosis (Ank) family. Among these, the multidrug resistance MATE family transporters are most ubiquitous, being present in all domains of life: *Archaea*, *Bacteria* and *Eukarya*. As secondary active transporters, they utilize transmembrane electrochemical ion gradients of Na⁺ and/or H⁺ in order to drive the efflux of xenobiotics or cytotoxic metabolic waste products with specificity mainly for polyaromatic and cationic substrates. Active efflux of drugs and toxic compounds carried out by multidrug transporters is one of the strategies developed by bacterial pathogens to confer multidrug resistance. MATE proteins provide resistance to, e.g., fluoroquinolone, aminoglycoside antibiotics, and anticancer chemotherapeutic agents, thus serving as promising pharmacological targets for tackling a severe global health issue.

Based on their amino acid sequence similarity, the MATE family members are classified into the NorM, the DNA-damage-inducible protein F (DinF), and the eukaryotic subfamilies. Structural information on the alternate conformational states and knowledge of the detailed mechanism of the MATE transport are of great importance for the structure-aided drug design. Over the past decade, the crystal structures of representative members of the NorM, DinF and eukaryotic subfamilies have been presented. They all share similar overall architecture

comprising 12 transmembrane helices (TMs) divided into two domains, the N-terminal domain (TMs 1-6) and the C-terminal domain (TMs 7-12), connected by a cytoplasmic loop between TM6 and TM7 (Fig. II.1). Since all available MATE family structures are known only in V-shaped outward-facing states with the central binding cavity open towards the extracellular side, a detailed understanding of the complete transport cycle has remained elusive. In order to elucidate the underlying steps of the MATE transport mechanism, structures of distinct intermediates, particularly inward-facing conformation, are required.

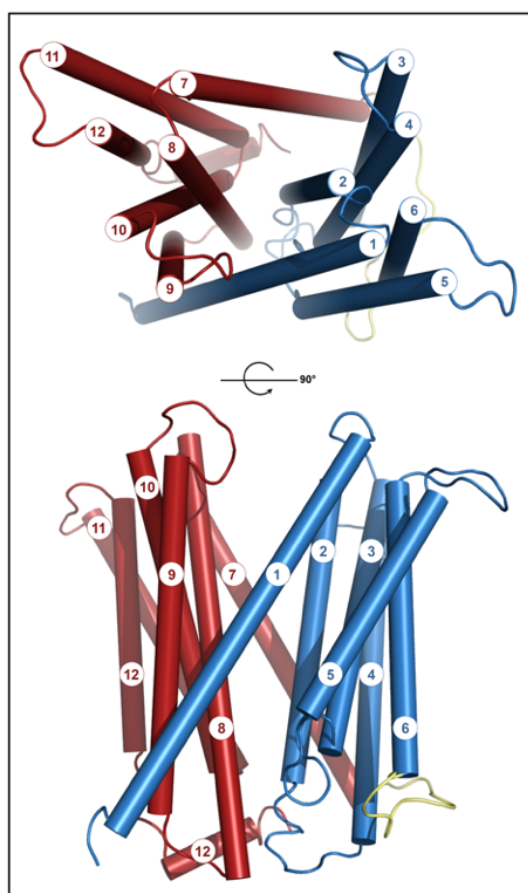


Fig. II.1. Overall architecture of the MATE transporters. Schematic representation of the NorM_VC crystal structure, which shows the presence of 12 TMs divided into two lobes, the N-domain (TMs 1-6, blue) and the C-domain (TMs 7-12, red). The cytoplasmic loop connecting these two domains between TM6 and TM7 is depicted in yellow.

In my PhD project, structural and functional studies have been performed on a MATE family (DinF subfamily) transporter, PfMATE, from the hyperthermophilic and anaerobic archaeon *Pyrococcus furiosus*. This protein was produced homologously in *Pyrococcus furiosus* as well as heterologously in *Escherichia coli*, and used for the subsequent purification and crystallization trials by the vapor diffusion (VD) and lipidic cubic phase (LCP) method. To the best of my knowledge, PfMATE is the first example of a successful homologous production of a membrane protein in *P. furiosus*. Due to the very low final amount of the purified protein from the native source, the heterologously produced PfMATE samples were typically used for the extensive structural studies.

Crystal structures of PfMATE have been previously determined in an outward-facing conformation in two distinct states (bent and straight) defined on the arrangement of TM1. A pH dependent conformational transition of this helix regulated by the protonation state of the conserved aspartate residue Asp41 was proposed. However, it has been discussed controversially, leading to the hypothesis about TM1 bending to be rather affected by interactions with exogenous lipids (monoolein) present under the crystallization conditions.

Based on these open questions, an experimental approach to investigate the role of lipids as structural and functional modulators of PfMATE has been taken in the course of my PhD project. The interplay between membrane proteins and lipids can affect membrane protein topology, structure and function. Considering differences between archaeal and bacterial lipid composition, cultivation of *P. furiosus* cells and extraction of its lipids was followed by the mass spectrometry (MS) based lipidomics for identification of individual lipid species in the archaeal extract. In order to assess the effects of lipids on PfMATE, different lipid molecules were used for co-purification and co-crystallization trials. This dissertation presents a workflow leading to the structure determination of a MATE transporter in the long sought-after inward-

facing state, which has been achieved upon purification and crystallization of the heterologously produced PfMATE in the presence of lipids from its native source *P. furiosus*. Also, the PfMATE outward-facing state obtained from the crystals grown at the acidic pH conditions sheds light on the previously proposed pH-dependent structural alterations within TM1.

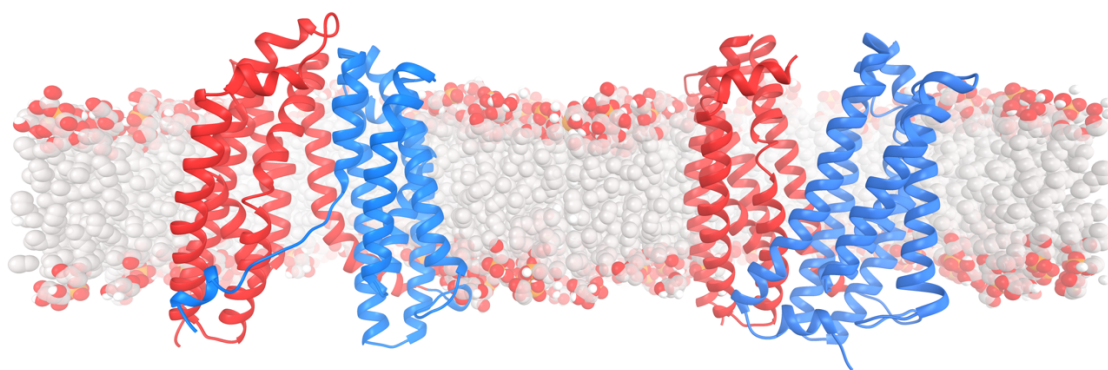


Fig. II.2. Schematic presentation of PfMATE structures. PfMATE in the outward-facing (right) and inward-facing conformation (left). The N-terminal domain is depicted in blue, the C-terminal domain in red.

It is interesting to note that the inward and outward-facing states of PfMATE were obtained from the crystals grown under similar conditions, but in the presence and absence of native lipids, respectively. This observation supports the hypothesis about physiologically relevant lipids to act as conformational modulators or/and a new class of substrates, expanding the substrate spectrum of the MATE family transporters. Comparative analysis of two PfMATE states reveals that transition from the outward to the inward-facing state involves rigid body movements of TMs 2-6 and 8-12 to form an inverted V, facilitated by a loose binding of TMs 1 and 7 to their respective bundles and their conformational flexibility. Local fluctuations within TM1 in the inward-facing structure, including bending and unwinding in the intracellular half of the helix, invoke its highly flexible nature, which is suitable for ion and substrate gating. To investigate the interplay between PfMATE and its native lipids, molecular dynamics (MD)

simulations of the transporter embedded in an archaeal-type lipid bilayer were also performed. These simulations under native-like environment demonstrate the presence of a lipid access pathway towards the central cavity of PfMATE, pointing towards its functional relevance. Native mass spectrometry (MS) results of the protein supplemented with *P. furiosus* lipids reveal that the archaeal lipid molecules remain tightly associated with PfMATE despite extensive purification and delipidation steps, and provides evidence for the lipid binding events occurring in this transporter.

In the light of ion coupling, the combination of the data from the SAD (single-wavelength anomalous dispersion) experiments using anomalous signals of Cs^+ with the results of the MD simulations indicates the presence of a monovalent cation binding site at Asp41, pointing towards a sodium ion dependent transport mechanism of PfMATE.

Based on the results of my PhD project, a framework of the structural changes that occur during the transition between the inward and outward-facing conformation of the MATE transporter could be established. PfMATE from *Pyrococcus furiosus* is now the only MATE family member having structures known in two distinctly opposite conformations. This work provides the first structural evidence for the alternating access mechanism for the MATE family proteins, in consequence enhance the current understanding about their transport cycle, and form an important basis for structure-based rational development of inhibitors.

Usage of native lipids for co-crystallization allowed trapping and visualization of PfMATE in an inward-facing conformation. This approach may be broadly applicable to different MATE, MOP proteins or even secondary active transporters in general. The novel insights from the aforementioned structural and computational analyses highlight a significance of physiologically relevant lipids and open an alternative view on the function and mechanism of action for the MATE family transporters.

III. LIST OF FIGURES

- Abb. II.1. Gesamtarchitektur der MATE-Transporter
- Abb. II.2. Schematische Darstellung von PfMATE-Strukturen
- Fig. II.1. Overall architecture of the MATE transporters
- Fig. II.2. Schematic presentation of PfMATE structures
- Fig. 1.1. Schematic illustration of the multidrug transporters superfamilies in Gram-negative bacteria
- Fig. 1.2. Representative substrates of the MATE family transporters
- Fig. 1.3. Overall architecture of the MATE transporters
- Fig. 1.4. Overall architecture (ribbon model) of the NorM_VC
- Fig. 1.5. Overall architecture (ribbon model) of the NorM_NG
- Fig. 1.6. Overall architecture (ribbon model) of the VcmN
- Fig. 1.7. Overall architecture (ribbon model) of the PfMATE
- Fig. 1.8. Overall architecture (ribbon model) of the Dinf_BH
- Fig. 1.9. Overall architecture (ribbon model) of the ClbM
- Fig. 1.10. Crystal structures of the MATE transporters in complex with substrates or inhibitors
- Fig. 1.11. Overall architecture (ribbon model) of PfMATE in the straight and bent conformations
- Fig. 1.12. Overall architecture (ribbon model) of the AtDTX14
- Fig. 1.13. Overall architecture (ribbon model) of the CasMATE
- Fig. 1.14. Transmission electron micrographs (TEM) of *P. furiosus* cells
- Fig. 1.15. Comparison of bacterial and archaeal lipids
- Fig. 1.16. Lipid composition of *P. furiosus* membrane
- Fig. 1.17. Progress of membrane protein structure determination
- Fig. 1.18. Crystallization methods of membrane proteins
- Fig. 2.1. Schematic representation of the Gibson Assembly method
- Fig. 2.2. Equipment for homologous expression in *P. furiosus*
- Fig. 2.3. Thoma cell counting chamber

- Fig. 2.4. DSF experiment
- Fig. 3.1. Schematic representation of the constructs for the homologous expression of *pf0708*
- Fig. 3.2. 1% (w/v) agarose gels showing DNA fragments after PCR
- Fig. 3.3. 1% (w/v) agarose gel representing the DNA fragments after restriction enzyme digestion of the plasmid DNA
- Fig. 3.4. Growth curve of *P. furiosus* Δ 1623 *pf0708* overexpression strain at 95 °C
- Fig. 3.5. Western Blots and SDS-PAGE gels of PfMATE produced under control of the *gdh* promoter after solubilization and affinity chromatography
- Fig. 3.6. Western Blots and SDS-PAGE gels of PfMATE produced under control of the *pf0613* promoter after affinity chromatography
- Fig. 3.7. Affinity chromatography columns elution profiles of PfMATE produced homologously
- Fig. 3.8. Affinity chromatography columns elution profiles of PfMATE produced homologously under the control of two different promoters
- Fig. 3.9. Gel filtration profiles of PfMATE produced homologously under control of two different promoters
- Fig. 3.10. Gel filtration profile of PfMATE produced homologously
- Fig. 3.11. Western Blot and SDS-PAGE gel after gel filtration
- Fig. 3.12. Western Blot and SDS-PAGE gel after gel filtration
- Fig. 3.13. The cubicon method
- Fig. 3.14. Crystals of PfMATE produced homologously
- Fig. 3.15. Western Blot of the *E. coli* TOP10 cell lysate after heterologous gene expression
- Fig. 3.16. Affinity chromatography column elution profiles of PfMATE produced heterologously
- Fig. 3.17. Gel filtration profile of the heterologously produced PfMATE
- Fig. 3.18. Native PAGE gel of PfMATE after gel filtration
- Fig. 3.19. Lipid composition of *P. furiosus* WT analyzed by mass spectrometry based lipidomics
- Fig. 3.20. Gel filtration of the heterologously produced PfMATE co-purified with the lipids
- Fig. 3.21. Crystals of PfMATE produced heterologously after co-crystallization with the total lipid extract from *P. furiosus*
- Fig. 3.22. Crystals of PfMATE produced heterologously after co-crystallization with diether-PG
- Fig. 3.23. Crystals of PfMATE produced heterologously after co-crystallization with NBD-PG

- Fig. 3.24. Crystals of PfMATE produced heterologously after co-crystallization with NAC6
- Fig. 3.25. Crystals of PfMATE produced heterologously after co-crystallization with DOPG
- Fig. 3.26. Rastering scan of a PfMATE crystal
- Fig. 3.27. Diffraction pattern of PfMATE crystals
- Fig. 3.28. Diffraction anisotropy plot and Xtriage results
- Fig. 3.29. Crystal packing of PfMATE
- Fig. 3.30. Overall architecture of PfMATE in the outward-facing conformation
- Fig. 3.31. Overall architecture of PfMATE in the inward-facing conformation
- Fig. 3.32. Electrostatic surface potential of the central cavity of MOP superfamily proteins
- Fig. 3.33. MIC test in *E. coli* KAM32
- Fig. 3.34. MIC test in *E. coli* C43(DE3) Δ *acrAB*
- Fig. 3.35. MIC test in *E. coli* BL21(DE3) Δ *acrAB*
- Fig. 3.36. MIC test in *E. coli* C41(DE3) Δ *acrAB*
- Fig. 3.37. *P. furiosus* WT and the *pf0708* deletion mutant
- Fig. 3.38. Drug susceptibility assay in *P. furiosus*
- Fig. 4.1. Side view of PfMATE structures
- Fig. 4.2. Schematic presentation of PfMATE structures
- Fig. 4.3. TM1 bending
- Fig. 4.4. PfMATE structures in the outward-facing conformation
- Fig. 4.5. Average B-factors
- Fig. 4.6. Structural superposition of the N- and C-domains of the two distinct PfMATE states
- Fig. 4.7. Fluctuations within the helices TM1 and TM7
- Fig. 4.8. Conformational changes of PfMATE
- Fig. 4.9. Water accessibility
- Fig. 4.10. Extracellular barrier in PfMATE inward-facing conformation
- Fig. 4.11. Charged residues in the inward and outward-facing structure of PfMATE
- Fig. 4.12. Intracellular barrier in PfMATE outward-facing conformation
- Fig. 4.13. Sodium ion binding site

Fig. 4.14. Chloride ion binding site

Fig. 4.15. Structural superposition of MurJ and PfMATE structures

Fig. 4.16. Lipid access pathway

Fig. 4.17. Spontaneous binding and flipping of an archaeal lipid in molecular dynamics simulations of PfMATE

Fig. 4.18. Stability curves from DSF results of PfMATE with and without the native lipid extract

Fig. 4.19. Stability curves from DSF results of the OGNG-treated PfMATE with and without the native lipid extract

Fig. 4.20. Mass spectra of PfMATE in the presence and absence of the archaeal lipids

Fig. 4.21. Mass spectra of PfMATE with increasing concentrations of the archaeal lipids

Fig. 4.22. Genetic changes in *E. coli* C41(DE3) and C43(DE3) comparative with BL21(DE3)

Fig. 5.1. X-ray anomalous scattering coefficients of phosphorous (P) and sulfur (S)

IV. LIST OF TABLES

Table 1.1. Summary of all available MATE family crystal structures

Table 2.1. Prokaryotic strains and plasmids

Table 2.2. PCR reagents

Table 2.3. PCR oligonucleotides

Table 2.4. PCR protocol

Table 2.5. Gibson Assembly reaction

Table 2.6. Restriction enzyme reaction

Table 2.7. SME medium

Table 2.8. Wolfe's minerals

Table 2.9. $\frac{1}{2}$ SME medium

Table 2.10. Purification steps

Table 2.11. SDS-PAGE buffers

Table 2.12. Western blot buffers

Table 2.13. Crystallization screens

Table 2.14. Software for data processing and visualization

Table 3.1. Data collection and refinement statistics

V. ABBREVIATIONS

A	absorbance
ABC	ATP-binding cassette
AbgT	p-aminobenzoyl-glutamate transporter
AgnG	agrocin 84 antibiotic exporter
Ank	progressive ankylosis
ATP	adenosine-5'-triphosphate
b	path length
BCA	bicinchoninic acid
BCIP	5-Bromo-4-chloro-3-indolyl phosphate disodium salt
BN	blue native
BSA	bovine serum albumin fraction V
c	protein concentration
cryoEM	electron cryomicroscopy
CV	column volume
Cymal-6	6-Cyclohexyl-1-Hexyl- β -D-Maltoside
DAPI	2-(4-Amidinophenyl)-6-indolecarbamide dihydrochloride
DDA	data-independent acquisition
DGD	diglycosyl diether
DinF	DNA damage-inducible protein F
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DPA	diphytanyl phosphatidic acid
DPG	diphytanyl phosphatidyl glycerol
DPI	diphytanyl phosphatidyl inositol)
DSC	Differential Scanning Calorimetry
DSF	Differential Scanning Fluorimetry
E	extinction coefficient
EDTA	ethylenediaminetetraacetic acid
eMATE	eukaryotic MATE
FEM	feature enhanced map
g	gravitational acceleration (9.80665 m/s ²)
G-1-P	<i>sn</i> -glycerol-1-phosphate
G-3-P	<i>sn</i> -glycerol-3-phosphate
GPCR	G protein-coupled receptor

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HexNac-PD	diphytanyl phosphatidyl <i>N</i> -acetyl hexose
HPLC	high performance liquid chromatography
IFC	inward facing conformation
IM	inner membrane
isoGDGT-0	isoprenoidal glycerol dialkyl glycerol tetraether
KO	knock-out strain
LB	lysogeny broth
LCP	lipidic cubic phase
LDAO	<i>N,N</i> -dimethyl- <i>N</i> -dodecylamine- <i>N</i> -oxide
LTQ	linear trap quadrupole
MATE	multidrug and toxic compound extrusion
MBH	membrane-bound hydrogenase
MD	molecular dynamics
MES	2-(<i>N</i> -morpholino) ethanesulfonic acid
MFP	membrane fusion protein
MFS	major facilitator superfamily
MGD	monoglycosyl diether
MIC	Minimum Inhibitory Concentration
MOP	multidrug/oligosaccharidyl-lipid/polysaccharide
MOPS	3-(<i>N</i> -morpholino) propanesulfonic acid
MS	mass spectrometry
MTBE	methyl tert-butyl ether
MVF	mouse virulence factor
MWCO	molecular weight cut-off
NAC6	<i>N</i> -acetyl- <i>D</i> -glucosamine 6-phosphate
NADPH	β -Nicotinamide adenine dinucleotide 2'-phosphate, reduced
NBD-	nitro-2-1,3-benzoxadiazol-4-yl
NBT	nitrotetrazolium blue chloride
NCBI	National Center for Biotechnology Information
NTP	nucleoside triphosphate
OD ₆₀₀	optical density at 600 nm
OFC	outward facing conformation
OGNG	Octyl Glucose Neopentyl Glycol
OLC	1-Oleoyl- <i>R</i> -glycerol
OLF	oligosaccharidyl-lipid flippase
OM	outer membrane

OMF	outer membrane factor
PACE	proteobacterial antimicrobial compound efflux
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
P _i	inorganic phosphate
PI	phosphatidylinositol
POPA	palmitoyl oleoyl phosphatidic acid
POPG	palmitoyl oleoyl phosphatidyl-glycerol
POPI	palmitoyl oleoyl phosphatidyl-inositol
PP _i	intracellular inorganic pyrophosphate
PST	polysaccharide transport
PVDF	polyvinylidene fluoride
r.m.s.d.	root-mean-square deviation
R _{free} /R _{work}	reliability factors
RNA	ribonucleic acid
RNase	ribonuclease
RND	resistance-nodulation-division
rpm	revolutions per minute
SAD	single wavelength anomalous dispersion
SDS	sodium dodecyl sulfate
SHI, SHII	soluble hydrogenase I, II
SMR	small multidrug resistance
SOC	Super Optimal broth with Catabolite repression
ssNMR	solid-state nuclear magnetic resonance
TBST	Tris buffered saline with Tween 20
TM	transmembrane helix
T _m	melting point
TPP	Tetraphenylphosphonium
Tris	2-Amino-2-(hydroxymethyl)-1.3-propanediol
UV-Vis	ultraviolet-visible
VD	vapor diffusion
WT	wild type
β-DDM	n-dodecyl-β-D-maltoside

1. INTRODUCTION

1.1. Multidrug transporters

The past few decades have witnessed increasingly rapid advances in the research field of membrane proteins, which, e.g. are responsible for communication between the internal and external environment of the cell¹. It is estimated that genes encoding membrane proteins comprise 20-30% of the majority of genomes in all kingdoms of life², around 23% of the human genome^{3,4}. Based on the membrane-protein interaction mode, membrane proteins can be classified into two main categories: peripheral (extrinsic) membrane proteins, located at the surface of the membrane and integral (intrinsic) membrane proteins, containing one or more segments embedded in the phospholipid bilayer⁵.

Membrane proteins are involved in the fundamental life processes and diverse physiological functions, such as energy conversion, signal transduction, cell adhesion, catalyzing enzymatic reactions or transport of ions and molecules across the membrane. Particularly, the ability to facilitate transport of drugs and toxic compounds has promoted them as an attractive therapeutic target. Currently, more than 60% of all pharmaceutical drugs available on the market target membrane proteins^{4,6}. One of the main challenges in the modern pharmaceutical research is multidrug resistance of different pathogenic genera, for example *Enterobacteriaceae*, *Pseudomonas*, *Burkholderia*, *Acinetobacter*, *Neisseria*, *Klebsiella*^{7,8}.

There are multiple strategies utilized by these organisms to develop multidrug resistance, such as decreased cellular permeability of the drug through the outer membrane, drug degradation or inactivation, modification of the drug target and active drug extrusion, which is considered to be the major mechanism of drug resistance⁹⁻¹². The latter one is carried out by a class of integral membrane proteins called multidrug efflux pumps or transporters, ubiquitously

distributed throughout all domains of life. They are also referred as xenobiotic transporters due to their ability to expel a broad range of hazardous compounds, e.g., organic acids, detergents, dyes, bile acids and hormones^{8,13}.

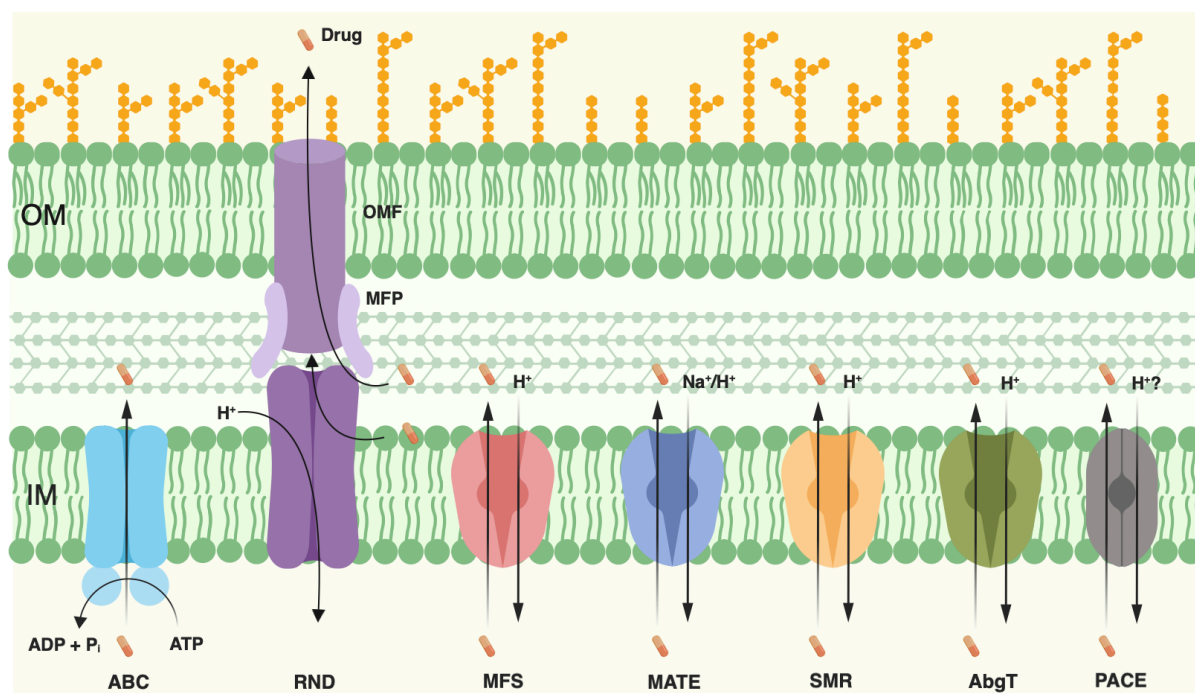


Fig. 1.1. Schematic illustration of the multidrug transporters superfamilies in Gram-negative bacteria. The pathway of substrate and proton (or ATP) dependent transport is indicated by arrows. Except for the PACE family, atomic structures are available for all of the representative members of the ABC, RND, MFS, MATE, SMR and AbgT superfamily. OM, outer membrane; IM, inner membrane; Pi, inorganic phosphate; MFP, membrane fusion protein; OMF, outer membrane factor. Figure adapted from Masi *et al.*¹⁴ and created with BioRender.com.

On the basis of the energy source, type of substrates and overall structural composition, multidrug transporters are divided into seven families^{15,16} (Fig. 1.1): the ATP-binding cassette (ABC)^{17–19}, the major facilitator superfamily (MFS)^{20–22}, the resistance-nodulation-division (RND)^{23–25}, the multidrug and toxic compound extrusion (MATE)^{26–29}, the small multidrug resistance (SMR)^{30,31} and recently identified the proteobacterial antimicrobial compound efflux (PACE)³² as well as the p-aminobenzoyl-glutamate transporter (AbgT) family³³. With the

exception of the RND superfamily, which is exclusively found in Gram-negative bacteria and connects the inner and outer membrane (by forming a tripartite efflux system with the periplasmic membrane fusion protein and an outer membrane channel), these superfamilies are located in the inner membrane and distributed among both Gram-positive and Gram-negative organisms. Nonetheless, there are also a few examples of the tripartite ABC and MFS transporters spanning the whole cell envelope. RND superfamily members are believed to cooperate together with the other superfamilies to form a drug efflux assembly, which fully transports a substrate firstly to the periplasm and subsequently across the outer membrane to the cell exterior¹⁴.

The mechanism of action of the ABC transporters is based on a primary active transport energized by the hydrolysis of ATP, whereas members of all the other superfamilies act as secondary active transporters, catalyzing transport of a substrate against its concentration gradient by using the electrochemical gradient of another ion across the membrane³⁴.

Recent developments in structural biology of membrane proteins, including advances in electron cryomicroscopy (cryoEM), X-ray crystallography and solid-state nuclear magnetic resonance (ssNMR), provide high resolution structures of multidrug transporters, which are essential to expand our understanding of the mechanistic details behind their transport mechanism. Structural insights are also important to develop new strategies against multidrug resistance by rational design of novel compounds inhibiting the function of multidrug efflux pumps³⁵⁻³⁸.

1.2. MATE family transporters

This dissertation investigates structural and functional aspects of a multidrug transporter from the MATE family. According to the Transporter Classification Database (TCDB; www.tcdb.org)³⁹, the MATE family belongs to the multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) superfamily (2.A.66)⁴⁰.

In addition to the MATE family⁴¹, five distantly related families of the MOP superfamily are functionally characterized: the polysaccharide transport (PST)⁴², the oligosaccharidyl-lipid flippase (OLF)⁴³, the mouse virulence factor (MVF)⁴⁴, the agrocin 84 antibiotic exporter (AgnG)⁴⁵ and the progressive ankylosis (Ank) family⁴⁶.

Members of the PST family are involved in transport of polysaccharides and/or their lipid-linked precursors in both prokaryotic and eukaryotic organisms. OLF family proteins have a similar function as they mediate transport of lipid-linked oligosaccharide precursors of glycoproteins, however they are present exclusively in the endoplasmic reticular membranes of eukaryotic cells. The AgnG family includes a single functionally characterized member from bacterium *Agrobacterium radiobacter* K84, exporting the antibiotic agrocin 84, a unique disubstituted analogue of adenosine, which has an inhibitory effect only against *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*⁴⁷. A member of the bacterial MVF family, MviN from *Salmonella typhimurium*, has been shown to be an important virulence factor for mouse and to act as a putative lipid flippase for the lipid II peptidoglycan precursor⁴⁸. Another MVF family transporter, MurJ from *Thermosipho africanus*, has been also classified as a flippase for the lipid-linked peptidoglycan precursor lipid II^{49,50}. The Ank family transporter drives the efflux of intracellular inorganic pyrophosphate (PP_i), thereby controlling bone formation and remodeling⁵¹.

Among the MOP superfamily, MATE proteins represent the most ubiquitous members widely distributed throughout both prokaryotic and eukaryotic organisms⁴¹. They confer drug

resistance to various antibiotics and anticancer chemotherapeutic agents⁵². Therefore, MATE family transporters comprise a significant research field as propitious targets for the development of drugs.

1.2.1. Substrate spectrum of the MATE transporters

Based on the phylogenetic analysis of their amino acid sequences, the MATE family members are classified into the NorM, the DNA damage-inducible protein F (DinF) and the eukaryotic (eMATE) subfamilies⁵³. Members of all subfamilies contribute to resistance against a diverse array of drugs and toxic compounds such as antibiotics and anticancer agents. MATE proteins are rather polyspecific, thus able to transport structurally diverse classes of antibiotics, including fluoroquinolones such as norfloxacin, ciprofloxacin and enoxacin, aminoglycosides, e.g. kanamycin and streptomycin, β -lactams such as ampicillin, chloramphenicol, and many other compounds like ethidium, rhodamine 6G, acriflavine, crystal violet, berberine, doxorubicin, novobiocin⁵⁴⁻⁵⁷.

Apart from drug extrusion, MATE orthologs in plants export secondary metabolic products from the cells such as nicotine, flavonoids and proanthocyanidin precursors⁵⁸. They are also key players in aluminium tolerance by citrate efflux, vacuole sequestration of plant-derived alkaloids and flavonoids, iron homeostasis and regulation of local auxin biosynthesis⁵⁹. In mammalian cells, MATE proteins drive extrusion of various xenobiotic cations. Human MATEs, which are highly expressed particularly in liver and kidney, facilitate the efflux of toxic chemotherapeutic agents, such as platinum containing cisplatin, oxaliplatin as well as metformin, which is primarily used against type 2 diabetes^{60,61}.

The first identified and functionally characterized MATE family member is the NorM family protein from *Vibrio parahaemolyticus*, which mediates multidrug extrusion of norfloxacin,

ciprofloxacin as well as ethidium, kanamycin and streptomycin⁶². Previously published studies of drug susceptibility tests have shown that almost all MATE family transporters provide resistance against fluoroquinolones, for example norfloxacin. Among cationic dyes, ethidium and acriflavine are substrates of several MATE transporters. Notably, many of the characterized MATE proteins exhibit disability to transport negatively charged compounds. Taken together, MATE family members are postulated to be responsible for polyspecific transport of structurally and chemically diverse compounds, although they are typically classified as polyaromatic and cationic substrates (Fig. 1.2).

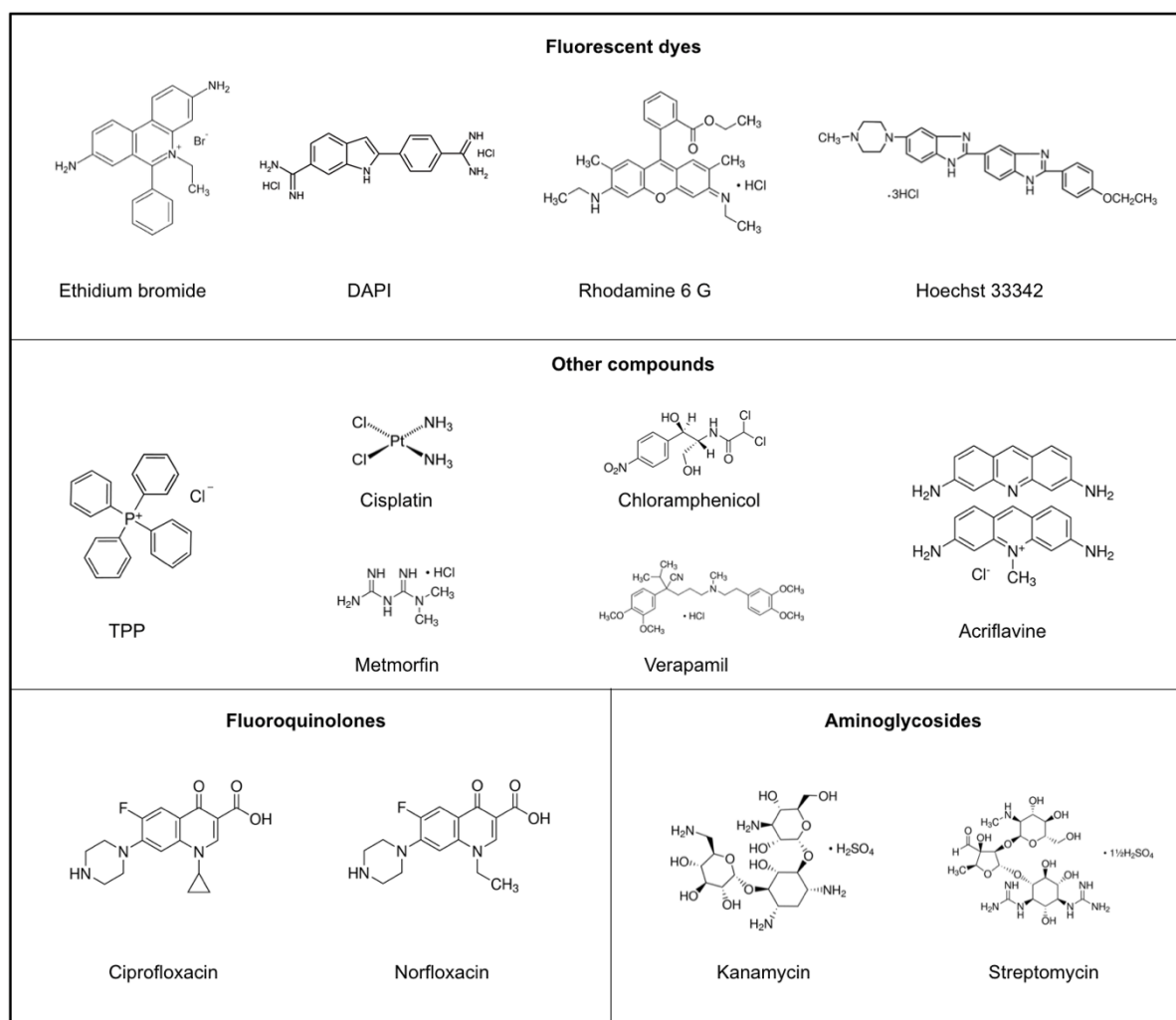


Fig. 1.2. Representative substrates of the MATE family transporters.

1.2.2. Ion coupling in the MATE transporters

MATE proteins act as secondary active transporters and utilize transmembrane electrochemical ion gradients (Na^+ and/or H^+) to drive the uphill movement of substrates across the membrane. Initially, it was believed that transport by MATE proteins is coupled to a sodium ion gradient. The first example of a Na^+ -dependent multidrug transporter was NorM from *Vibrio parahaemolyticus*⁶². Thereafter, functional studies on PmpM from *Pseudomonas aeruginosa* revealed that MATE proteins can also utilize H^+ as a coupling ion⁶³. For many years, MATE transporters were considered to be dependent on exclusively either sodium ion motive force or proton motive force. However, recently published experimental and computational data have shown that some of the MATE family members, for example NorM_VC from *Vibrio cholerae*, previously thought to be Na^+ -dependent, can be coupled to both cations, sodium ions or protons^{63,64}. The same dual ion specificity was recently proposed for ClbM from *E. coli* and PfmATE from *P. furiosus*⁶⁵, which were originally recognized as Na^+ and H^+ -driven transporters, respectively.

Members of the NorM and DinF subfamilies can use a Na^+ and/or H^+ electrochemical gradient to extrude drugs, whereas transport by eMATE transporters, including human MATE transporters, hMATE1 and hMATE2, are typically coupled to proton flux⁶¹. Generally, the coupling mechanism can be either governed by an allosteric coupling of ion and substrate binding in topographically distinct binding sites, or follows the concept of a direct competitive interaction due to an occurrence of a shared ion/substrate binding site.

Except for sodium ions and protons, other monovalent ions including K^+ , Rb^+ and Li^+ were reported to influence the activity of drug efflux of some MATE transporters, e.g. KetM from *Klebsiella pneumoniae* shows enhanced DAPI extrusion in the presence of Na^+ , K^+ or Li^+ ⁶⁶, and in case of the MATE protein from rat, transport of norfloxacin is facilitated by K^+ and Rb^+ . Identification of distinct ion and substrate binding sites from both sodium ion and proton

coupled MATE transporters illustrate their mechanistic diversity, even among the same MATE subfamily. A detailed coupling mechanism of a drug efflux and a counterion influx hasn't been fully elucidated yet.

1.2.3. Crystal structures of the MATE transporters

Over the past decade a major progress in structural biology of the MATE transporters has been made. In recent years all of the high-resolution structures of representative members of the NorM⁶⁷⁻⁷⁰, DinF⁷¹⁻⁷⁴ and eMATE^{75,76} subfamilies were generated by X-ray crystallography (Tab. 1.1). They share a similar V-shaped architecture, consisting of 12 transmembrane helices (TMs), which are divided into two six-helix bundles, referred as the N-terminal (TMs 1-6) and the C-terminal (TMs 7-12) domain. These domains are connected by the cytoplasmic loop between TM6 and TM7, as shown in Fig. 1.3, illustrating the first MATE family crystal structure, NorM_VC from *Vibrio cholerae*⁶⁷. The N and C-terminal domains are related by an intramolecular two-fold pseudo-symmetry along an axis perpendicular to the membrane, which is likely a result of gene duplication.

All available MATE crystal structures are captured in one conformational state only, namely an outward-facing conformation with a central binding vestibule accessible from the extracellular side and both termini, the N and C-terminus, are exposed to the cytoplasmic side. This conformation favors high affinity binding of monocations and low affinity binding for substrates. According to the basic principles of the alternating access model^{77,78}, also referred as a rocker-switch mechanism, MATE family transporters sequentially expose their substrate binding cavity to either side of the membrane during a transport cycle by a rigid-body movement of the N and C-domains. However, due to the lack of structural data of different intermediates corresponding to distinct steps underlying the MATE transport cycle, the mechanistic insights that govern transition between an outward and inward-facing state, remain

enigmatic. The inward-facing conformation has not been determined yet and it is still unclear how a MATE transporter switches to the inward-facing state in order to acquire a substrate from the cytoplasm. Comprehensive structural knowledge of the alternate conformations and the detailed molecular mechanism of the MATE transporter cycle are of pressing clinical relevance.

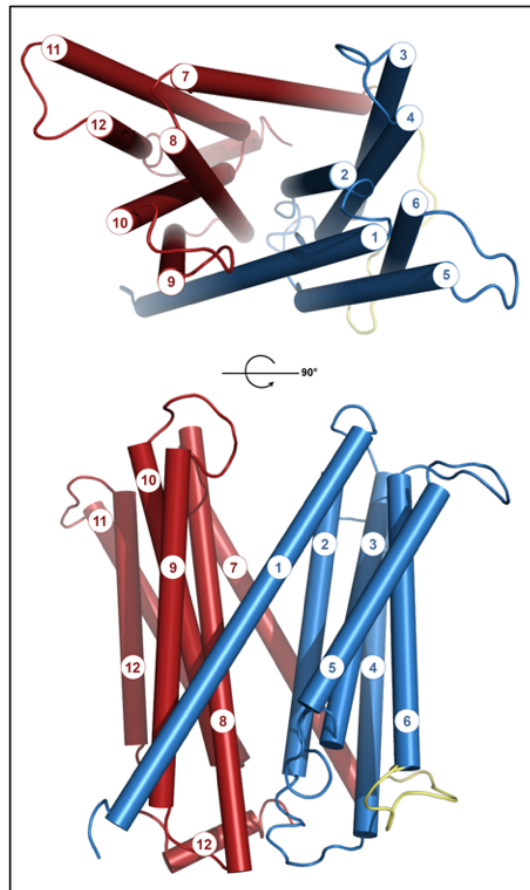


Fig. 1.3. Overall architecture of the MATE transporters. Schematic representation of the NorM_VC crystal structure, which shows the presence of 12 TMs divided into two domains, the N-domain (TMs 1-6, blue) and the C-domain (TMs 7-12, red). The cytoplasmic loop connecting these two domains between TM6 and TM7 is depicted in yellow.

Tab. 1.1. Summary of all available MATE family crystal structures. All structures from the NorM, DinF and eMATE subfamilies were obtained by X-ray crystallography in the presence or absence of different ligands and/or ions. Crystallization experiments were performed by vapor diffusion (VD) or lipidic cubic phase (LCP) method.

Protein	Source organism	Ligands/ions in the crystal structure	Resolution (Å)	Method	Space group	PDB ID	Ref.
NorM subfamily							
NorM_VC	<i>Vibrio cholerae</i>	-	3.65	VD	P 21 21 2	3MKT	67
		Rubidium ion (Rb)	4.2	VD	P 21 21 2	3MKU	67
NorM_NG	<i>Neisseria gonorrhoeae</i>	Tetraphenylphosphonium (TPP)	3.59	VD	P 32 2 1	4HUK	68
		Rhodamine 6G (R6G)	3.59	VD	P 32 2 1	4HUN	68
		Ethidium (ET)	3.49	VD	P 32 2 1	4HUM	68
		Cesium ion (Cs)	3.81	VD	P 32 2 1	4HUL	68
		Verapamil (4YH)	3.0	VD	P 32 2 1	5C6P	69
VcmN	<i>Vibrio cholerae</i>	1-Oleoyl-R-glycerol (OLC)	2.205	LCP	P 21 21 21	6IDP	70
		1-Oleoyl-R-glycerol (OLC)	2.502	LCP	P 21 21 21	6IDR	70
		1-Oleoyl-R-glycerol (OLC)	2.79	LCP	P 21 21 21	6IDS	70
DinF subfamily							
PfMATE	<i>Pyrococcus furiosus</i>	1-Oleoyl-R-glycerol (OLC)	2.398	LCP	C 1 2 1	3VVN	73
		1-Oleoyl-R-glycerol (OLC)	2.504	LCP	C 1 2 1	3VVO	73
		1-Oleoyl-R-glycerol (OLC) MaD3S	2.60	LCP	C 1 2 1	3VVS	73
		Br-Norfloxacin (BNU)	2.91	LCP	P 6	3VVP	73
		1-Oleoyl-R-glycerol (OLC) MaD5	3.0	LCP	C 1 2 1	3VVR	73
		1-Oleoyl-R-glycerol (OLC) MaL6	2.45	LCP	C 1 2 1	3WBN	73
		1-Oleoyl-R-glycerol (OLC)	2.096	LCP	C 1 2 1	3W4T	73
		PEG monodecyl ether (CXE) Chloride ion (Cl)	2.349	VD	C 1 2 1	4MLB	79
DinF-BH	<i>Bacillus halodurans</i>	-	3.20	VD	P 21 21 21	4LZ6	71
		Rhodamine 6G (R6G)	3.70	VD	P 21 21 21	4LZ9	71
		-	3.0	VD	P 21 21 21	5C6N	69
		Verapamil (4YH)	3.0	VD	P 21 21 21	5C6O	69
CibM	<i>Escherichia coli</i>	1-Oleoyl-R-glycerol (OLC) Cacodylate ion (CAC)	2.70	LCP	P 31 2 1	4Z3N	72
		Cacodylate ion (CAC) Rubidium ion (Rb)	3.30	LCP	P 31 2 1	4Z3P	72
Eukaryotic MATE subfamily							
CasMATE	<i>Camelina sativa</i>	1-Oleoyl-R-glycerol (OLC)	2.90	LCP	P 21 21 21	5XJJ	75
		1-Oleoyl-R-glycerol (OLC)	2.3	LCP	P 21 21 21	5YCK	75
AtDTX14	<i>Arabidopsis thaliana</i>	-	2.60	LCP	P 21 21 21	5Y50	76

1.2.3.1. NorM subfamily

In 2010, the first MATE crystal structure, that of NorM_VC from *Vibrio cholerae* (Fig. 1.4) was determined in the presence and absence of Rb⁺, referred as a Na⁺ congener⁸⁰. Afterwards, structures of NorM_NG from *Neisseria gonorrhoeae* (Fig. 1.5) have been solved, not only bound to Cs⁺ (another analog of sodium ion), but also in complex with three distinct substrates: ethidium, rhodamine 6G (R6G) and tetraphenylphosphonium (TPP). These structures illuminated for the first time interactions of a MATE transporter with its substrates, particularly between the negatively charged residues Asp41, Asp355 and Asp356 in NorM_NG and the partially-positive aromatic rings of the substrates.

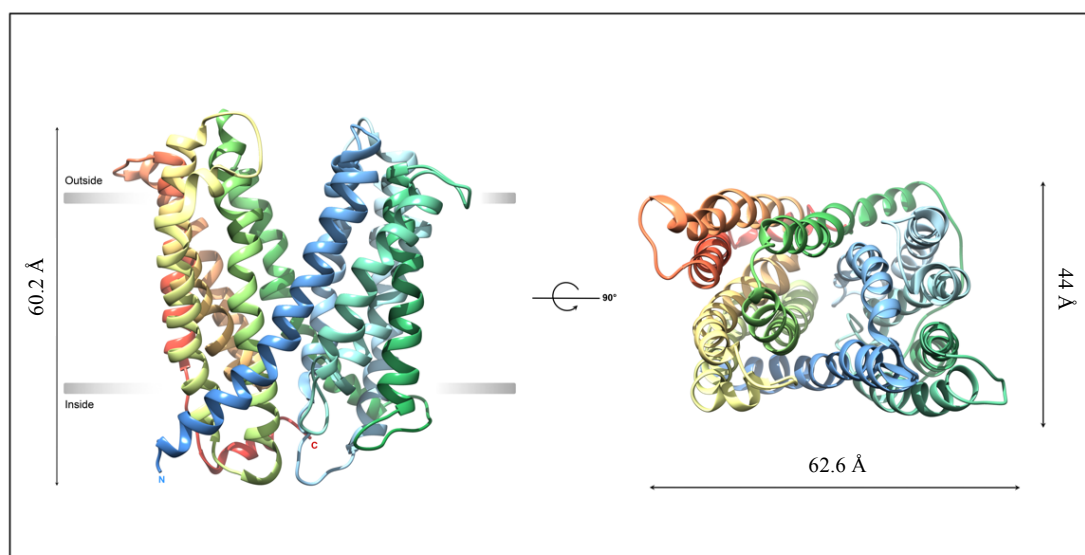


Fig. 1.4. Overall architecture (ribbon model) of the NorM_VC. The figure represents the side view and top view of NorM_VC.

A substrate-binding cavity has been identified near the membrane-periplasm interface between the N and C-domains. The most recently published structure of the NorM subfamily represents NorM_NG in complex with verapamil, an ion channel blocker and inhibitor of numerous multidrug transporters. All of the NorM subfamily structures adopt an outward-facing conformation, arranged in two bundles of six helices forming an internal cavity accessible from the extracellular side. The crystal structures of the Rb⁺ bound NorM_VC and the Cs⁺ bound

NorM_NG demonstrate a putative Na⁺ binding site located in the C-domain, where a monovalent cation is coordinated by glutamate (Glu261 in NorM_NG), tyrosine (Tyr367 in NorM_VC, Tyr294 in NorM_NG) and phenylalanine (Phe259 and Phe429 in NorM_VC). In case of NorM_NG, the three different substrates occupy a similar location in the main cavity and interact with the residues from both domains. Because the drug and inhibitor share a similar position in NorM_NG structures, the mechanism of action of verapamil is presumably based on preventing the substrate from binding by preoccupying its binding pocket. Although the bottom of the central cavity in NorM_NG is tightly closed from the cytoplasm, the drug binding site is partially sheltered from the extracellular side by two flexible loops, one connecting TM3 and TM4 and the second loop between TM9 and TM10. In case of the substrate-free NorM_VC structure, these loops move away from the drug binding cavity, indicating their putative involvement in the conformational change upon substrate binding/release. Additionally, a structural comparison between the substrate-free NorM_VC and the substrate-bound NorM_NG reveals that Phe259 (TM7), Gln278 and Ser285 (TM8) in NorM_VC corresponding to Phe265 (TM7), Gln284 and Ser288 (TM8) in NorM_NG are shifted away from the drug binding pocket in NorM_VC, leading to the speculation about drug release when TM7 and TM8 shift towards TM10.

The model of transport mechanism for the NorM proteins is based on proton or sodium binding to the conserved acidic residues in the C-domain, which in turn may induce a conformational transition of TM7 between a straight and bent state. In the light of the recent studies, which revealed dual ion coupling behavior of NorM_VC and the identification of a sodium binding site near Asp36 in TM1 of the N-domain, an alternate transport mechanism was proposed related to an ion induced structural changes within TM1.

Very recently, in February 2019, another crystal structure of VcmN MATE transporter from a pathogenic bacterium *Vibrio cholerae* was reported⁷⁰ (Fig. 1.6). The authors proposed a H⁺-

dependent conformational change, which induces TM1 bending upon protonation of the conserved Asp35 and in turn leads to rearrangement of the hydrogen bond network and collapse of the substrate-binding pocket, presumably releasing the bound substrate as the last step of the substrate extrusion mechanism. The proposed mechanism is based on the mutually exclusive events of the substrate and ion binding to this transporter.

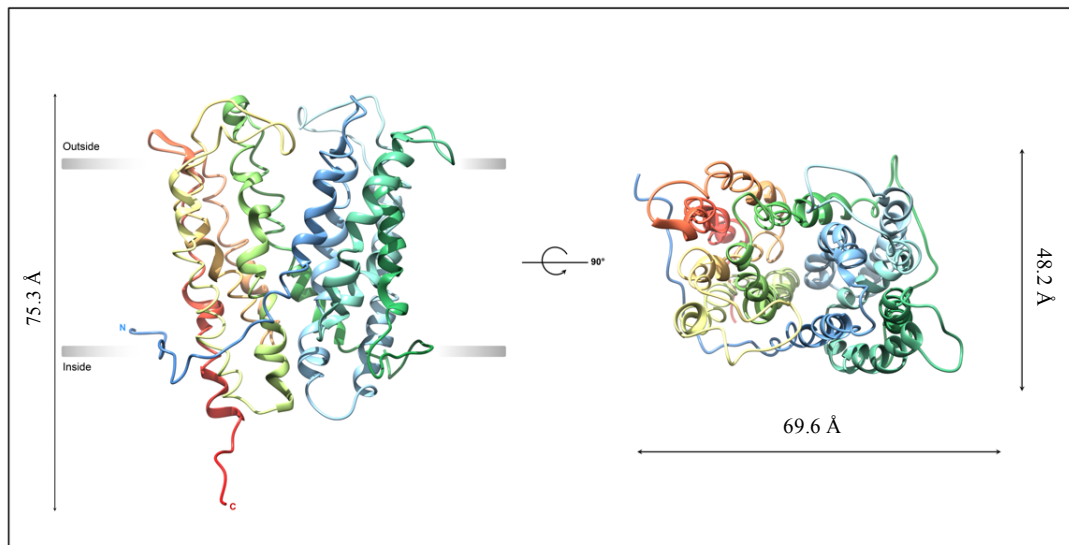


Fig. 1.5. Overall architecture (ribbon model) of the NorM_NG. The figure represents the side view and top view of NorM_NG.

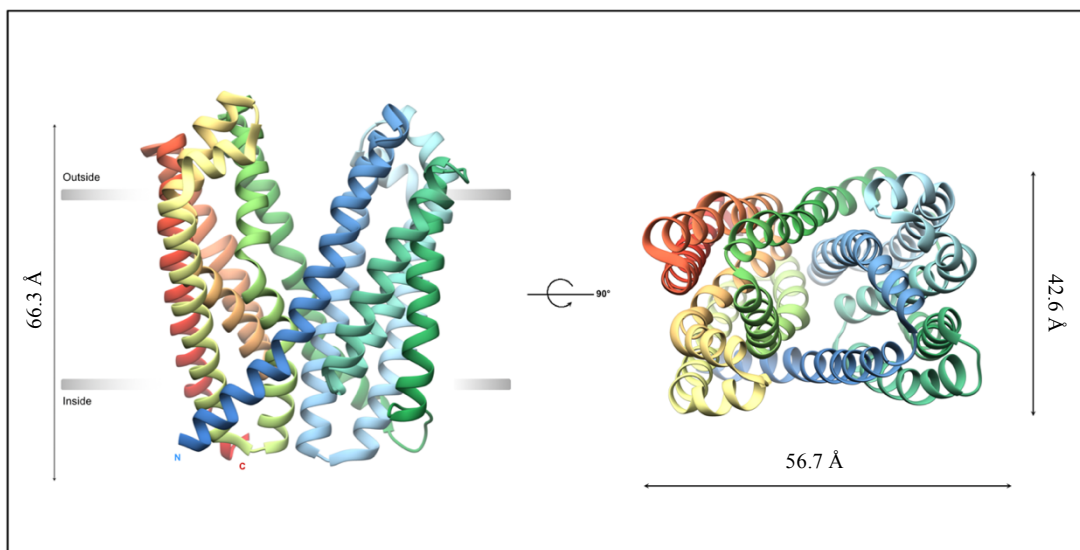


Fig. 1.6. Overall architecture (ribbon model) of the VcmN. The figure represents the side view and top view of VcmN.

1.2.3.2. DinF subfamily

In 2013, two crystal structures of the MATE transporters, initially assigned as H⁺-dependent transporters, were presented, PfMATE from *Pyrococcus furiosus* (Fig. 1.7) and DinF_BH from *Bacillus halodurans* (Fig. 1.8). A few years later, the crystal structure of the ClbM transporter from *Escherichia coli* (Fig. 1.9) was determined in the native and Rb⁺ bound states. DinF_BH has been also captured in complex with a substrate, rhodamine 6G (R6G), as well as with an inhibitor, verapamil, which is occupying the same binding site as the drug. Whereas the PfMATE structure was obtained in the native state in complex with a substrate, an analogue of norfloxacin (Br-NRF), as well as with three macrocyclic peptide inhibitors, MaL6, MaD5 and MaD3S. In contradiction to other known MATE transporters, DinF_BH reveals an asymmetric arrangement, where the pseudo-symmetry is broken by bending of helices TM7 and TM8 near their midsections. The transporter ClbM shares 27% sequence identity with PfMATE and 13% sequence identity with DinFBH. Despite the relatively high structural similarity within the N-domain to the H⁺-dependent PfMATE, ClbM was initially annotated as a Na⁺-coupled transporter based on the ethidium transport activity. ClbM is also responsible for transport of precolibactin, a drug precursor involved in the metabolism of the genotoxic molecule colibactin.

The structural comparison of the substrate-bound DinF subfamily structures revealed that PfMATE traps the substrate exclusively in the N-domain near the cation binding site, whereas DinF_BH acts as the prokaryotic NorM transporters, binding the substrates in the middle of the main vestibule formed by the N and C-domains (Fig. 1.10). Similarly to the NorM transporters, the Rb⁺ bound ClbM structure reveals the coordinating residues Tyr277 in TM7 and Phe292, Asp299 in TM8, which can also correspond to the alkali ion binding site, the ion being either Na⁺, K⁺ or Rb⁺.

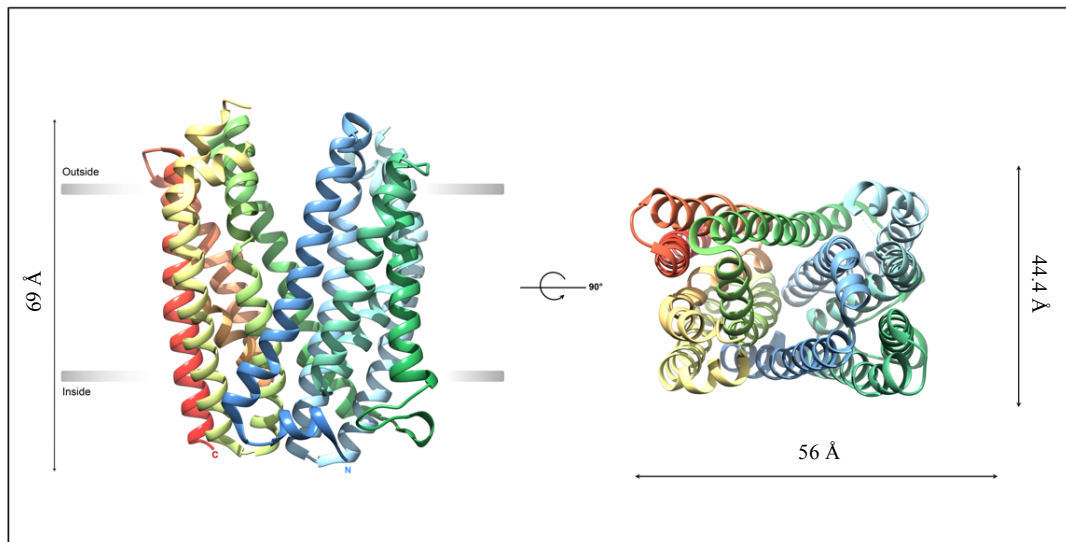


Fig. 1.7. Overall architecture (ribbon model) of the PfMATE. The figure represents the side view and top view of PfMATE.

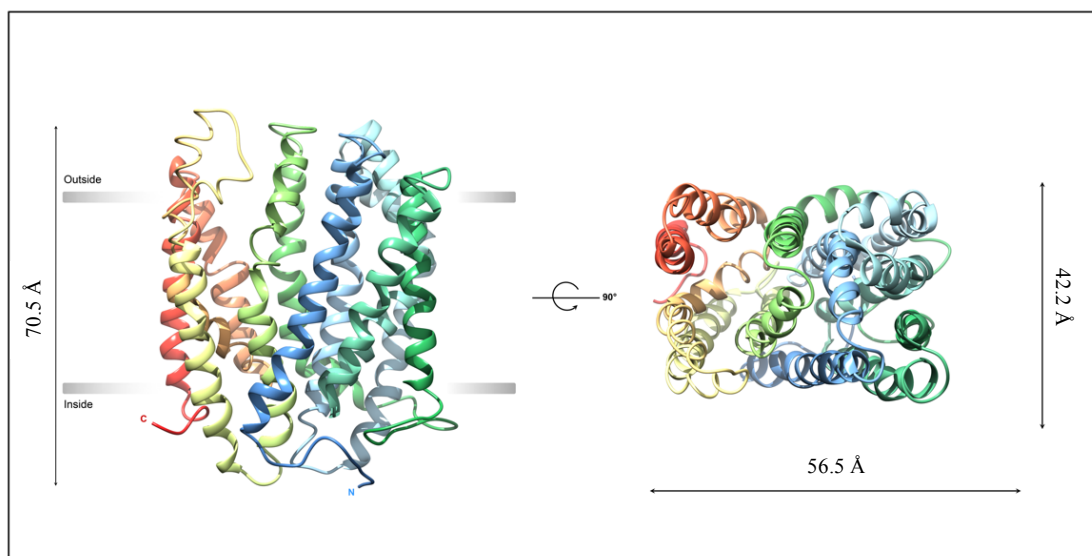


Fig. 8 Overall architecture (ribbon model) of the DinF_BH. The figure represents the side view and top view of DinF_BH.

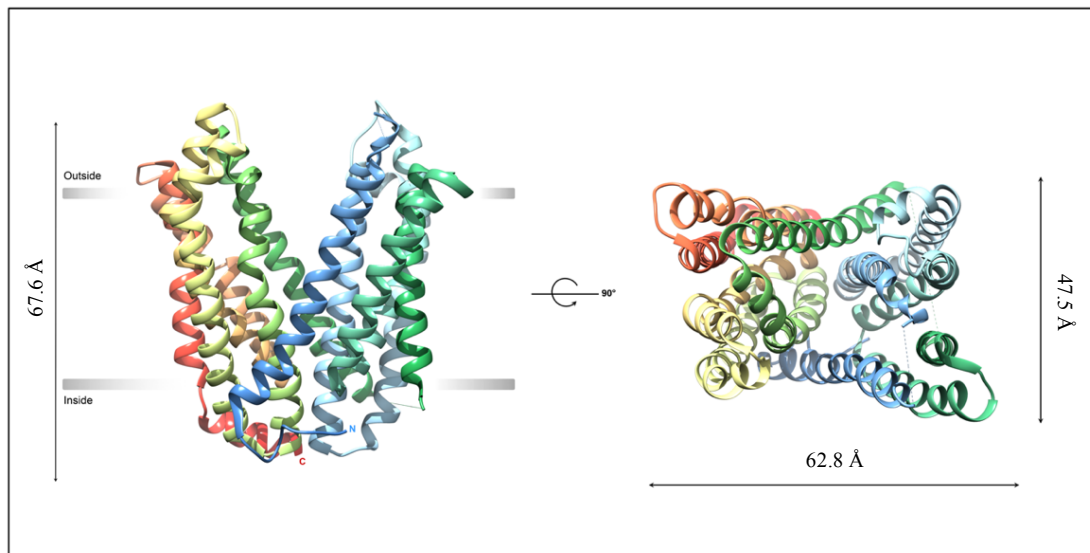


Fig. 9 Overall architecture (ribbon model) of the ClbM. The figure represents the side view and top view of ClbM.

The transport mechanism of the DinF subfamily is postulated to involve structural alterations within TM1 in the N-domain, however, TM1 bending is not commonly found in the other transporters. In case of ClbM, two alternate mechanisms were proposed. The first one is based on protonation/deprotonation of Asp53 (TM1) or Asp199 (TM8), which induces structural alterations of TM1, whereas the second one implies that after a conformational shift of TM7 and TM8 towards TM10, a cation binds to Tyr277 or Phe292, which leads to a structural rearrangement of TM1. DinF_BH and PfMATE also contain a conserved aspartate residue in TM1 corresponding to Asp53 in ClbM, namely Asp40 in DinF-BH and Asp41 in PfMATE, which will be described in more detail in the next paragraph.

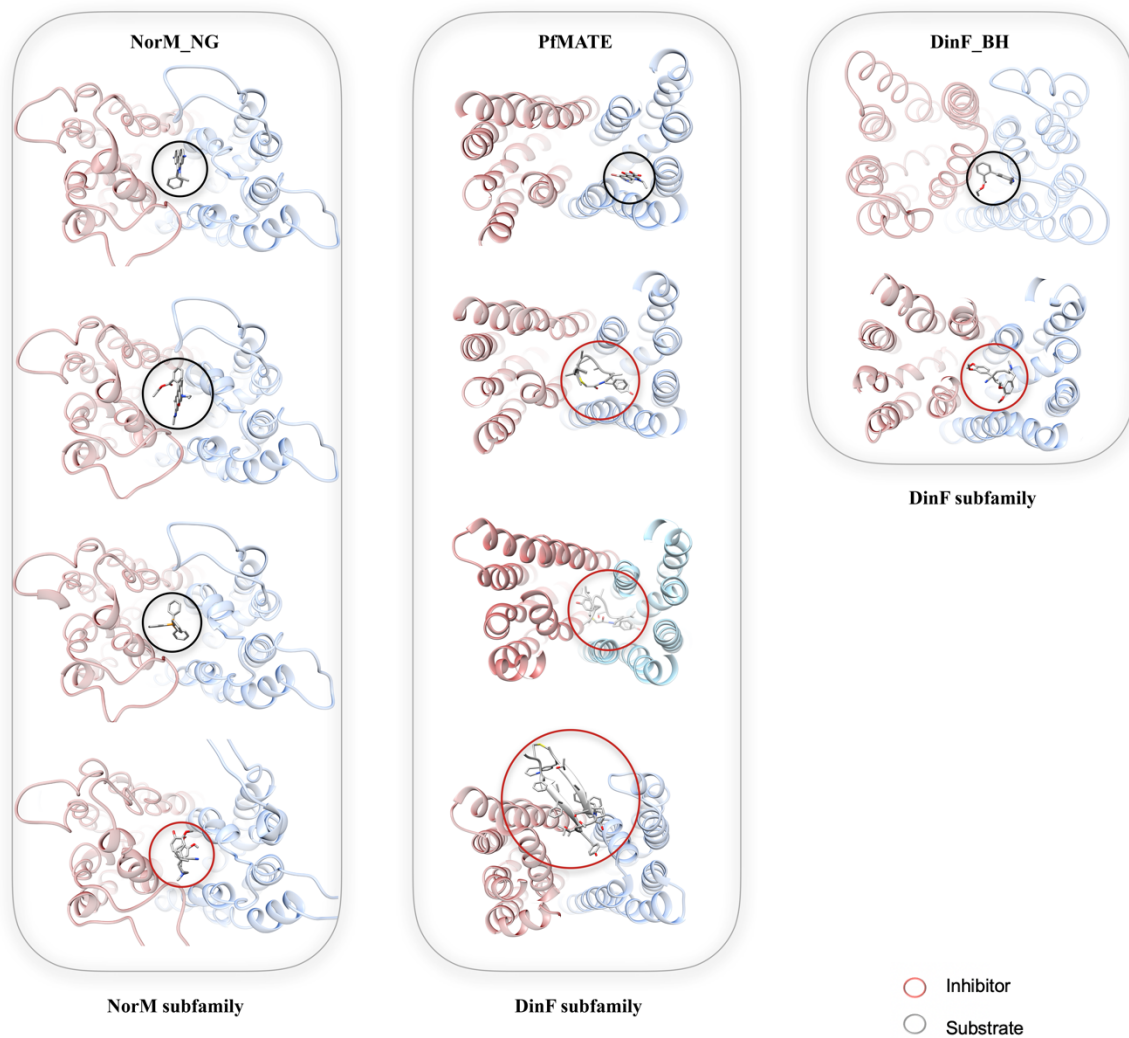


Fig. 1.10. Crystal structures of the MATE transporters in complex with substrates or inhibitors. Red circles indicate an inhibitor, whereas black circles illustrate a substrate of the corresponding MATE transporter.

1.2.3.2.1. The PfMATE transporter from *Pyrococcus furiosus*

Crystal structures of PfMATE have been previously determined in an outward-facing conformation. Two distinct states (bent and straight) were defined based on the arrangement of TM1, which is slightly bent in the low pH structure compared to the straight conformation in the high pH structure (Fig. 1.11). Based on this observation, it was concluded that the helix TM1 bending is correlated to the pH of crystallization media. Therefore, the authors proposed a PfMATE transport mechanism, in which TM1 undergoes a pH dependent conformational transition between the straight and bent state upon protonation of the largely conserved aspartate residue Asp41 in TM1. Coupling between H⁺ and substrate was suggested to be mediated by structural rearrangements of TM1 and to be allosteric in nature. However, this feature is not shared with the other transporters and there is no evidence for a proton dependent TM1 bending in case of DinF_BH. Instead, H⁺ and a substrate are postulated to compete for a shared binding site in DinF_BH and protonation of the conserved Asp40 in TM1 may induce a substrate release. Another critical difference refers to the drug binding site. In DinF_BH this site is formed by the residues from both, the N and C-domains, and exhibits a striking negatively charged electrostatic surface, being consistent with the preference for transport of cationic substrates. The inhibitor, verapamil, shares a similar mode of interaction with DinF_BH and occupies the same binding site as the substrate.

Notably, the drug binding site observed in PfMATE is formed exclusively within the N-domain cavity located halfway in the membrane, and it is dominated by polar amino acids forming a positively charged vestibule. The binding site of the large thioether-cyclic peptide inhibitors, MaD5 and MaD3S, is located in the N-domain, whereas MaL6 occupies the central region between two domains of PfMATE (Fig. 1.10). Based on the fluorescence-based assay using *E. coli* cells and spheroplasts lacking the outer membrane, norfloxacin and ethidium bromide were

identified as substrates for this transporter and MaD5 exhibited the largest inhibitory activity against extrusion of ethidium bromide.

Taken together, the proposed PfMATE transport mechanism involves a deprotonated and substrate bound state in the outward-facing conformation, which undergoes a structural rearrangement within TM1 upon protonation of Asp41. In consequence, bending of TM1 at Pro26 induces extrusion of the substrate from the N-domain cavity into the extracellular environment. Finally, a protonated and substrate-free PfMATE converts into the inward-facing state to acquire the drug from the cytoplasmic space.

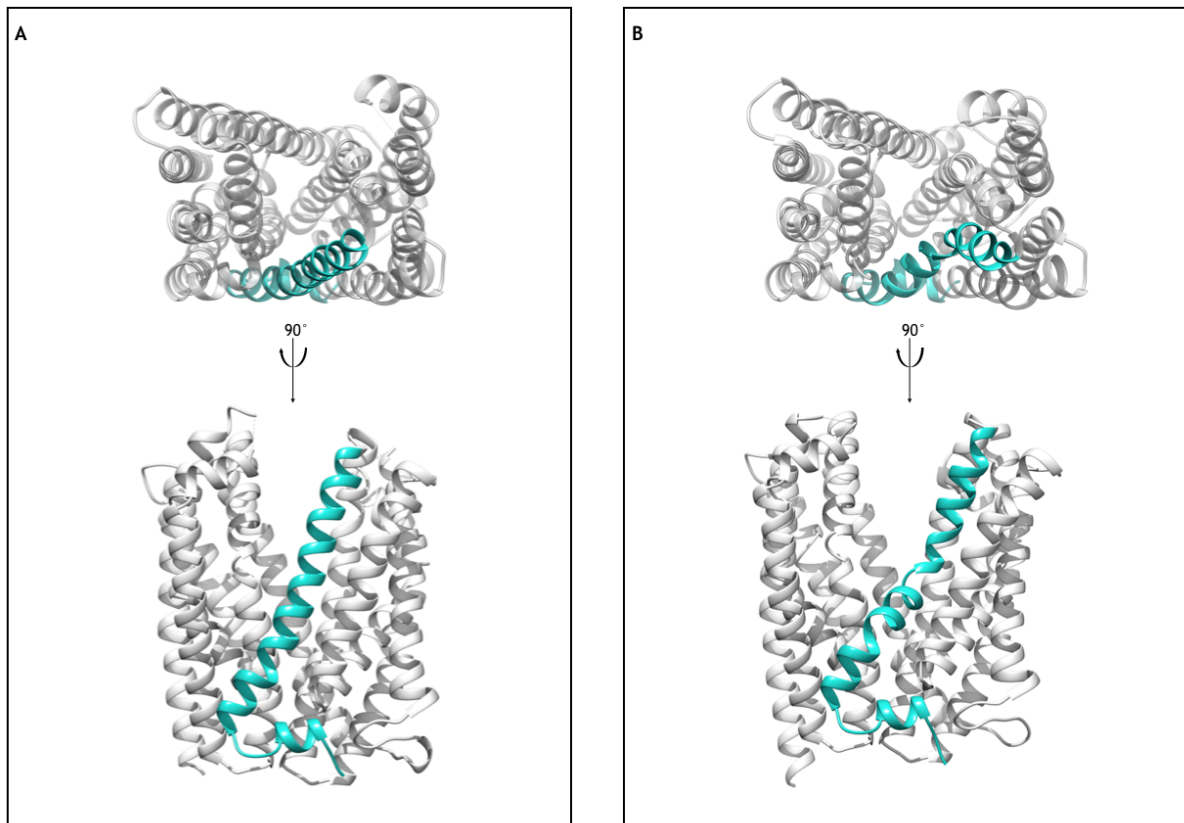


Fig. 1.11. Overall architecture (ribbon model) of PfMATE in the straight and bent conformations. The figure represents the side view and top view of the straight (A) and bent (B) conformation. TM1 is highlighted in blue, whereas the remaining helices are colored in grey.

1.2.3.3. Eukaryotic MATE subfamily

The most recently published crystallographic structures of the MATE family concern the eMATE subfamily members from plants, AtDTX14 from *Arabidopsis thaliana* (Fig. 1.12) and CasMATE from *Camelina sativa* (Fig. 1.13). Similar to the other structures of MATE family members, the internal cavity is accessible from the extracellular space and the overall arrangement resembles an outward-facing state. Based on the amino acid sequence alignment, eMATE and NorM proteins contain the conserved acidic residues in the C-domain (Glu265 and Asp383 in AtDTX14 corresponding to Glu255 and Asp371 in NorM_VC) being critical for the transport activity.

Eukaryotic MATE proteins contain a cation-binding motif, which is also identified in NorM_NG. Although acidic residues are observed in the N-domains of NorM and DinF, they are not conserved in case of the eMATEs. These observations may indicate some similarity for the transport mechanism with the NorM and differences with the DinF transporters. Despite an amino acid sequence similarity between the NorM and eMATE subfamily, there are structural differences, especially within TM7. This helix is significantly bent in the AtDTX14 crystal structure, presumably due to the protonation of the conserved residues. CasMATE reveals the central pocket shaped in an asymmetrical manner, with the N-domain side being narrower than the C-domain side, where the latter one is proposed to be the substrate binding site. Moreover, the extracellular entrance of AtDTX14 is narrower than those of all bacterial structures, thus possibly representing a partially occluded outward-facing structure. The C-domain cavity is formed between TM7, TM8, and TM10, and is suitable for the recognition of the positively charged substrates. The N-domain of the eMATE transporters could play a different role compared to the prokaryotic MATEs.

A structural comparison between AtDTX14 and the NorM proteins leads to a plausible hypothesis about the transport mechanism for the eMATE subfamily, where the protonation of

the conserved acidic residue in the C-domain (Glu265 or Asp383 in AtDTX14) induces the conformational change within TM7.

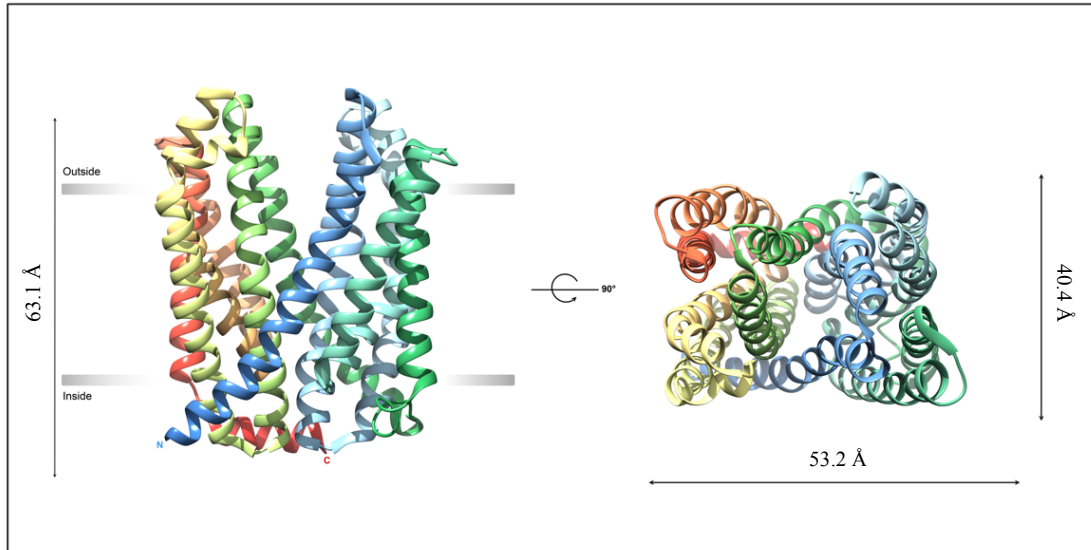


Fig. 1.12. Overall architecture (ribbon model) of the AtDTX14. The figure represents the side view and top view of AtDTX14.

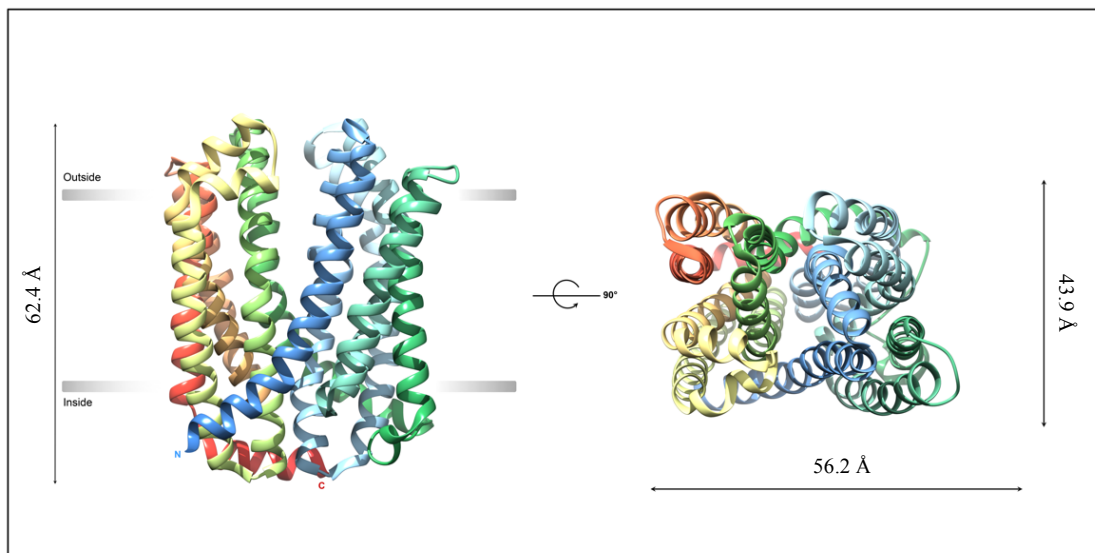


Fig. 1.13. Overall architecture (ribbon model) of the CasMATE. The figure represents the side view and top view of CasMATE.

1.3. *Pyrococcus furiosus* - an anaerobic hyperthermophile

The PfMATE source organism, *Pyrococcus furiosus*, is a strictly anaerobic and hyperthermophilic archaeon from the phylum *Euryarchaeota*, class **Thermococci** and order *Thermococcales*. Its genome size is around 1.9 Mb and contains 2192 open reading frames (ORFs)⁸¹. The name *Pyrococcus furiosus*, meaning a ‘rushing fireball’ in Latin, refers to its high motility and a doubling time of around 37 min under its optimal growth conditions. It was isolated from the shallow marine sediments of Vulcano Island in Italy and characterized for the first time in 1986 by G. Fiala and K. O. Stetter from the University of Regensburg⁸². This microorganism exhibits an ability to grow at the elevated temperatures ranging from 70 to 103 °C, with an optimum temperature of 95-100 °C, and at the pH range 5-9, with an optimum pH of 7. In contrast to many other hyperthermophiles, *P. furiosus* prefers sugars, α -glycosidic carbohydrates, like starch and maltose or cellobiose, laminarin or even chitin over amino acids and proteins⁸³. Efficient growth on proteins is possible in presence of elemental sulfur (S^0) as an electron acceptor, producing hydrogen sulfide (H_2S), but growth on oligosaccharides can be carried out without S^0 . Starch-derived sugars are fermented to organic acids like acetate, CO_2 and H_2 as metabolic end products.

P. furiosus exhibits one of the simplest forms of a respiratory system based on a membrane-bound hydrogenase (MBH) complex, using protons as terminal electron acceptors^{84,85}. Energy generated from the proton reduction to hydrogen gas (H_2) is utilized to create an electrochemical proton gradient across the membrane, which is directly coupled to ATP synthesis. Besides MBH, *P. furiosus* harbors two cytoplasmic hydrogenases (SHI and SHII), which are able to recycle H_2 in order to provide NADPH for biosynthetic purposes^{84,86}. Another interesting feature of *P. furiosus* is the presence of enzymes dependent on tungsten (W), which is rarely used in biological systems⁸⁷.

This archaeon is also considered to be a source of highly thermostable enzymes. Due to their stability and activity under harsh conditions, they are valuable tools in molecular biology and biotechnology. The most popular example of such an enzyme might be DNA polymerase I commonly used in the polymerase chain reaction (PCR)⁸⁸. DNA or RNA ligases as well as β -glucosidase and α -amylase isolated from *P. furiosus* are essential in numerous industrial sectors.

After 2010, when a functional plasmid-based transformation system for this hyperthermophile was developed, genetic modifications including gene expression in *P. furiosus* became possible⁸⁹. From the morphological point of view, the cells appear as regular to slightly irregular cocci of 0.8 μm to 2.5 μm diameter with numerous, up to 50, monopolar flagella. In contrast to bacteria, the cellular envelope of this archaeon lacks peptidoglycan and consists of a single cellular membrane as well as a paracrystalline surface layer (S-layer), composed of a glycoprotein (Fig. 1.14). Another feature distinguishing them from their bacterial counterparts is the membrane lipid composition.

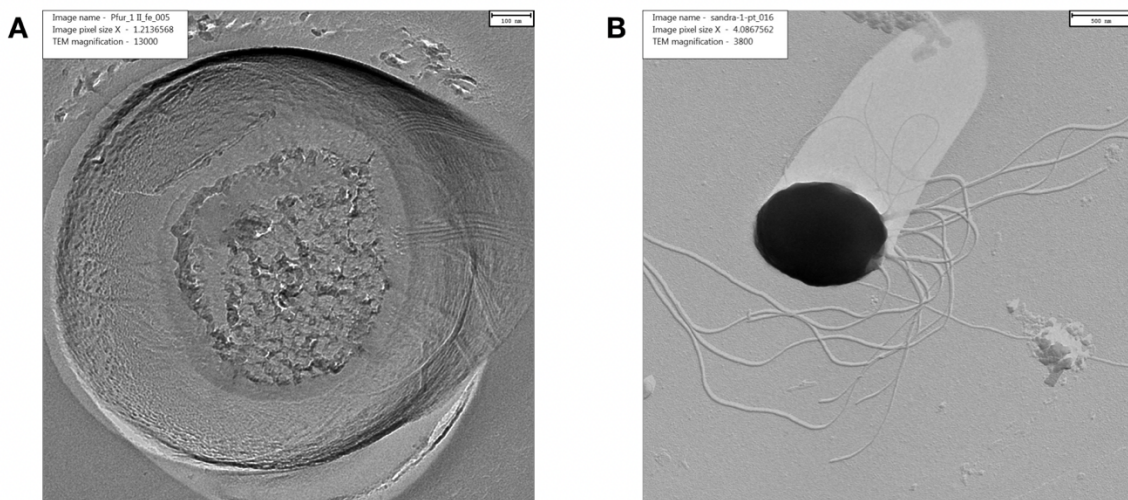


Fig. 1.14. Transmission electron micrographs (TEM) of *P. furiosus* cells. (A) Micrograph of a freeze-etched slice through a frozen-hydrated *P. furiosus* cell exposing the S-layer. (B) Micrograph of a coccoid cell of *P. furiosus* with multiple flagella at one pole of the cell. The *P. furiosus* WT cells samples were prepared by myself and the TEM micrographs were kindly acquired with assistance from Prof. Dr. Reinhard Rachel from the University of Regensburg.

1.4. Archaeal lipids

Cell membranes diversified the chemical structure of their lipids during the evolution of life. *P. furiosus* as a hyperthermophilic archaeon has evolved certain molecular mechanisms to cope with the high temperature, hydrostatic pressure and lack of oxygen⁹⁰. One of the essential thermal adaptations to allow surviving at such extreme conditions is an unusual membrane lipid composition^{91,92}. Archaeal membrane lipids are distinguished from those found in eukaryotic and bacterial cells in four major ways (Fig. 1.15).

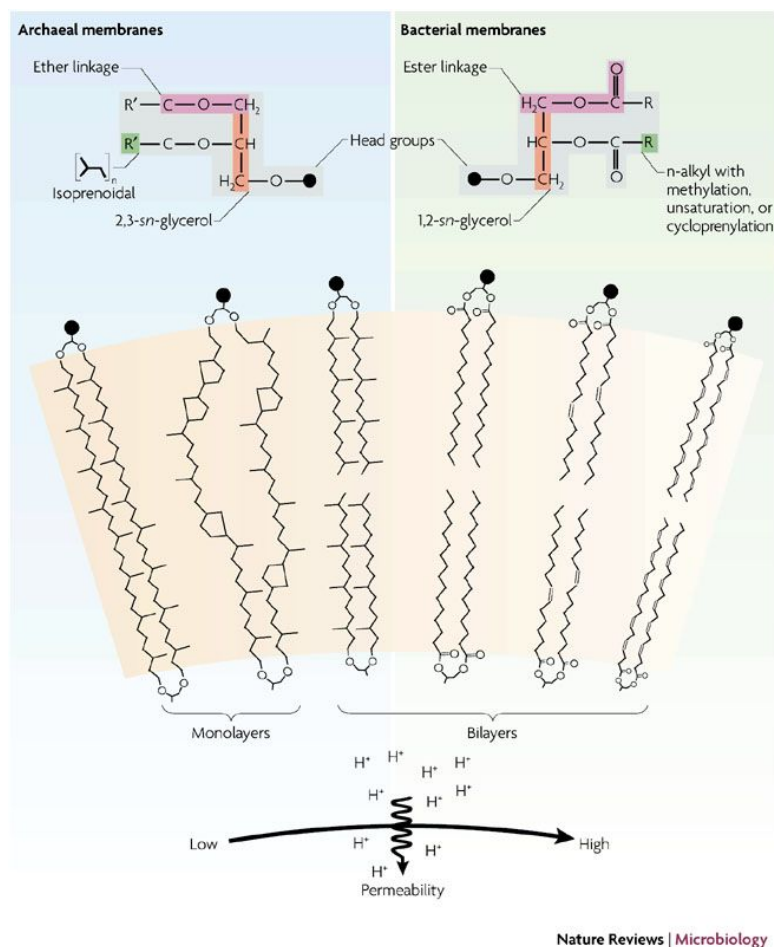


Fig. 1.15. Comparison of bacterial and archaeal lipids. The basic chemical structures of archaeal (left) and bacterial (right) membrane lipids are shown to illustrate the typical chemical differences. Examples of intact membrane structures used by archaea (left) and bacteria (right) are shown below, including the monolayers that are produced by some archaea and the highly unsaturated membranes present in some bacteria. The arrow (bottom) indicates a general trend of increasing permeability to ions such as protons and sodium. Eukaryotic membranes are similar to those of bacteria. Picture taken from ⁹⁰.

One of the fundamental differences between the archaeal and bacterial lipid composition concerns the linkage between the glycerol backbone and the hydrophobic lipid tail. In archaeal lipids these moieties are connected by an ether linkage rather than an ester linkage present in their bacterial counterparts⁹³. Additionally, the archaeal lipid tail comprises branched alkyl chains of 20-40 carbon atoms, built up from usually saturated repeats of isoprenoid units, which are more resistant to hydrolysis and oxidation. In contrast, the side chains of bacterial lipids are composed of fatty acids, usually consisting of 16-18 carbon atoms.

Another crucial distinction between these two life domains is related to the stereochemistry and chirality of the glycerol moiety. Isoprenoid chains are linked by an ether linkage to the *sn*-glycerol-1-phosphate (G-1-P) backbone with *sn*-2 and *sn*-3 configuration, forming a diether (archaeol)⁹⁴, whereas the bacterial lipids possess a backbone moiety composed of two fatty acid hydrocarbon chains esterified to the *sn*-1 and *sn*-2 positions of glycerol-3-phosphate (G-3-P).

Finally, archaeal lipids can form not only bilayers of diethers but also monolayers composed of tetraethers (caldarchaeol), spanning the whole membrane. Due to a restricted motility of the hydrocarbon chains they are highly stable and are less permeable to ions. The differences between the aforementioned lipids are depicted in Fig. 1.15. Polar head groups are similar in archaea and bacteria and consist of mixtures of glyco groups (mainly disaccharides), and/or phosphogroups, primarily phosphoglycerol, phosphoserine, phosphoethanolamine or phosphoinositol. Phosphocholine headgroups are rarely found. Recent studies have revealed a certain degree of overlap, for example some bacteria can harbor membrane spanning lipids, which can be considered as a form of homeoviscous adaptation and regulation of membrane fluidity.

Although ether linkages have been found to be a dominant feature of the archaeal membranes, they can also be observed in their bacterial counterparts, and isoprenoid units are present in all three domains of life.

Analysis of the *P. furiosus* lipidome reveals that the vast majority 80-90% of the total lipid extract comprises diethers, whereas the remaining fraction consists of tetraethers, spanning the whole membrane. The polar lipid species identified by the TLC and MALDI-TOF/MS analyses are diphytanylglycerol analogues (simply called diethers) of phosphatidylinositol (PI) and phosphatidylglycerol (PG), *N*-acetylglucosaminediphytanylglycerol phosphate (*N*-Acetyl-hexose-P-D), diether analogue of phosphatidic acid (PA), monoglycosyl diether (MGD), diglycosyl diether (DGD) and tetraethers: (Hexose₂-PG-T) and (*N*-Acetyl-hexose-hexose-P-T-cyclic)⁹⁵. The chemical structures of the aforementioned lipids are presented in Fig. 1.16.

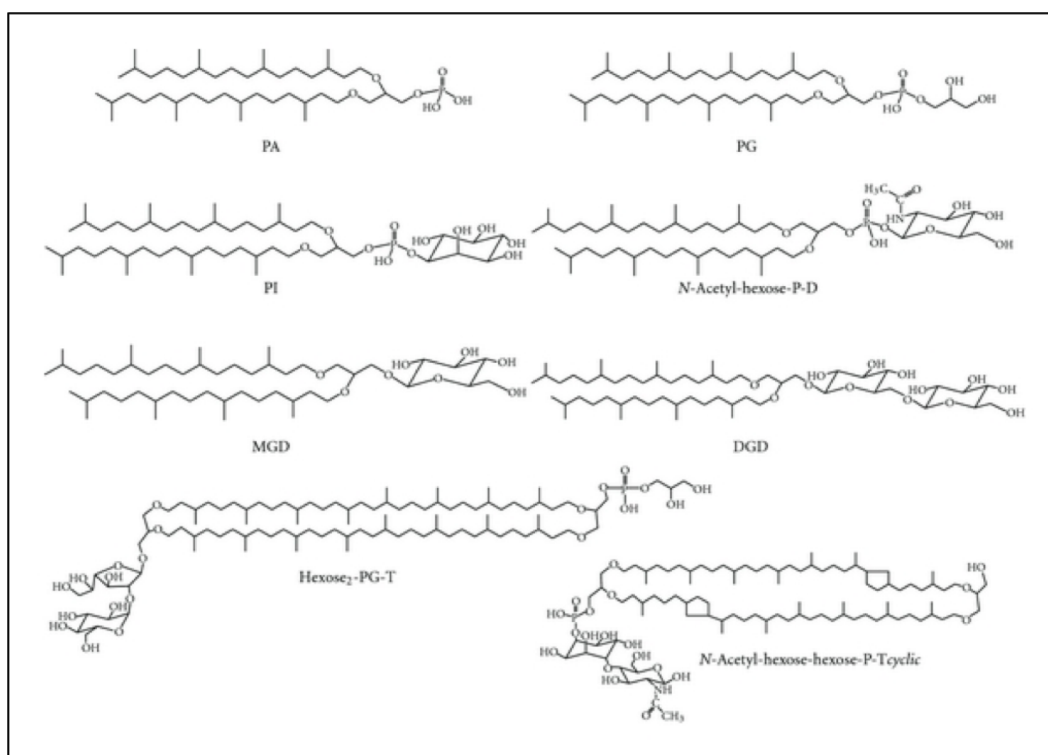


Fig. 1.16. Lipid composition of *P. furiosus* membrane. Chemical structures of lipid species identified in *P. furiosus* total lipid extract by MS/HPLC method. Figure adapted from ⁹⁵.

1.5. Structural biology of membrane proteins

Membrane proteins are amphipathic molecules, containing hydrophobic and polar regions. Owing to their chemical nature, flexibility as well as bottlenecks associated with their production and purification in sufficient quantities, membrane proteins comprise a challenging target for structural biology, especially for X-ray crystallography. A prerequisite to generate crystals diffracting to high resolution is sufficient crystal contacts to form well-ordered packing. These hurdles can be tackled by multiple strategies, including mutagenesis of the protein to increase its stability, generation of chimeras or co-crystallization with antibody fragments^{96,97} or nanobodies⁹⁸, which may increase crystal contacts or capture alternate conformational states.

The landmark in the field of membrane protein X-ray crystallography was in 1985, when J. Deisenhofer, O. Epp, K. Miki, R. Huber and H. Michel published the first high-resolution (3 Å) crystal structure of an integral membrane protein, namely the photosynthetic reaction center from the purple bacterium *Rhodospseudomonas viridis*⁹⁹. The Nobel Prize in Chemistry in 1988 was awarded jointly to H. Michel, R. Huber and J. Deisenhofer for this discovery.

The electron microscopic analysis of purple membranes from *Halobacterium halobium* by R. Henderson and P. N. T. Unwin had provided the first structural identification of membrane spanning helices in a membrane protein, namely of bacteriorhodopsin, the protein constituent of purple membrane (determined up to 7 Å in 1975)¹⁰⁰. The X-ray crystal structure of this protein, reported in 1997¹⁰¹, was the first example of a successful crystallization of a membrane protein using lipidic cubic phase (LCP), a phase with a toothpaste-like consistency. The LCP method was developed by E. M. Landau and J. P. Rosenbusch¹⁰² and optimized in the following years by M. Caffrey and V. Cherezov¹⁰³. Since then, the LCP technique has been commonly applied to determine the structures of multiple membrane proteins^{104–106}, including G protein-coupled receptors (GPCRs)^{107–110}.

The cumulative number of the integral membrane protein structures deposited in the Protein Data Bank has been increasing over the past decade (Fig. 1.17)¹¹¹. The vast majority has been elucidated by X-ray crystallography, whereas the remaining ones have been contributed by NMR spectroscopy and by cryoEM. K. Wüthrich was awarded the Nobel Prize in Chemistry in 2002 for his development in the field of biological NMR spectroscopy¹¹², whereas R. Henderson, J. Frank and J. Dubochet received the Nobel Prize in Chemistry in 2017 for their contributions to the development of single particle cryoEM¹¹³.

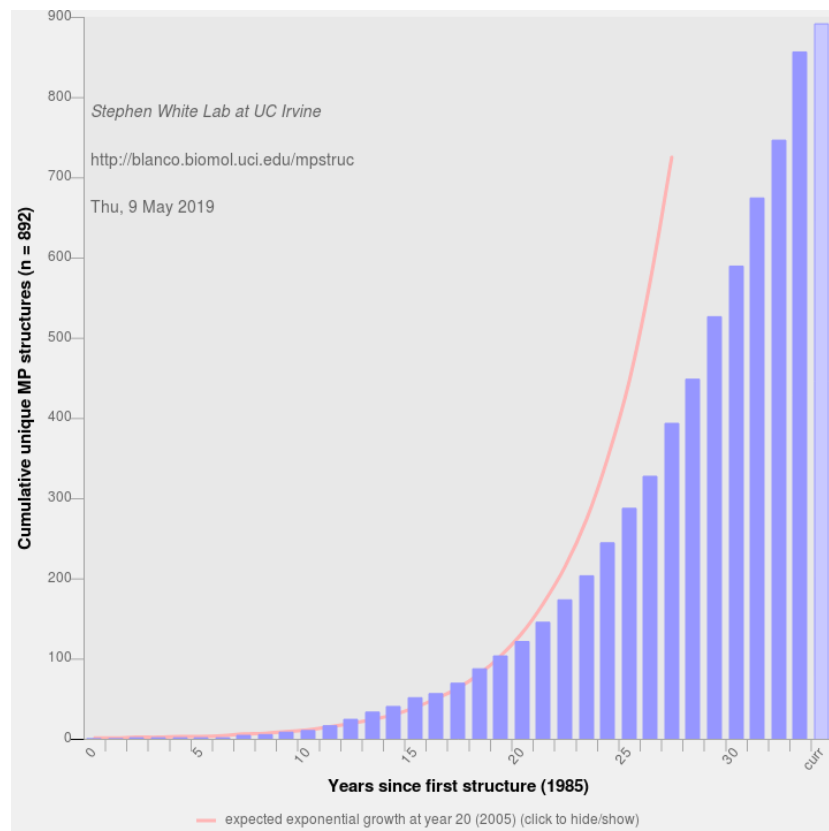


Fig. 1.17. Progress of membrane protein structure determination.

X-ray crystallography is still the only method capable of providing very high-resolution data and detailed insights into how ligands, cofactors, and ions interact with proteins. Most of the structural studies have been carried out using crystals obtained by vapor diffusion with protein-detergent complexes (*in surfo* method), however, in recent years there is a significant increase

in the number of structures that have been solved using crystals obtained in lipidic mesophase (*in meso* method) (Fig. 1.18).

Among all published MATE crystal structures, the most recently reported member of the NorM subfamily (VcmN), two representatives of the DinF subfamily (CibM and PfMATE) as well as all eukaryotic MATE transporters, AtDTX14 and CasMATE, were determined using crystals obtained by the LCP method.

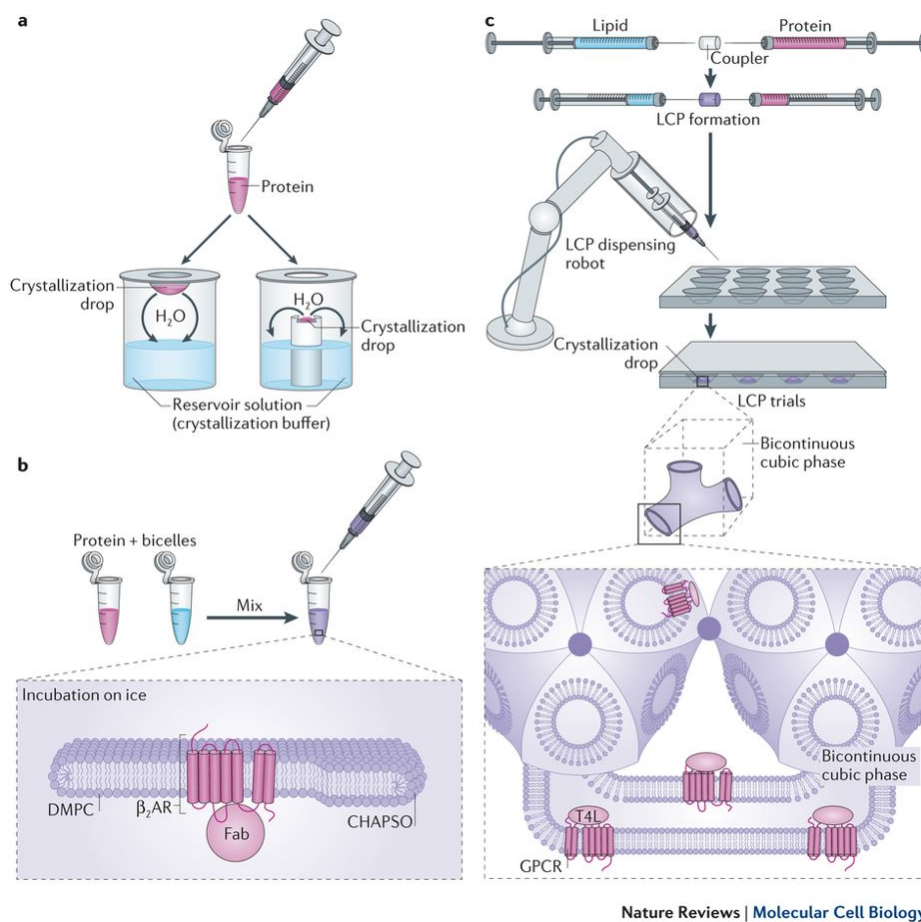


Fig. 1.18. Crystallization methods of membrane proteins. Schematic representation of (A) vapor diffusion (*in surfo*), (B) bicelle and (C) LCP (*in meso*) methods. (A) The protein is mixed with the crystallization buffer and set up as a hanging drop (left) or sitting drop (right). (B) The protein is mixed with bicelles, usually composed of dimyristoyl phosphatidylcholine (and then crystallized by the vapor diffusion approach). (C) The protein is mixed with the exogenous lipid, typically monoolein and dispensed on the crystallization plate. Picture adapted from Ghosh *et al.*¹¹⁴

1.6. Motivation and aim of this work

Although structural and functional data of several MATE transporters are currently available, there are still many open questions to fully understand their mechanism of action at the atomic level. Since all crystal structures show the transporters only in an outward-facing conformation, the detailed transport mechanism of drugs and metabolites remains poorly understood. Despite resemblance in the overall architecture of all MATE transporter structures, members of different MATE subfamilies are involved in the efflux of various structurally and chemically dissimilar substrates. In order to elucidate the underlying steps of the MATE transport cycle, structures of distinct intermediates are required. Particularly, structural data of an inward-facing conformation would provide crucial insights into a substrate high affinity state, which is physiologically relevant and important for rational drug design. It would also deliver fundamental knowledge on how the MATE transporters follow the originally proposed concept of alternating access^{77,115}. The main objective of this work is structural and functional characterization of the DinF subfamily member PfMATE, (UniProtKB code: Q8U2X0), which is one of the four annotated MATE transporters of *Pyrococcus furiosus*.

Upon the beginning of my PhD studies, PfMATE had been studied by members of our department. This protein had been produced heterologously in *Escherichia coli* and successfully crystallized by Cornelia Münke, which led to the structure determination at the high resolution of 2.35 Å (PDB ID code: 4MLB) achieved by Tsuyoshi Nonaka. The overall structure of PfMATE is very similar to that of the NorM subfamily members available at that time. However, the substrate-binding cavity of PfMATE exhibits a strongly positive electrostatic surface potential in contrast to the negatively charged ones of the NorM transporters. Therefore, it was concluded that this cavity may not be suited for accommodation of the typical cationic substrates of the classical MATE family. Indeed, a PhD student, Viveka Nand Malviya, who worked on the functional aspects of PfMATE at that time, could not

identify any typical cationic substrate for this protein based on the MIC (Minimum Inhibitory Concentration) assays in *E. coli* and DSC (Differential Scanning Calorimetry) experiments. Due to the positively charged cavity of PfMATE and the similarities with the homology modeled Wzx lipid flippase, it was initially speculated that PfMATE might be a transporter of negatively charged lipophilic substrates, in particular of lipids. In 2013, during preparation of the manuscript by our group, the crystal structure of PfMATE was published by Tanaka *et al.* The authors presented functional data that PfMATE is a classical MATE transporter causing multidrug resistance in *E. coli* against ethidium bromide and norfloxacin. Tanaka *et al.*⁷³ presented two outward-facing structures of PfMATE, which differ in the conformation of TM1 (straight and bent). A pH dependent conformational transition of TM1 regulated by the protonation state of the conserved aspartate residue Asp41 was proposed, which, however, has been discussed controversially. First of all, recent results of computational investigations indicate a Na⁺ binding site located near Asp41. Furthermore, there is no structural evidence for a pH dependent conformational change in case of the DinF_BH transporter from the same DinF subfamily⁷¹. Divergent methods applied for crystallization of PfMATE (LCP) and DinF-BH (vapor diffusion), raised the hypothesis about TM1 bending to be rather affected by interactions with exogenous lipids (monoolein) present under the crystallization conditions¹¹⁶.

Based on these open questions, an experimental approach to investigate the role of lipids as structural and functional modulators of PfMATE has been taken in the course of my PhD project. As reported previously, native lipid environment as well as the physiological conditions might be significant for the structure and function of membrane proteins, especially from hyperthermophilic archaeal organisms¹¹⁷. Therefore, the initial stage of my PhD project was mainly focused on the homologous expression of the gene coding for PfMATE in *P. furiosus* under anaerobic conditions along with the elevated temperature in order to provide a native environment. Subsequently, the homologously produced protein was subjected for purification

and crystallization trials. The low molecular weight of PfMATE, around 50 kDa, is currently a limiting factor for the cryoEM studies. Therefore, X-ray crystallography has been a method of choice for the structure determination. Due to the low amount of the purified protein sample for extensive crystallization experiments, another strategy was developed in parallel to work on the protein produced heterologously in *E. coli*. In order to investigate the effects of lipids on PfMATE, endogenous lipids extracted from the native organism *P. furiosus* were used for co-purification and co-crystallization trials. This approach led to the structure determination in a novel conformational state for the MATE family transporters. In order to assess molecular details of lipid access pathways, molecular dynamics (MD) simulations were also performed. Furthermore, combination of the crystallographic data from the single-wavelength anomalous dispersion (SAD) experiments using anomalous dispersion of Cs⁺ performed previously by Tsuyoshi Nonaka with the results of the MD simulations included in my thesis shed light on a sodium ion binding site at Asp41. Based on these data, a framework of the structural changes that occur during transition between the inward and outward-facing conformation of the MATE transporter could be established.

2. MATERIALS AND METHODS

2.1. Prokaryotic strains and plasmids

Table 2.1. Prokaryotic strains and plasmids.

Prokaryotic strain	Application
<i>Escherichia coli</i> NEB 5-alpha	Cloning
<i>Escherichia coli</i> TOP10	Heterologous gene expression
<i>Pyrococcus furiosus</i> WT	Genomic DNA and lipid extraction
<i>Pyrococcus furiosus</i> Δ1623	Homologous gene expression
Plasmid	Application
pYS14	Cloning and homologous gene expression
pYS17	Amplification of the <i>pf0613</i> promoter
pBAD- <i>PfMATE</i>	Heterologous gene expression

2.2. Extraction of genomic DNA from *P. furiosus*

0.1 g *P. furiosus* WT cell pellet was resuspended in 567 µl TE buffer (10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA). Cells were lysed by addition of 28 µl 10% (w/v) SDS, 6 µl proteinase K (20 mg/ml) and 12 µl RNase (10 mg/ml). After incubation at 60 °C for 15 min, 37 µl 5 M NaCl was added. Due to the usage of toxic volatile compounds, the following steps were performed in a fume hood. 700 µl of the organic solvent mixture ((25:24:1, v/v) phenol/chloroform/isoamyl alcohol) was added. The sample was vortexed and centrifuged at 14,000 rpm at 4 °C for 5 min. The upper aqueous phase containing the DNA was carefully removed and transferred into a new tube. After addition of 700 µl chloroform the sample was thoroughly mixed and centrifuged at 14,000 rpm at 4 °C for 5 min. The upper phase was transferred into a new tube. 0.6 volumes of isopropanol were added and mixed gently by tube inversion. After addition of 1 ml of 70% ethanol, the sample was vortexed and centrifuged at 14,000 rpm at 4 °C for 2 min. The supernatant was discarded, the DNA pellet was air-dried, resuspended in 50 µl 10 mM Tris-HCl (pH 8.5) and subsequently stored at 4°C.

2.3. Determination of DNA concentrations

Qualitative and quantitative analysis of genomic and plasmid DNA was performed photometrically by absorbance measurements at 260 nm using a NanoDrop spectrophotometer (ThermoFisher Scientific). 10 mM Tris-HCl (pH 8.5) served as a blank as it was used for isolation of plasmid and genomic DNA. The purity of the DNA can be estimated by the ratio of absorbance at 260 nm and 280 nm (OD_{260}/OD_{280}), which should be in the range of 1.8-2.0. Another indicator for the presence of contaminating substances, such as chaotropic salts and phenol is the OD_{260}/OD_{230} ratio which is expected to be around 2.0 in case of pure DNA.

2.4. Polymerase chain reaction (PCR)

Oligonucleotides with and without overlap extension were synthesized by Eurofins Genomics (Tab. 2.2). The reactions were set up in 0.2 ml tubes on ice by mixing the components summarized in Tab. 2.3. PCR was carried out using Phusion DNA polymerase (ThermoFisher Scientific) in a T Gradient thermocycler (Biometra) according to the protocol in Tab. 2.4.

Table 2.2. PCR reagents.

Component	50 μ l reaction	Final concentration
5X Phusion HF	10 μ l	1X
10 mM dNTPs	1 μ l	200 μ M
10 μ M forward primer	2.5 μ l	0.5 μ M
10 μ M reverse primer	2.5 μ l	0.5 μ M
Phusion DNA polymerase [2 U/ μ l]	0.5 μ l	1 U/50 μ l reaction
<i>P. furiosus</i> genomic DNA or	variable	200 ng
pYS14 plasmid DNA	variable	10 ng
Nuclease-free water	to 50 μ l	

Table 2.3. PCR oligonucleotides.

Primer name	Sequence 5'-3'	Template	Application
PF0708_for	accgaaatgaggtggtatgaatgagtgaaaaaaccaccaagggc	<i>P. furiosus</i> genomic DNA	<i>pf0708</i> amplification
PF0708_rev	tttcagtgatggtgatggtgatgccttttaccctgtagcggag		
PF0709_for	accgaaatgaggtggtatgaatgcgaaaagcctcgaagggag	<i>P. furiosus</i> genomic DNA	<i>pf0709-pf0708</i> amplification
PF0708_rev	tttcagtgatggtgatggtgatgccttttaccctgtagcggag		
pYS14_for	catcaccatcaccatcactgaaatc	pYS14 plasmid DNA	pYS14 backbone amplification with the <i>gdh</i> promoter
pYS14_rev	tcataccacctcatttcggaatc		
pYS14_F_seq	tctccgagaattggccttg	pYS14 plasmid DNA	DNA sequencing
pYS14_R_seq	tgccgctgcaagcattatgg		
ogdh_R	agcgagaaggatcctgcag	pYS17 plasmid DNA	<i>pf0613</i> promoter amplification
inglu_F	ctgcaggatacctctcgcctgatatccttaacatttct		
PF0708_gdh_F	atgagtgaaaaaaccaccaagggc	pYS14 plasmid DNA	pYS14 <i>pf0708</i> backbone amplification without the <i>gdh</i> promoter
8PYS_Glu_R	cgcccttggtggtttttcactcattttcacctcagaattttac		
PF0709_gdh_F	atgcgaaaagcctcgaagggagg	pYS14 plasmid DNA	pYS14 <i>pf0709-pf0708</i> backbone amplification without the <i>gdh</i> promoter
9PYS_Glu_R	cctccctcgaaggcttttcgattttcacctcagaattttac		

Table 2.4. PCR protocol.

Step	Temperature (°C)	Time (s)
Initial denaturation	98	30
Denaturation	98	10
Annealing	50-65	30
Extension	72	15-30 per kb
Final extension	72	600
Hold	4	0

2.5. Agarose gel electrophoresis

5/6 volume of the DNA sample was mixed with 1/6 volume of 6X DNA Gel Loading Dye (ThermoFisher Scientific, 10 mM Tris-HCl, pH 7.6), 0.03% bromophenol blue, 0.03% xylene cyan FF, 60% glycerol, 60 mM EDTA) and loaded onto a 1% (w/v) agarose gel. Electrophoresis was carried out in a tank filled with 1X TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) at 90 V at room temperature for 1-1.5 h. Visual tracking of DNA migration was possible due to the presence of bromophenol blue and xylene cyan FF, which co-migrate with DNA fragments of around 300 bp and 4000 bp, respectively. After electrophoretic separation, the agarose gel was stained in 0.0002% (w/v) EtBr solution for 15 min. To remove the excess of the staining solution, the gel was washed in H₂O for 5 min. EtBr intercalates between the base pairs of double stranded DNA and allow visualization of DNA fragments under the UV-light using an UV transillumination gel doc (Biorad).

2.6. Gel extraction of DNA

DNA corresponding to the PCR products was extracted from agarose gel after electrophoretic separation by a scalpel blade around a band of interest. Subsequent purification was carried out using QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's protocol. DNA sample was eluted with 10 mM Tris (pH 8.5).

2.7. Gibson Assembly cloning

Assembly of two DNA fragments for a shuttle vector generation was performed using the Gibson Assembly cloning kit (NEB). Schematic representation of this method is illustrated in Fig. X. The amount of DNA for the Gibson Assembly reaction was calculated according to the following formula: pmols = (weight in ng) x 1,000 / (base pairs x 650 daltons). The total amount of DNA fragments should be 0.02-0.5 pmols. The Gibson Assembly reaction was set

up on ice by mixing the Gibson Assembly Master Mix (2X) with the insert and vector DNA (Tab. 2.5). Afterwards, samples were incubated in a T gradient thermocycler (Biometra) at 50 °C for 60 min and subjected to transformation into NEB 5-alpha competent *E. coli* cells.

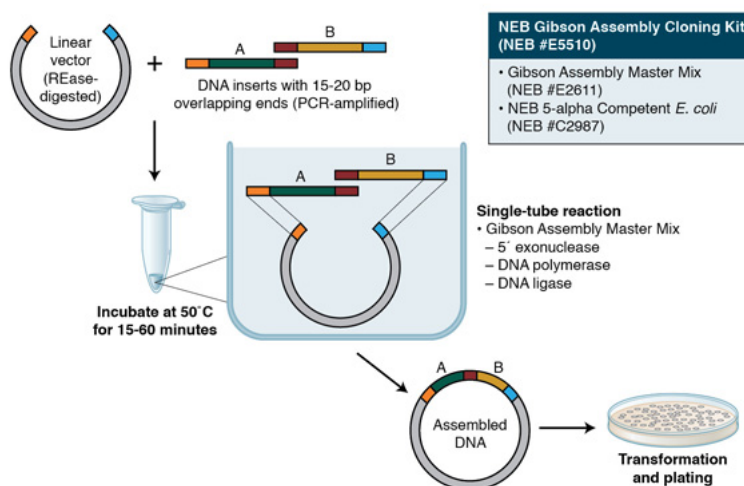


Fig. 2.1. Schematic representation of the Gibson Assembly method.

Table 2.5. Gibson Assembly reaction.

Component	20 µl	Final concentration	
Gibson Assembly Master Mix (2X)	10 µl	1X	
PCR product - backbone of pYS14 (24.40 ng/µl)	7 µl	183 ng, 0.039 pmols	0.175 pmols
PCR product - <i>pf0709</i> and <i>pf0708</i> insert (35.68)	3 µl	107.04 ng, 0.097 pmols	
Component	20 µl	Final concentration	
Gibson Assembly Master Mix (2X)	10 µl	1X	
PCR product - backbone of pYS14 (24.40 ng/µl)	7.5 µl	183 ng, 0.039 pmols	0.141 pmols
PCR product - <i>pf0708</i> insert (37.86 ng/µl)	2.5 µl	94.65 ng, 0.102 pmols	

2.8. Transformation of *E. coli* cells

For plasmid DNA transformation, 2 μ l of the Gibson assembly reaction was mixed gently with 50 μ l *E. coli* competent NEB 5-alpha cells thawed on ice, provided with the Gibson assembly kit. After incubation on ice for 30 min, the cells were subjected to a heat shock at 42 °C for 90s, followed by incubation on ice for 5 min. For the cell growth, 300 μ l of SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose) were added and the cells were incubated at 37 °C with vigorous shaking for 1 h. Positive clones were selected by spreading the cell suspension (10-50 μ l) onto the pre-warmed lysogeny broth (LB) agar (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar) plates with an appropriate antibiotic. After overnight incubation at 37 °C, plates were visually analyzed for presence of single colonies.

In case of *pf0708* heterologous expression, 100 ng of plasmid DNA containing the gene coding for PfMATE (pBAD-PfMATE) was mixed with *E. coli* TOP10 competent cells (Invitrogen). Transformation of *E. coli* competent TOP10 cells was performed in the same manner as described above.

2.9. Isolation of plasmid DNA

After transformation, a single colony was picked by a pipette tip as an inoculum for 3 ml LB medium (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) containing 50 μ g/ml carbenicillin. Cultures were incubated at 37 °C with shaking overnight and then centrifuged at 14,000 rpm for 30 min. Plasmid DNA was isolated from the *E. coli* cell pellet using the Plasmid DNA isolation kit (Qiagen) according to the manufacturer's protocol. Plasmid DNA was eluted with 10 mM Tris-HCl (pH 8.5).

2.10. Restriction enzyme reaction

Enzymatic digestion of plasmid DNA was performed according to the protocol provided by the manufacturer (Fermentas). Plasmid DNA was incubated with the enzymes *EcoRI* and *NotI* for 2 h at 37 °C (Tab. 2.6).

Table 2.6. Restriction enzyme reaction.

Component	10 µl reaction	Final concentration
10X Buffer 0	1 µl	1X
<i>EcoRI</i> (10 U/µl)	0.2 µl	0.2 U/µl
<i>NotI</i>	0.2 µl	0.2 U/µl
Plasmid DNA	variable	200 ng
Nuclease-free water	to 10 µl	

2.11. DNA sequencing

The full-length sequences of the cloned genes were confirmed by DNA sequencing carried out by Microsynth SeqLab. Plasmid DNA (15 µl of 50-100 ng/µl; 750-1500 ng in total) and the corresponding primers (2 µl of 10 µM) were premixed and sent for DNA sequencing. The results were analyzed using the MultiAlign Software.

2.12. Homologous gene expression in *P. furiosus*

2.12.1. ½ SME medium preparation

Firstly, SME solution (Tab. 2.7) and Wolfe's minerals (Tab. 2.8) were prepared. Afterwards, the remaining ingredients, except $\text{Na}_2\text{S} \times 3 \text{H}_2\text{O}$, were added and thoroughly mixed (Tab. 2.9). Upon addition, the starch had to be dissolved in water by boiling the solution in the Erlenmeyer glass flask for 1-2 min. The 1-liter glass bottle (Duran) containing ½ SME medium was closed with a rubber stopper and sparged with N_2 at 0.5 bar for 30 min. Afterwards, $\text{Na}_2\text{S} \times 3 \text{H}_2\text{O}$ was added as a reducing agent and thoroughly stirred for 30 min. The pH was adjusted to 7.0 with 0.5 ml 1 M NaOH. The medium along with 50 empty 120 ml glass bottles (Stute GmbH) and rubber stoppers was introduced into an anaerobic tent (Coy Laboratory Products Inc.), which had been pressurized with N_2/H_2 (95/5) before. 20 ml of ½ SME medium were dispensed into each bottle, which was then tightly closed with a rubber stopper. The medium containing bottles were taken out from the anaerobic tent, closed with aluminium caps and sparged three times with N_2 at 1.0 bar before being autoclaved for in an autoclave (LaM-MCS-J, SANOclav) at 1.5 bar and 121 °C for 20 min. Solid ½ SME medium was prepared in the same manner, but 1.5% (w/v) Gelrite (Roth) was added, due to its higher thermostability compared to agar, commonly used for solid LB media. After autoclaving, the Gelrite medium had to be cooled down to around 60 °C and then poured into Petri dishes (ca. 30 ml per dish). The plates were stored in the anaerobic tent for the subsequent experiments (Fig. 2.2).

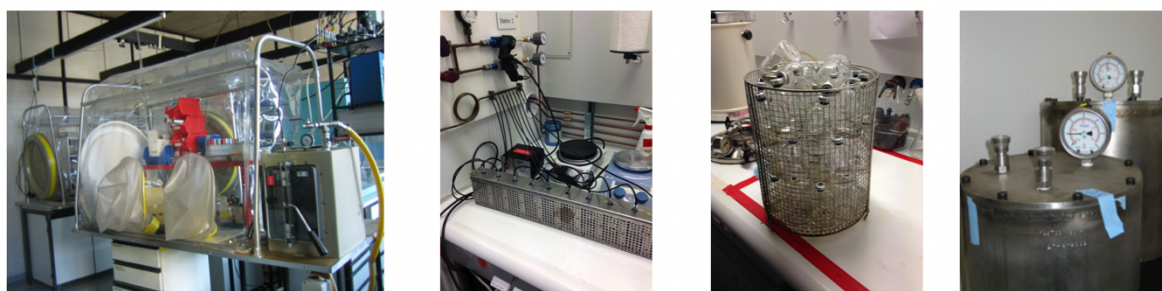


Fig. 2.2. Equipment for homologous expression in *P. furiosus*. (A) Anaerobic tent. (B) Gas station. (C) Glass bottles filled with medium before being autoclaved. (D) Anaerobic jar.

Table 2.7. SME (*Synthetisches Meerwasser/synthetic seawater*).

Substance	Amount	Final concentration (mM)
NaCl	27.7 g	473.99
MgSO ₄ x 7 H ₂ O	7.0 g	28.4
MgCl ₂ x 6 H ₂ O	5.5 g	27.1
CaCl ₂ x 2 H ₂ O	0.75 g	5.1
KCl	0.65 g	8.7
NaBr	0.1 g	0.97
H ₃ BO ₃	0.03 g	0.49
SrCl ₂ x 6 H ₂ O	0.015 g	0.056
KI	0.1 ml	0.00060
H ₂ O	up to 1,000 ml	

Table 2.8. Wolfe's minerals. pH 1.0 adjusted with 1 M H₂SO₄.

Substance	Amount	Final concentration (mM)
MgSO ₄ x 7 H ₂ O	30.0	121.70
MnSO ₄ x H ₂ O	5.0	29.60
NaCl	10.0	171.10
FeSO ₄ x 7 H ₂ O	1.0	3.60
CoSO ₄ x 7 H ₂ O	1.8	6.40
CaCl ₂ x 2 H ₂ O	1.0	6.80
ZnSO ₄ x 7 H ₂ O	1.8	6.30
CuSO ₄ x 5 H ₂ O	0.1	0.40
KAl (SO ₄) ₂ x 12 H ₂ O	0.18	0.38
H ₃ BO ₃	0.1	1.62
Na ₂ MoO ₄ x 2 H ₂ O	0.1	0.41
(NH ₄) ₂ Ni(SO ₄) ₂ x 6 H ₂ O	2.80	7.09
Na ₂ WO ₄ x 2 H ₂ O	0.1	0.30
Na ₂ SeO ₄	0.1	0.53
H ₂ O	up to 1,000 ml	

Table 2.9. ½ SME medium.

Substance	Amount	Final concentration
SME	500 ml	1/2X
KH ₂ PO ₄	0.5 g	3.67 mM
NH ₄ Cl	0.5 g	9.35 mM
Wolfe's minerals/10x	1.0 ml	0.01X
0.2% Resazurin	0.5 ml	0.0001 % (v/v)
Na ₂ S x 3 H ₂ O	0.3 g	2.27 mM
Yeast extract	1 g	0.1% (w/v)
Peptone	1 g	0.1% (w/v)
Starch	1 g	0.1% (w/v)
Na-pyruvate	4.4 g	40 mM
H ₂ O	up to 1,000 ml	

2.12.2. Transformation of plasmid DNA into *P. furiosus* cells

P. furiosus Δ1623 cells were incubated in the glass bottles containing ½ SME medium supplemented with either pyruvate or starch at 95 °C with shaking at 100 rpm. The preculture was enriched with 0.8 mM agmatine sulfate. For the transformation of *P. furiosus* cells the overnight cultures in the exponential growth phase were used. The following steps were performed in the anaerobic tent. 2 ml of the cell culture was centrifuged at 6,000 g for 10 min. The cell pellet was resuspended in 100 µl of the transformation buffer composed of ½ SME without KH₂PO₄, 40 mM Na-pyruvate, 4.7 mM NH₄Cl and 80 mM CaCl₂. After incubation at 4 °C for 30 min, 1 µg of plasmid DNA containing the gene *pf1623* was added in a volume of 10 µl. Incubation was carried out at 4 °C for 60 min, 85 °C for 3 min, followed by 4 °C for 10 min. Afterwards, 1:10 and 1:100 dilutions of the transformation solution in ½ SME medium were prepared for spreading the cells on the Petri dishes containing Gelrite medium without agmatine sulfate. Plates were incubated in the anaerobic jar at 85 °C for 72 h, which was pressurized 3 times with N₂ (0.5 bar). A single colony was picked by a syringe needle,

transferred into 20 ml medium and subsequently incubated at 85 °C overnight. 1.3 ml of the culture were used for the genomic DNA isolation by the PureYield Plasmid Miniprep System (Promega) according to the protocol of the manufacturer. The culture was centrifuged at 6,000 g for 10 min and the pellet was resuspended in 100 µl of TE buffer. After addition of 20 µl proteinase K (20 mg/ml), the reaction mix was incubated at 65 °C for 30 min and vortexed. 4 µl of RNaseA (100 mg/ml) were added, followed by incubation at room temperature for 2 min and vortexing. The remaining part of the culture (0.7 ml) was used for inoculating the fresh medium.

2.12.3. Cultivation of *P. furiosus* cells

The up-scaled *pf0708* overexpression was performed in 16- and 100-liter bioreactors (HTE, Bioengineering). The *P. furiosus* strain DSM3638 (WT) was grown under anaerobic conditions at 95 °C in ½ SME medium as described previously. Cultivation of the precultures was performed in serum bottles, which were used as inoculum for ½ SME medium in bioreactors at the University of Regensburg, Germany. Growth of *P. furiosus* was controlled by analyzing cell numbers under a light microscope (Axioskop 2 Plus, Zeiss) using a Thoma counting chamber of 0.02 mm depth (Marienfeld) (Fig. 2.3). Cells were harvested in the late exponential growth phase by a centrifuge, (Type CEPA Z41G, Padberg), flash-frozen in liquid nitrogen and subsequently stored at -80 °C.

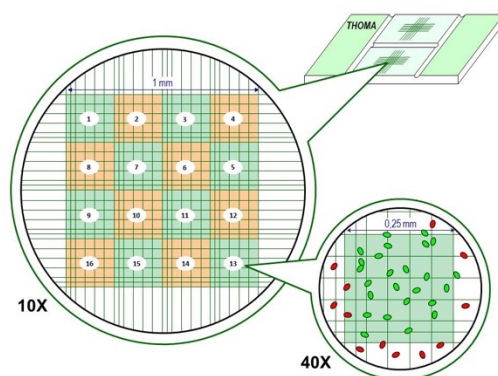


Fig. 2.3. Thoma cell counting chamber. Large central square is divided into 16 smaller squares.

2.13. Heterologous gene expression in *E. coli*

The gene encoding PfmMATE was expressed in *Escherichia coli* TOP10 cells (Invitrogen). Transformation with the plasmid DNA (pBAD-PfmMATE) was performed according to the protocol described in the section 2.8. A single colony was picked as an inoculum for the 5 ml LB medium preculture containing carbenicillin. After 6 h of incubation at 37 °C, 50 µl of this culture was used to inoculate 50 ml of fresh LB medium with 50 µg/ml carbenicillin and incubated at 37 °C overnight. This preculture was used as an inoculum for 2-liter of LB medium supplemented with carbenicillin in 5 liter unbaffled flasks. The cell growth was recorded spectrophotometrically. Once an optical density at 600 nm (OD₆₀₀) reached 0.6-0.8, gene expression was induced by addition of L-arabinose to a final concentration of 0.02% (w/v). After 3 h of incubation at 37 °C, cells were harvested by centrifugation at 4 °C for 15-30 min.

2.14. Membrane preparation and solubilization of membrane proteins

Bacterial cell pellets were resuspended in a fivefold excess (v/w) of lysis buffer composed of 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 5 mM MgCl₂, a pinch of deoxyribonuclease I (DNaseI) and 1 mM phenylmethylsulfonyl fluoride (PMSF). Cells were disrupted by three passages through a Microfluidizer (Microfluidics). The cell lysate was centrifuged in a GSA rotor to remove cell debris twice, first at 10,000 rpm at 4 °C for 30 min followed by 12,000 rpm at 4 °C for 30 min. To collect the membrane pellet, the clarified supernatant was ultracentrifuged in a Type 45 Ti rotor at 43,000 rpm at 4 °C for 3 h. The membrane fraction was resuspended in a buffer composed of 20 mM HEPES-NaOH (pH 8.0) and 300 mM NaCl, flash frozen in liquid nitrogen and subsequently stored at -80 °C. All following steps were performed at 4 °C.

Membrane proteins were solubilized by addition of 2% (w/v) β-DDM (n-dodecyl-β-D-

maltoside, Glycon Biochemicals) in 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl and 30 mM imidazole (pH 7.4), so that a final membrane protein concentration of 10 mg/ml was reached. After 2 h of rotation, the solubilized membrane proteins were separated by ultracentrifugation in a type 70.1 Ti rotor at 55,000 rpm for 2 h and filtered through a 0.20 µm centrifugal filter (Millipore).

2.15. Protein purification

PfMATE was purified using an ÄKTA system (ÄKTA prime or purifier) in two chromatographic steps, affinity chromatography followed by size exclusion chromatography (Tab 2.10).

Table 2.10. Purification steps.

Name	Purification step	Manufacturer
HisTrap 1 ml	Affinity chromatography	GE Healthcare
Superdex 200 10/300 GL	Size exclusion chromatography	GE Healthcare
Superdex PC 3.2/30	Size exclusion chromatography	GE Healthcare

2.15.1. Affinity chromatography

For affinity chromatography the resulting supernatant after ultracentrifugation of the solubilized membrane proteins containing the C-terminally His-tagged PfMATE was loaded onto a HisTrap HP column previously equilibrated with 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 30 mM imidazole (pH 7.4) and 0.05% (w/v) β-DDM. Two buffers were used, Buffer A composed of 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 0.05% (w/v) β-DDM and Buffer B composed of 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 500 mM imidazole (pH 7.4), 0.05% (w/v) β-DDM.

The purification consisted of the following elution steps:

- | | | | |
|----|-------|----------------------|-----------------|
| 1. | 5 CV | 30 mM imidazole | 6% Buffer B |
| 2. | 10 CV | 100 mM imidazole | 20% Buffer B |
| 3. | 5 CV | 100-400 mM imidazole | 20-80% Buffer B |
| 4. | 10 CV | 500 mM imidazole | 100% Buffer B |

Protein elution was monitored by measuring the UV absorbance at 280 nm. The protein was eluted with a linear gradient of imidazole (100-400 mM) or 500 mM imidazole in the purification buffer and concentrated using a centrifugal filter device (30,000 MWCO or 50,000 MWCO).

2.15.2. Size exclusion chromatography

The eluted and concentrated protein was further purified by size exclusion chromatography at 4°C or room temperature and loaded onto a Superdex PC 3.2/30 or Superdex 200 10/300 GL column, which was previously equilibrated with the gel filtration buffer containing 20 mM HEPES-NaOH (pH 8.0), 15 mM NaCl and 0.06% (w/v) 6-cyclohexyl-1-hexyl- β -D-maltoside (Cymal-6, Anatrace) or 20 mM HEPES-NaOH (pH 8.0), 150 mM NaCl 0.03% (w/v) β -DDM. Protein elution was monitored by measuring the UV absorbance at 280 nm. Protein peak fractions were pooled and concentrated with a centrifugal filter device (30,000 MWCO or 50,000 MWCO) to a final concentration of 10-20 mg/ml for crystallization experiments.

2.16. Protein concentration measurement

Quantitation of total protein in the isolated membranes was performed by a colorimetric method, namely the BCA Protein assay. The principle behind this experiment relies on the

reduction of Cu^{2+} to Cu^+ by protein in an alkaline medium and the selective colorimetric detection of Cu^+ by a reagent containing bicinchoninic acid. Chelation of one Cu^+ by two molecules of BCA results in a purple-colored reaction product. The BCA reagents A and B are mixed with the protein sample and incubated at 37 °C covered with an aluminum foil. After 30 min the protein concentration in the membrane fraction is acquired by measuring the absorbance at 562 nm using a spectrophotometer (GE Healthcare). The protein concentration is determined based on a standard curve with reference to a series of dilutions of a bovine serum albumin fraction V (BSA) solution (Thermo Fisher Scientific).

For the purified protein sample a spectrophotometric method was used by measuring the UV absorption at 280 nm using a NanoDrop device and calculating the concentration using the Beer-Lambert equation: $A = E \cdot b \cdot c$, where A, absorbance at 280 nm; E, extinction coefficient ($46870 \text{ M}^{-1}\text{cm}^{-1}$ for PfMATE); b, path length (cm); c, protein concentration (M). A protein sample of 2 μl was applied) and the purification buffer was used as a blank.

2.17. Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated by electrophoresis through the polyacrylamide gel under denaturing conditions using sodium dodecyl sulfate (SDS) according to the protocols described by Laemmli. Due to an even binding of SDS to the peptides, the proteins are separated according to their molecular weight. Sample aliquots were mixed with 4x SDS sample buffer (106 mM Tris-HCl, 141 mM Tris base, pH 8.5, 2% (w/v) SDS, 10% (v/v) glycerol, 0.51 mM EDTA, 0.22 mM SERVA® Blue G250, 0.175 mM phenol red) from Invitrogen (Darmstadt, Germany) in a ratio of 1:4. The electrophoretic separation was performed using Bolt® 4-12% Bis-Tris polyacrylamide gels with 1X MES or 1X MOPS SDS running buffer from Invitrogen (Darmstadt, Germany) (Tab 2.11) at 180 V for approximately 45-60 min. Separated proteins were stained with PageBlue™ protein staining solution from Thermo Fisher Scientific (St. Leon-Rot, Germany) and the banding pattern was inspected and recorded.

Table 2.11. SDS-PAGE buffers.

20X MOPS (pH 7.7)	20X MES (pH 7.3)
50 mM MOPS	50 mM MES
50 mM Tris base	50 mM Tris base
0.1% (w/v) SDS	0.1% (w/v) SDS
1 mM EDTA	1 mM EDTA

2.18. Non-denaturing polyacrylamide gel electrophoresis (BN-PAGE)

To determine native mass and oligomeric states, electrophoretic separation of folded protein was carried out using non-denaturing polyacrylamide gels according to the protocol by Schagger *et al.* The migration distances depend on the overall surface charge and the hydrodynamic radius of the protein particle in solution. Sample aliquots were mixed with 4x BN-PAGE sample buffer (50 mM Bis-Tris, pH 7.0, 50 mM NaCl, 10% (v/v) glycerol, 0.01% (w/v) bromophenol blue) in a ratio of 1:4. Two running buffers were utilized for BN-PAGE runs: 1x cathode buffer (50 mM Tricine, pH 7.0, 15 mM Bis-Tris) and 1x anode buffer (50 mM Bis-Tris, pH 7.0). The electrophoretic separation was performed using Novex® 4-16% Bis-Tris polyacrylamide gels at 150 V at room temperature for approximately 60 minutes and was continued at 4 °C at 250 V for 30 to 90 minutes.

2.19. Immunoblotting (Western Blot)

After electrophoretic separation of proteins via SDS-PAGE, the proteins were transferred to a PVDF membrane using the i-Blot system, Invitrogen (Darmstadt, Germany) for 7 min following the manufacturer's instructions. After blotting, the PVDF membrane was incubated in a 1X TBST buffer (Tab 2.12) containing bovine serum albumin fraction (2% (w/v) BSA) at room temperature for 60 minutes in order to reduce non-specific interactions between the antibody and the proteins. Afterwards, the PVDF membrane was washed with 1X TBST buffer and incubated with the monoclonal mouse anti-polyhistidine antibody conjugated to

alkaline phosphatase (1 to 2,000 diluted in 1X TBST buffer) at room temperature for 1-2 h. Thereafter the membrane was washed thrice with 1X TBST buffer for 5 min and then with the alkaline phosphatase (AP) reaction buffer (Tab 2.12).

In order to trigger a chromogenic alkaline phosphatase reaction, a development solution prepared in the AP buffer was added (250 µg/ml 5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt (BCIP) and 50 µg/ml nitro-blue tetrazolium chloride (NBT)). The colorimetric reaction was stopped by addition of water. Purple bands of an insoluble precipitate are caused by the His-tagged fusion protein.

Table 2.12. Western blot buffers.

10X TBST (pH 7.5)	10X AP (pH 9.5)
100 mM Tris	100 mM Tris
150 mM NaCl	100 mM NaCl
0.5% (v/v) Tween 20	5 mM MgCl ₂

2.20. Lipid extraction

Lipids of *P. furiosus* were extracted according to an MTBE (methyl tert-butyl ether) protocol¹¹⁸. Briefly, the *P. furiosus* cell pellet was mixed with 13 ml of methanol/MTBE (1:3.33, (v/v)) and homogenized on ice by an Ultra-Turrax homogenizer for 2 min. After addition of 2.5 ml doubly distilled water and 8 ml MTBE/methanol/doubly distilled water (10:3:2.5, (v/v/v)), the upper phase was taken off and the lower phase was re-extracted with another 13 ml MTBE/methanol/ doubly distilled water (10:3:2.5, (v/v/v)). The combined organic phases were dried and resuspended in 500 µl CHCl₃/MeOH (1:1 (v/v)). Of this extract 100 µl were once more evaporated and again resuspended in 100 µl IPA/MeOH/CHCl₃ (90:5:5, (v/v/v)).

2.21. Mass spectrometry-based lipidomics

2 μ l of the lipid extract were injected into a Waters BEH C8, 100 x 1 mm, 1.7 μ m HPLC column used with an Ultimate 3000 UHPLC (Thermo Scientific). Solvent A was water with 1% ammonia acetate and 0.1% formic acid, and solvent B was acetonitrile/2-propanol 5:2 with 1% ammonia acetate and 0.1% formic acid. A gradient elution started at 50% mobile phase B, rising to 100% B over 40 min; 100% B were held for 10 min and the column was re-equilibrated with 50% B for 8 min before the next injection. The flow rate was 150 μ l/min.

Data acquisition was performed according to the previously reported protocol by Orbitrap-MS (LTQ-Orbitrap, Thermo Scientific) full scan in the preview mode at a resolution of 100,000 and < 2 ppm mass accuracy with external calibration. The spray voltage was set to 4500 V and the capillary temperature was at 300 °C. From the FT-MS preview scan the 10 most abundant m/z values were picked in the data dependent acquisition (DDA) mode, fragmented in the linear ion trap analyzer and ejected at nominal mass resolution. Normalized collision energy was set to 50%, the repeat count was 2 and the exclusion duration 10s. Data analysis was performed using Lipid Data Analyzer, a custom developed software tool¹¹⁹.

2.22. Crystallization

Crystals of PfMATE in the inward and outward-facing conformations (PDB ID: 6FHZ and 6GWH, respectively) were obtained by the LCP technique^{103,120}. The protein solution in 20 mM HEPES-NaOH (pH 8.0), 15 mM NaCl and 0.06% (w/v) Cymal-6 was filtered through 0.20 μ m spin filters (Millipore) before being mixed with molten monoolein (9.9 MAG, Nu-Chek Prep) with a 2:3 (v/v) protein:monoolein ratio using a coupled syringe mixer. 96-well crystallization trays were set up using the ProCrys Meso lipidic mesophase dispenser (Zinsser Analytic). The size of the precipitant and protein-monoolein solution drop were 1.5 μ l and 100 nl, respectively. The usage of the MemMeso HT-96 screen (Molecular Dimensions) resulted in

the successful crystallization hits. The plates were stored in the incubator of the CrystalMation system (Rigaku) at 295 K.

For the inward-facing structure (PDB ID: 6FHZ), the protein sample was incubated with the total lipid extract from *P. furiosus* after the affinity chromatography step, subsequently co-purified on a Superdex 200 10/300 GL column and then crystallized. Crystals of PfMATE in the inward-facing conformation were grown in 0.1 M sodium chloride, 0.1 M magnesium chloride, 30% (v/v) PEG600 and 0.1 M sodium citrate (pH 5.0). PfMATE in the outward-facing state was captured under almost identical conditions (0.1 M sodium chloride, 30% (v/v) PEG500 DME and 0.1 M sodium citrate (pH 5.0)), however in absence of the native lipids from *P. furiosus*. After crystals matured to their full size, they were harvested and directly flash-frozen in liquid nitrogen without additional cryoprotectant. Crystallization screens used for crystallization of PfMATE in the presence or absence of different lipid species are summarized in the Table 2.13.

Table 2.13. Crystallization screens.

Name	Manufacturer
Screens	
MemMeso HT-96	Molecular Dimensions
MemGold Eco	Molecular Dimensions
LCP_Screen_PEG400_Citrate	In-house designed screen
LCP_Screen_Salts	In-house designed screen
sazakrze_screen1	In-house designed screen
sazakrze_screen2	In-house designed screen
sazakrze_screen3	In-house designed screen
sazakrze_screen4	In-house designed screen
sazakrze_screen5	In-house designed screen
sazakrze_screen6	In-house designed screen
sazakrze_screen7	In-house designed screen
MPI Custom Screen pH 5.6	In-house designed screen

2.23. Data collection and structure refinement

The X-ray diffraction datasets for PfMATE in the inward and outward-facing conformation (PDB ID: 6FHZ and 6GWH, respectively) were collected at the PXII/X10SA beamline of the synchrotron Swiss Light Source (SLS) in Villigen, Switzerland. Data acquisition was performed at 100 K. The diffraction data were processed with XDS package for indexing, merging and scaling¹²¹. A PILATUS 6M pixel detector from Dectris (Baden-Dättwil, Switzerland) was used. All datasets were collected in the high redundancy mode (360 ° rotation) with a rotation angle increment of 0.2 °, an exposure time of 0.1 s, and 50 filter transmission.

For the determination of PfMATE in the outward-facing conformation (PDB ID: 6GWH), the PfMATE structure obtained with crystals using the VD method (PDB ID: 4MLB), namely a monomer from the tetrameric asymmetric unit, was used as the search model for molecular replacement with Phaser from the CCP4 suite¹²². The quality of the electron density map allowed to localize most of the amino acid residues, with the exception of the loop regions connecting TM3 and TM4 as well as TM9 and TM10. The model was subjected to manual adjustment in COOT¹²², and subsequent iterative refinement at a resolution limit 2.8 Å with phenix.refine^{123,124}. Parameters applied for the refinement strategy: XYZ coordinates, group B-factors, occupancies with optimized X-ray/stereochemistry and ADP weight. In the final model, 98.39% of the residues were in the favoured region and the rest in the allowed region in the Ramachandran plot.

For the determination (PDB ID: 6FHZ) of the structure of PfMATE in the inward facing conformation the initial phases were also obtained by molecular replacement¹²⁵⁻¹²⁷. The same reference model (PDB ID: 4MLB) was split into two halves at the centre: N-terminal fragment and C-terminal fragment (TMs 1-6 and TMs 7-12, respectively) and these two separate ensembles were used for the model generation in Phaser. Except for the perturbed fragment of

TM1, all TMs were assigned unambiguously. Due to the weak electron density map in the intramembrane region of TM1, the main chain tracing was validated by simulated annealing composite omit maps and feature-enhanced map (FEM) generated by PHENIX^{128–130}. After multiple cycles of manual model rebuilding in COOT, iterative refinement at a resolution limit 2.8 Å was performed with phenix.refine. Parameters applied for the refinement strategy: XYZ coordinates, group B-factors, occupancies with optimized X-ray/stereochemistry and ADP weight. The final model has 97.03% of the residues in the favored region in the Ramachandran plot, and 2.51% of the residues in the allowed region. The quality of the models was assessed using MolProbity¹³¹ and refinement statistics. The R.M.S.D. was calculated by TM-align¹³². All structural figures were prepared with PyMOL¹³² and UCSF Chimera¹³² (Tab. 2.14).

Table 2.14. Software for data processing and visualization.

Name	Application
Phenix program suite	Data processing and structure refinement
CCP4 program suite	Data processing
XDS	Data processing
XSCALE	Data processing
XDSCONV	Data processing
Pymol	Structure visualization
UCSF Chimera	Structure visualization
Coot	Manual model building
TM align	Structural alignment
UCLA-DOE LAB – Diffraction Anisotropy	Assessment of diffraction anisotropy

2.24. Molecular dynamics simulations

The PfMATE structures were embedded in lipid bilayers using CHARMM-GUI¹³³. Based on multi-conformation continuum electrostatics (MCCE)^{134,135} calculations at pH 7, the protonation states of the titrable Asp, Glu, and His residues were assigned. Asp184 and Glu331 were assigned neutral in the outward-facing conformation, while the other Asp and Glu residues were charged. In the inward-facing conformation Asp41, Asp184, Glu163, Glu273, and Glu331 were protonated in the absence of a sodium ion near Asp41/Asp184, whereas in the presence of a sodium ion, Asp41 was charged. All His residues were neutral and they were protonated at either the N δ or N ϵ atom according to the network of hydrogen bonds. All systems were hydrated with 150 mM NaCl as electrolyte. The all-atom CHARMM36 force field was used for protein, lipids, and ions, with TIP3P water^{136,137}

The MD trajectories were analysed using Visual Molecular Dynamics (VMD)¹³⁸. All simulations were performed using GROMACS 5.0.6¹³⁹. The starting systems were energy minimized for 5,000 steepest-descent steps and equilibrated initially for 500 ps of MD in an NVT ensemble and later for 8 ns in an NPT ensemble under periodic boundary conditions. During equilibration, the restraints on the positions of non-hydrogen protein atoms of initially 4000 kJ·mol⁻¹·nm² were gradually released. Particle-mesh Ewald summation¹⁴⁰ with cubic interpolation and a 0.12 nm grid spacing was used to treat long-range electrostatic interactions. The time step was initially 1 fs and then increased to 2 fs. The LINCS algorithm¹⁴¹ was used to fix all bond lengths. Constant temperature was set with a Berendsen thermostat¹⁴² with a coupling constant of 1.0 ps. A semiisotropic Berendsen barostat was used to maintain a pressure of 1 bar. During the production run, the Berendsen thermostat and barostat were replaced by the Nose-Hoover thermostat¹⁴³ and the Parrinello-Rahman barostat¹⁴⁴. Analysis was carried out on unconstrained simulations.

Initially, the outward-facing structure was embedded in a bilayer of 68 palmitoyl oleoyl

phosphatidyl-glycerol (POPG), 102 palmitoyl oleoyl phosphatidic acid (POPA), and 168 palmitoyl oleoyl phosphatidyl-inositol (POPI) lipids. 500 ns and 1,2000 ns of unconstrained production simulations were carried out for the system at 358 K. Afterwards, the lipids were replaced by their archaeal counterparts (i.e. POPA by DPA, POPG by DPG, and POPI by DPI). After the equilibration phase, the archaeal-lipid embedded PfMATE was simulated for 1.5 μ s at 373 K.

The inward-facing structure was embedded in a bilayer of 68 palmitoyl oleoyl phosphatidyl-glycerol (POPG), 102 palmitoyl oleoyl phosphatidic acid (POPA), and 170 palmitoyl oleoyl phosphatidyl-inositol (POPI) lipids. Unconstrained production simulation was carried out for 450 nsec at 358 K. To assess the putative sodium ion-binding site, a sodium ion was placed near Asp41/Asp148. After equilibration, unconstrained production simulation was carried out for 250-ns at 358 K.

To explore the conformational dynamics during the alternation of the access state, we performed a targeted MD simulation starting from the inward-facing structure and targeting the outward-facing structure. For the targeted MD, we used the PLUMED v2.1¹⁴⁵ patch in GROMACS 5.0.6. A bias potential acting on the C α RMSD pushed the initial structure, the membrane embedded inward-facing structure of PfMATE, toward the target structure, the outward-facing structure. The targeted MD run was 50 ns.

2.25. Differential scanning fluorimetry (DSF)

DSF (differential scanning fluorimetry) was performed using label-free nanoDSF Prometheus (NanoTemper Technologies, Munich)¹⁴⁶ in order to determine thermal stability of the purified protein in the presence and absence of lipids. The intrinsic fluorescence at 330 nm and 350 nm is used to monitor the fluorescence change of Trp and Tyr residues upon heat

unfolding without the need for additional dye. 10 μl of the protein sample was loaded into the glass capillary and the measurements were performed at a heating rate of 1 $^{\circ}\text{C}/\text{min}$.

The first derivative of the unfolding curves was used to determine the transition midpoint, T_m , at which 50% of the protein molecules are unfolded, reflecting the thermal stability of the protein (Fig. 2.4).

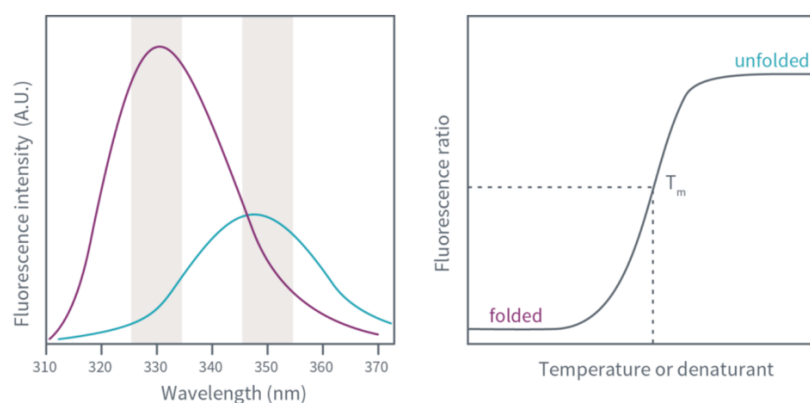


Fig. 2.4. DSF experiment. Prometheus monitors the intrinsic fluorescence signal of proteins as a measure of its folding state. A measurement of the fluorescent signal is plotted against increasing temperature or concentration of a chemical denaturant to determine the T_m of a protein. Image adapted from¹⁴⁶.

2.26. Native mass spectrometry

Prior to mass spectrometry analysis, protein solutions were buffer exchanged into 200 mM ammonium acetate pH 8.0, with 2X CMC of LDAO detergent of interest using a Biospin-6 (BioRad) column and introduced directly into the mass spectrometer using gold-coated capillary needles prepared in-house. Data were collected on a modified QExactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) optimized for analysis of high mass complexes, using methods previously described for membrane proteins¹⁴⁷. The instrument parameters were as follows: capillary voltage 1.2 kV, S-lens RF 100%, quadrupole selection from 2,000 to 20,000 m/z range, collisional activation in the HCD cell 100 V, argon UHV pressure 1.12×10^{-9} mbar, temperature 100 $^{\circ}\text{C}$, resolution of

the instrument at 17,500 at $m/z = 200$ (a transient time of 64 ms) and ion transfer optics (injection flatpole, inter- 5 flatpole lens, bent flatpole, transfer multipole: 8, 7, 6, 4 V respectively). The noise level was set at 3 rather than the default value of 4.64. No in-source dissociation was applied. Where required, baseline subtraction was performed to get a better-quality mass spectrum. For lipid binding analysis, lipids were first solubilized in 200 mM ammonium acetate (pH 8.0) and 0.05% LDAO and were added in increasing amounts while keeping the protein concentration constant. The optimized mass spectrometry conditions for lipid binding are same as described above.

3. RESULTS

The results part of this dissertation is divided into three main sections. The first one describes experiments leading to the homologous production of PfMATE: (i) cloning, (ii) expression in the native microorganism, *Pyrococcus furiosus*, as well as the results of the subsequent (iii) purification and (iv) crystallization trials. The second part presents the follow-up work on PfMATE produced heterologously in *Escherichia coli*, including (i) expression, (ii) purification and (iii) crystallization experiments in the absence and presence of native lipids from *P. furiosus* resulting in (iv) the structure determination of PfMATE in two alternate conformations, outward and inward facing state, respectively. In the last section results of the molecular dynamics (MD) simulations and native mass spectrometry (MS) are presented, which shed light on the interplay between PfMATE and lipids.

The aforementioned strategies were applied in order to achieve an ultimate goal, namely the structure determination of a MATE family transporter in a novel conformational state and expanding our understanding about the mechanistic and functional implications of lipids for PfMATE.

3.1. PfMATE produced homologously

3.1.1. Construction of the shuttle vector for the homologous expression of *pf0708*

Genetic manipulation, specified here as homologous expression in the hyperthermophilic and anaerobic archaeon, requires particular genetic tools. The *P. furiosus* transformation system used in this work is based on the vector pYS14 kindly provided by PD Dr. Winfried Hausner from the University of Regensburg. This shuttle vector possesses an auxotrophic selectable marker, which allows selection in *P. furiosus*, and the ampicillin selection system used in *E. coli*. Gene expression using pYS14 is possible in *P. furiosus* due to the presence of the constitutive promoter *gdh* (glutamate dehydrogenase) as well as of a resistance cassette composed of the gene *pf1623* coding for arginine decarboxylase, which is crucial for biosynthesis of polyamines. As a result of *pf1623* disruption in *P. furiosus* $\Delta 1623$ strain, a vector containing *pf1623* is essential for the cells to grow in complex media lacking the polyamine precursor, agmatine. A bioinformatic analysis of the *P. furiosus* genomic DNA revealed that the gene *pf0709*, coding for a putative transcriptional regulator, is located upstream of *pf0708*. Taking into consideration a possible regulatory effect on the *pf0708* gene expression, two constructs were generated. The first construct contains *pf0708* and the second *pf0708* together with *pf0709* (Fig. 3.1).

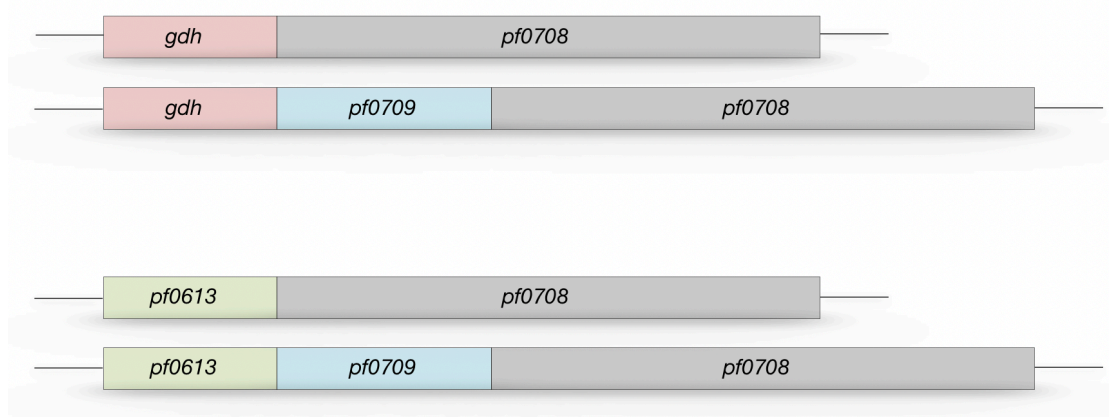


Fig. 3.1. Schematic representation of the constructs for the homologous expression of *pf0708*. Four different constructs containing *pf0708* coding for PfMATE or *pf0708* together with *pf0709* coding for a putative transcription regulator. *gdh*, glutamate dehydrogenase promoter; *pf0613*, fructose-1,6-bisphosphatase promoter.

The target genes *pf0708* and *pf0709-pf0708* were amplified by PCR using the genomic DNA extracted from *P. furiosus* as template. Amplification of the inserts and the pYS14 backbone was performed with the oligonucleotides containing overlapping ends (PF0708_for and PF0708_rev for *pf0708*; PF0709_for and PF0709_rev for *pf0709-pf0708*; pYS14_for and pYS14_rev for pYS14). The results of the gradient PCR at the annealing temperatures from 50 to 65 °C show that all genes were successfully amplified at 60.2-65 °C for *pf0708* and 64.6-65 °C for *pf0709-pf0708* (Fig. 3.2). In case of the pYS14 backbone, annealing temperatures of 60.2-65 °C resulted in a single band on the agarose gel. Therefore, 65 °C was chosen for the final PCR of all three DNA fragments.

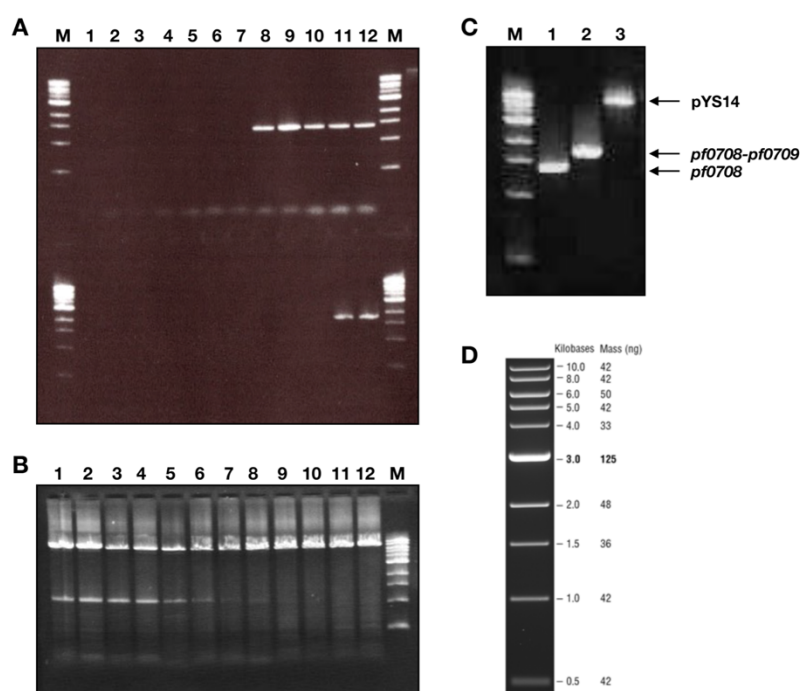


Fig. 3.2. 1% (w/v) agarose gels showing DNA fragments after PCR. (A-B) Each lane corresponds to the product of a gradient PCR at different annealing temperatures: 1, 50 °C; 2, 50.3 °C; 3, 51.4 °C; 4, 53 °C; 5, 54.8 °C; 6, 56.6 °C; 7, 58.4 °C; 8, 60.2 °C; 9, 61.9 °C; 10, 63.6 °C; 11, 64.6 °C; 12, 65 °C. (A) Upper lanes illustrate the PCR products of *pf0708* amplification (1426 bp). Lower lanes correspond to the PCR products of *pf0709-pf0708* amplification (1785 bp). (B) PCR products of pYS14 backbone amplification (7132 bp). (C) Each lane represents the final PCR products at 65 °C used for the Gibson Assembly cloning. 1, *pf0708*; 2, *pf0709-pf0708*; 3, backbone of pYS14. (D) M, 1 kb DNA ladder, NEB.

The agarose gel bands corresponding to the expected sizes of the amplified genes (1426 bp for *pf0708*; 1785 bp for *pf0709-pf0708*; 7132 bp for the pYS14 backbone) were cut out from the agarose gel for gel extraction and purification after enzymatic reactions. The final concentrations of the purified linear DNA fragments were measured using the NanoDrop spectrophotometer based on the absorbance at 260 nm. Fusion of the respective genes into the shuttle vector pYS14 was performed with the Gibson Assembly method, which allows joining multiple DNA fragments during one isothermic reaction.

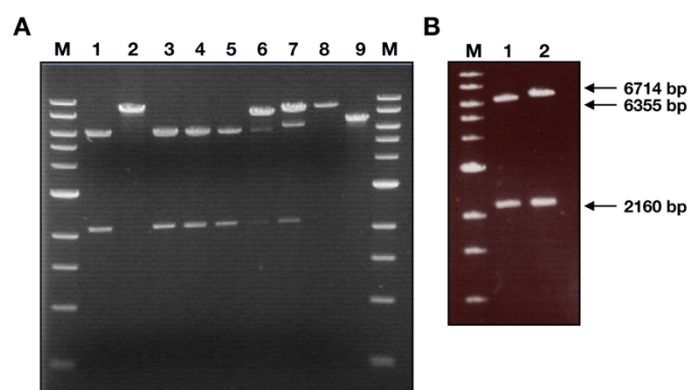


Fig. 3.3. 1% (w/v) agarose gel representing the DNA fragments after restriction enzyme digestion of the plasmid DNA. (A) Each lane corresponds to the DNA fragments obtained after digestion of the vector with restriction endonucleases, *EcoRI* and *NotI*. 1-5, pYS14-*pf0708*; 6-7, pYS14-*pf0709-pf0708*; 8, pYS14; 9, uncut PCR product of the pYS14 backbone amplification. Expected bands after the enzymatic hydrolysis: pYS14-*pf0708*, 6355 bp and 2160 bp; pYS14-*pf0708-pf0709*, 6714 bp and 2160 bp; pYS14, 9382 bp; uncut PCR product (pYS14 backbone), 7132 bp; uncut pYS14-*pf0708*, 8515 bp; uncut pYS14-*pf0708-pf0709*, 8874 bp. (B) Products of the restriction enzyme digests with *EcoRI* and *NotI*. 1, pYS14-*pf0708*; 2, pYS14-*pf0709-pf0708*. M, 1 kb DNA ladder, NEB.

The Gibson assembly reaction of two DNA fragments was performed by mixing the insert (*pf0708* or *pf0709-pf0708*) and the linear backbone of the pYS14 vector with the Gibson Assembly Reaction Mix. The amounts of the insert and vector were optimized accordingly, so that the final excess of the insert over the pYS14 vector was 2.58-fold in case of *pf0708* and 2.51-fold for *pf0709-pf0708*. After incubation at 50 °C for 60 min, the assembly product was

used for transformation of NEB 5-alpha competent *E. coli* cells. Selection of the positive clones was performed on the plates containing carbenicillin. Subsequently, plasmid DNA was isolated from single colonies grown on the plates and then doubly digested by the restriction endonucleases *EcoRI* and *NotI* in order to identify clones containing the gene of interest. The products of the restriction enzyme digestions were subjected to agarose gel electrophoresis (Fig. 3.3). Plasmid DNA samples, which resulted in the expected band sizes (6355 bp and 2160 bp for pYS14-*pf0708*; 6714 bp and 2160 bp for pYS14-*pf0709-pf0708*) were finally mixed with the sequencing oligonucleotides (pYS14_F_seq and pYS14_R_seq) and sent to Microsynth Seqlab for DNA sequencing.

Additionally, two constructs with the fructose-1,6-bisphosphatase promoter (*pf0613*) were generated in the same manner as described above. Overlapping oligonucleotides were designed accordingly to amplify and assemble the *pf0613* promoter from the vector pYS17 (PF0708_gdh_F and 8PYS_Glu_R or PF0709_gdh_F and 9PYS_Glu_R) and the backbone of the previously generated vectors pYS14-*pf0708* or pYS14-*pf0708-pf0709* without the *gdh* promoter (ogdh_R and inglu_F). The *pf0613* promoter is repressed under glycolytic conditions (in the presence of starch as an energy source) and induced under gluconeogenic conditions (in the presence of pyruvate), whereas the *gdh* promoter is induced in the presence of starch in complex growth media. The generated plasmid DNA was subsequently applied for the transformation of *P. furiosus* cells in order to produce PfMATE homologously.

3.1.2. Gene expression in *P. furiosus*

Genetic manipulation of *P. furiosus* was conducted at the University of Regensburg in the laboratory supervised by PD Dr. Winfried Hausner by myself. Plasmid DNA (pYS14-*pf0708* or pYS14-*pf0709-pf0708*) generated by the Gibson Assembly method was used for the transformation of *P. furiosus* Δ 1623 cells exhibiting auxotrophy for agmatine. Gene expression was tested under control of two different promoters (*gdh* and *pf0613*). Cells grown to a cell density of $0.8-1.0 \times 10^8$ cells/ml under anaerobic conditions at 85 °C were used for transformation of the plasmid DNA into *P. furiosus* Δ 1623 cells. Subsequently, selection of the positive clones was performed on the Gelrite plates with $\frac{1}{2}$ SME medium lacking agmatine. After around 72 h of incubation in the anaerobic jars at 87 °C, a single colony was transferred into 20 ml $\frac{1}{2}$ SME medium and cultivated for around 17 h. In case of the cells transformed with the vector containing the *gdh* promoter, the growth medium was supplemented with starch, whereas the cells containing the construct with the *pf0613* promoter were incubated in the presence of pyruvate. Afterwards, plasmid DNA was isolated from *P. furiosus* Δ 1623 cells by the PureYield Plasmid Miniprep System according to the protocol of manufacturer (Promega). Successful transformation was confirmed by PCR using primers corresponding to the genes *pf0708* or *pf0709-pf0708*. *P. furiosus* Δ 1623 culture containing the desired plasmid DNA was used as an inoculum for $\frac{1}{2}$ SME medium in a 100-liter bioreactor. PfMATE was homologously produced by cultivating *P. furiosus* Δ 1623 cells at 95 °C under anaerobic conditions in the bioreactor pressurized up to 2.0 bar with N₂. The cell growth was monitored every 2 h by counting the cells manually under the light microscope with a Thoma counting chamber. After reaching the stationary phase, characterized by the pivot point in the growth curve (Fig. 3.4), the cells were collected by centrifugation, flash-frozen in liquid N₂ and subsequently stored at -80 °C. Cultivation of 100-liter batches of *P. furiosus* overexpressing *pf0708* typically yielded 55-80 g of wet cell mass.

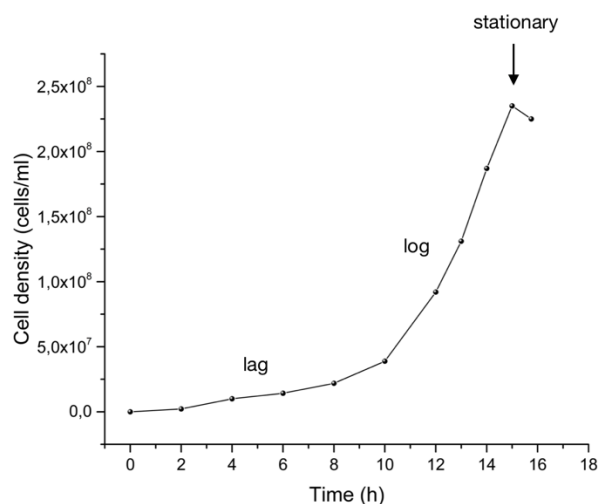


Fig. 3.4. Growth curve of *P. furiosus* Δ1623 *pf0708* overexpression strain at 95 °C. The growth curve can be divided into the lag phase (0-8 h), log phase (8-15 h) and stationary phase (from 15 h). The black arrow indicates the beginning of the stationary phase.

3.1.3. Purification of PfMATE produced homologously

The follow-up work was performed in Prof. Hartmut Michel's laboratory at the Max Planck Institute of Biophysics, Frankfurt am Main. The cell pellet of *P. furiosus* Δ1623 pYS14-*pf0708* or *P. furiosus* Δ1623 pYS14-*pf0709-pf0708* was resuspended in a buffer containing 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 5 mM MgCl₂, a pinch of DNaseI and 1 mM PMSF, and subsequently subjected to cell lysis by three passages through a Microfluidizer. Afterwards the membranes were isolated from the soluble fraction by ultracentrifugation. The total protein concentration in the membrane fraction was determined by the BCA assay in order to prepare aliquots of 20 mg/ml of the final protein concentration. Membrane proteins were solubilized using different concentrations of β-DDM at different temperatures, 2% (w/v) β-DDM at 50 °C, 2% (w/v) β-DDM at 40 °C, 3% (w/v) β-DDM at 40 °C and 2% (w/v) β-DDM at 4 °C.

Owing to the presence of the C-terminal HisTag, immunodetection of PfMATE was performed by Western Blotting using a monoclonal anti-polyhistidine-alkaline phosphatase coupled antibody, which yields an insoluble purple precipitate after reaction with the substrates

NBT and BCIP. Although there was no signal corresponding to the His-tagged protein in the cell lysate (presumably due to the amount of the protein below the detection limit) and a very weak signal in the solubilized membrane fraction (Fig. 3.5A and Fig. 3.6A), the solubilized membrane fraction was applied to a two-step purification.

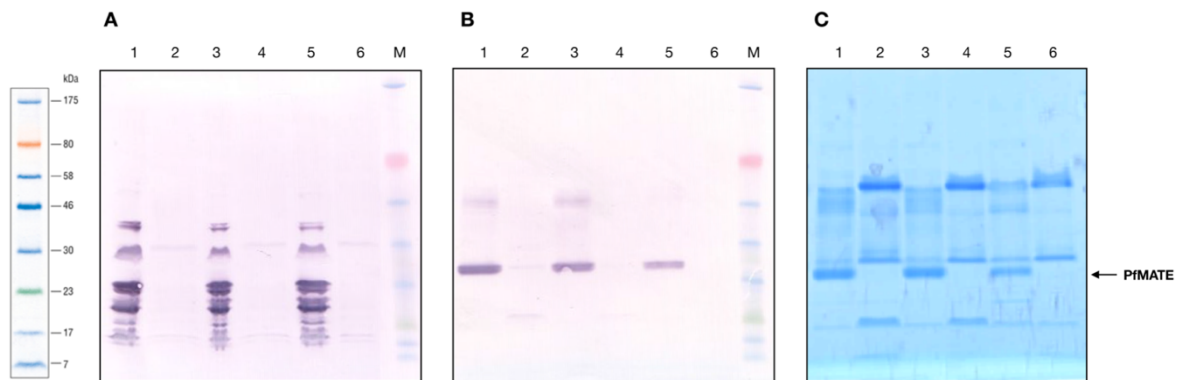


Fig. 3.5. Western Blots and SDS-PAGE gels of PfMATE produced under control of the *gdh* promoter after solubilization and affinity chromatography. (A-C) Lanes 1-2 correspond to the sample solubilized with 2% (w/v) β -DDM at 40 °C; lanes 3-4, with 3% (w/v) β -DDM at 40 °C and lanes 5-6, with 2% (w/v) β -DDM at 4 °C. (A) Western Blot results of the samples after solubilization. Lanes 1, 3, 5, pellet after ultracentrifugation of the solubilized samples. Lanes 2, 4, 6, supernatant after ultracentrifugation of the solubilized samples. (B) Western Blot results and (C) SDS-PAGE results of the purified samples after affinity chromatography. Lanes 1, 3, 5, correspond to the elution fraction E1, whereas lanes 2, 4, 6 represent the elution fraction E2.

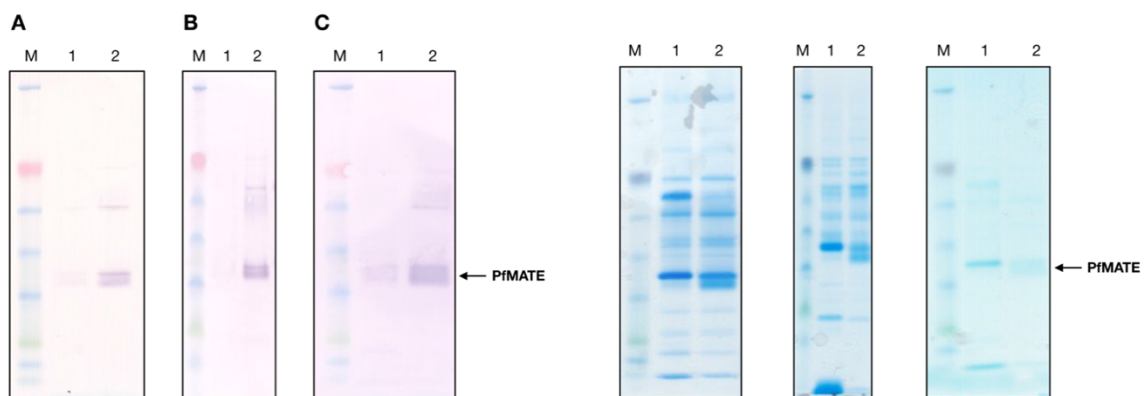


Fig. 3.6. Western Blots and SDS-PAGE gels of PfMATE produced under control of the *pf0613* promoter after affinity chromatography. Left images show the Western Blots, the right images the SDS-PAGE gels. PfMATE was solubilized with (A) 2% (w/v) β -DDM at 50 °C; (B) 2% (w/v) β -DDM at 40 °C; (C) 2% (w/v) β -DDM at 4 °C. Lane 1 corresponds to the elution fraction E1, whereas lane 2 represents the elution fraction E2.

Firstly, the affinity chromatography based on His tag/nickel interaction was performed on a HisTrap column followed by a polishing step via size exclusion chromatography on a Superdex PC 3.2/30 analytical column. Protein elution was monitored by measuring the absorbance at 280 nm. As shown in Fig. 3.7, representing a typical purification profile of PfMATE solubilized with 2% (w/v) β -DDM at 40 °C, the protein could be eluted with around 500 mM imidazole. This chromatogram shows the presence of two peaks, referred to as the elution fraction 1 (E1) and elution fraction 2 (E2).

The eluted samples were concentrated with a centrifugal filter device for electrophoretic analysis using an SDS-PAGE gel and subsequent detection by Coomassie staining and Western Blotting (Fig. 3.5 and Fig. 3.6). Based on the interaction with the anti-polyhistidine-alkaline phosphatase antibody, the His-tagged fusion protein was present exclusively in the E1 fractions. Affinity chromatographs of PfMATE samples solubilized at various conditions are summarized in Fig. 3.8.

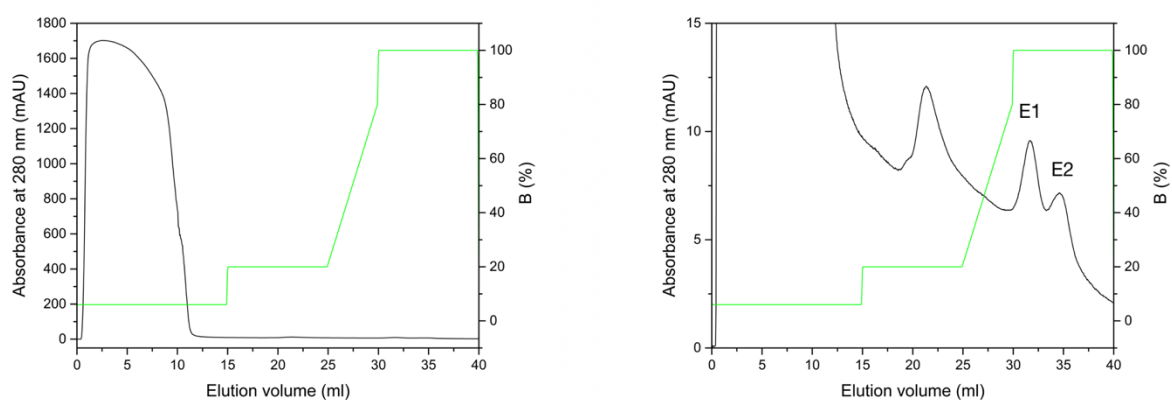


Fig. 3.7. Affinity chromatography columns elution profiles of PfMATE produced homologously. The black line shows the absorbance at 280 nm, whereas the green line shows the concentration gradient of buffer B composed of 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 0.05% (w/v) DDM and 500 mM imidazole. Left, the full purification profile. Right, close up of the purification profile between 0-15 mAU, including two elution peaks, referred here to as E1 and E2.

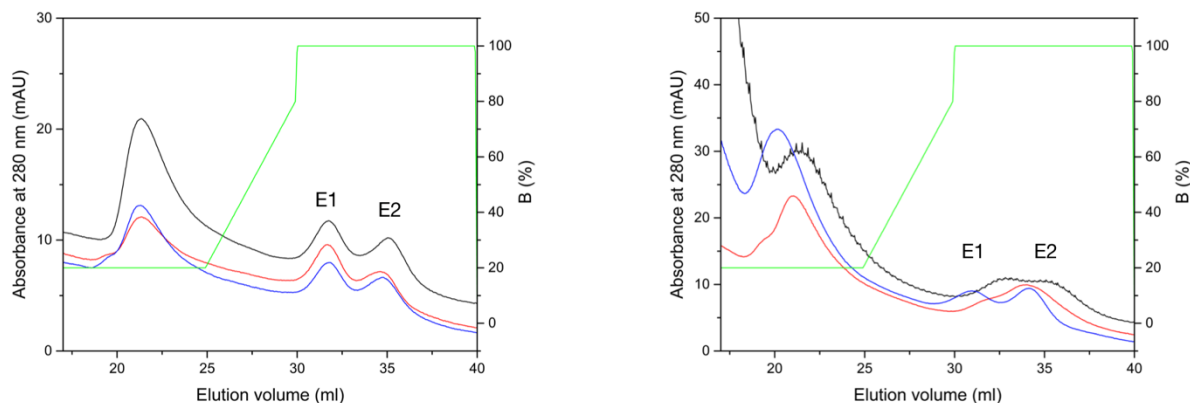


Fig. 3.8. Affinity chromatography columns elution profiles of PfMATE produced homologously under the control of two different promoters. Left, close up of the purification profile of PfMATE produced under control of the *gdh* promoter. Membrane protein solubilization was performed with 2% (w/v) β -DDM at 40 °C, 3% (w/v) β -DDM at 40 °C and 2% (w/v) β -DDM at 4 °C represented by black, blue and red lines, respectively. Right, close up of the purification profile of PfMATE produced under control of the *pf1623* promoter. Membrane protein solubilization was performed with 2% (v/w) β -DDM at 50 °C, 2% (w/v) β -DDM at 40 °C and 2% (w/v) β -DDM at 4 °C represented by black, blue and red lines, respectively. The elution fractions are referred to as E1 and E2. The green line shows the concentration gradient of buffer B composed of 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 0.05% (w/v) DDM and 500 mM imidazole, whereas the remaining lines represent the absorbance at 280 nm.

After affinity chromatography, fractions corresponding to E1 and E2 were concentrated with a centrifugal filter device, filtered through a 0.20 μ m filter and loaded onto a Superdex PC 3.2/30 column previously equilibrated with a buffer composed of 20 mM HEPES-NaOH (pH 8.0), 150 mM NaCl, 0.03% (w/v) β -DDM. The chromatogram corresponding to gel filtration of E1 (Fig. 3.9) contains a main peak with an elution volume at around 1.5 ml associated with some minor impurities.

Based on the Western Blot results, a purple band corresponding to the His-tagged PfMATE was detected. The homogeneity of the sample E1 was improved by changing the buffer to 20 mM HEPES-NaOH (pH 8.0), 15 mM NaCl, 0.06% (w/v) Cymal-6, as shown in Fig. 3.10. Despite a relatively high sample purity and homogeneity, illustrated by the presence of a single peak in the chromatogram and a single band on the SDS-PAGE gel, the final quantity of

PfMATE produced homologously was very low, in the μg range. The purification profile of sample E2 contains a peak with a retention volume at around 1.7 ml (Fig. 3.11). Due to the lack of a signal corresponding to the His-tagged protein by Western Blot analysis, sample E2 may be a contaminant with a high affinity for nickel ions in the HisTrap column.

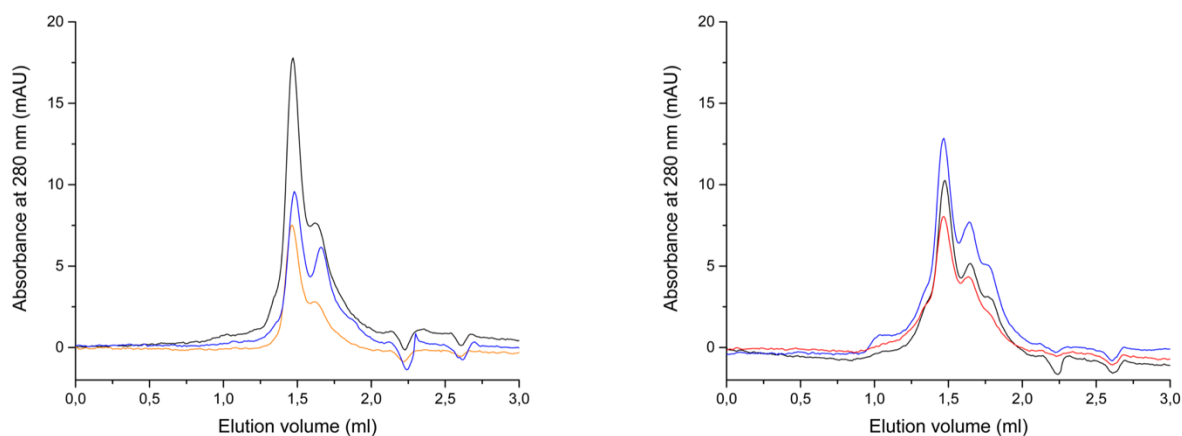


Fig. 3.9. Gel filtration profiles of PfMATE produced homologously under control of two different promoters. Left, purification profile of PfMATE produced under control of the *gdh* promoter. Samples were solubilized with 2% (w/v) β -DDM at 40 °C (black), 3% (w/v) β -DDM at 40 °C (blue) and 2% (w/v) β -DDM at 4 °C (red). Right, purification profile of PfMATE produced under control of the *pf0613* promoter. Membrane protein solubilization was performed with 2% (w/v) β -DDM at 50 °C (black), 2% (w/v) β -DDM at 40 °C (blue) and 2% (w/v) β -DDM at 4 °C (red).

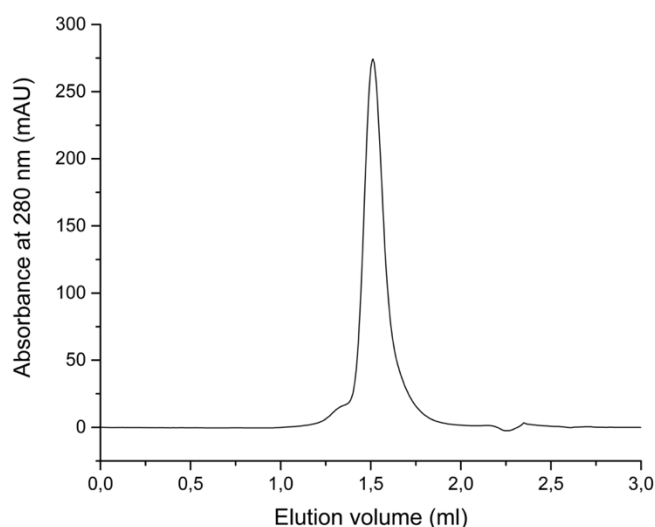


Fig. 3.10. Gel filtration profile of PfMATE produced homologously. The purification profile of PfMATE produced under control of the *gdh* promoter in the buffer 20 mM HEPES-NaOH (pH 8.0), 15 mM NaCl, 0.06% (w/v) Cymal-6.

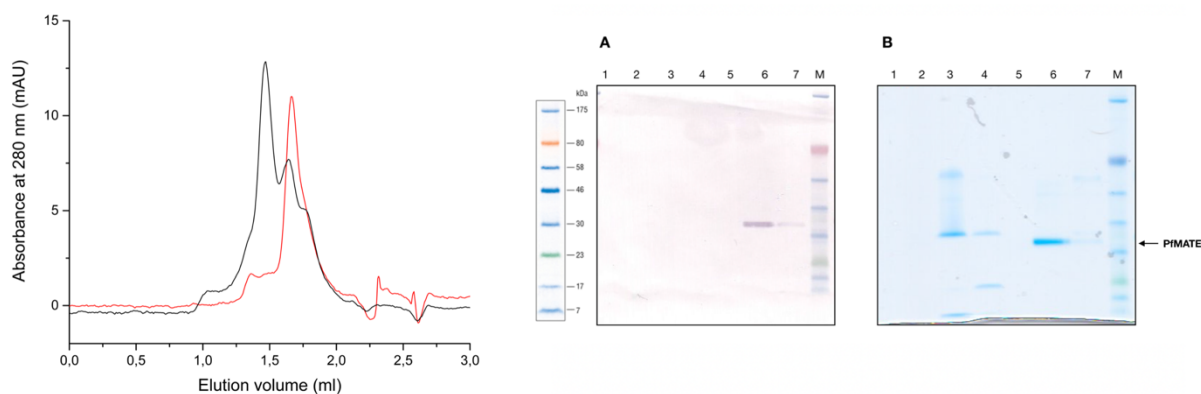


Fig. 3.11. Western Blot and SDS-PAGE gel after gel filtration. Left, gel filtration profiles of PfMATE produced under control of the *pf0613* promoter solubilized with 2% (w/v) β -DDM at 40 °C. Black line represents the elution fraction E1, whereas the red line depicts the elution fraction E2. (A) Western Blot and (B) SDS-PAGE results of the purified samples after gel filtration. Lanes 1-5 correspond to elution fractions of E2. Lanes 6-7 correspond to elution fractions of E1.

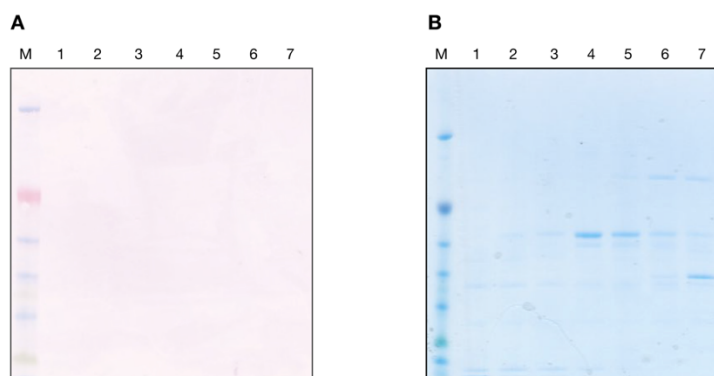


Fig. 3.12. Western Blot and SDS-PAGE gel after gel filtration. (A) Western blot and (B) SDS-PAGE results. Lanes 1-7 correspond to the elution fractions of the sample obtained after co-expression of *pf0708* with *pf0709* under control of the *gdh* promoter.

Neither quantity nor purity of PfMATE was improved by co-expression of *pf0708* with *pf0709*, as shown in Fig. 3.12, representing Western Blot and SDS-PAGE results of the purified sample. Purification of PfMATE produced by the *pf0708* expression under control of the *gdh* promoter resulted in a slightly higher yield and purity than the protein produced under control of the *pf0613* promoter. Therefore, the construct containing the *gdh* promoter without *pf0709* was typically used for the homologous expression of *pf0708* in *P. furiosus*.

3.1.4. Crystallization of PfMATE produced homologously

After size exclusion chromatography, protein concentration was quantified by measuring the absorbance at 280 nm using the NanoDrop spectrophotometer and then calculating the molarity from the Beer-Lambert equation ($A=E*b*c$, where A, absorbance at 280 nm; E, extinction coefficient ($46870\text{ M}^{-1}\text{cm}^{-1}$ for PfMATE); b, path length (cm); c, protein concentration (M)). Afterwards, the purified PfMATE sample was subjected to crystallization trials. The major hurdle was associated with the final concentration of the protein in the range of 0.2-2 mg/ml, being extremely low especially for the crystallization using the LCP method where protein concentrations in the range of 10-20 mg/ml are preferred. Previously published studies on PfMATE produced heterologously in *E. coli*, showed that crystallization by the LCP method yielded well-diffracting crystals. Therefore, apart from initial trials by the vapor diffusion method, the LCP technique has been mainly applied for crystallization attempts of PfMATE produced homologously. For that purpose, commercially available as well as in-house designed crystallization screens were used. The purified protein sample was mixed with the molten monoolein in a 2:3 (v/v) ratio and dispensed on 96-well crystallization plates with the reservoir solutions. After incubation in the CrystalMation robot at 22 °C, each droplet was inspected every few days manually under a light microscope with polarizing optics. Changing colors during rotation of the polarizing filter lens allows detection of birefringent crystals. Successful crystal growth was achieved mostly by usage of the MemMeso HT-96 and LCP_Screen_PEG400_Citrate screens. Precipitating agents in the cocktail solution, such as a variety of polyethylene glycols (PEGs) as well as sodium citrate buffer at acidic pH were found to be advantageous for the crystal growth of PfMATE. All grown microcrystals (typically up to 20-40 μm in size) were obtained at relatively low pH ranging from 5 to 6.

At the final stage of my PhD project, an alternative approach of LCP crystallization was published, namely the cubicon method (Fig. 3.13), which involves several rounds of protein

reconstitution in order to raise the final protein concentration in the lipid cubic phase¹⁴⁸. This approach led to the reconstitution of the homologously produced PfMATE in the lipid bilayer at concentrations up to 4 mg/ml and yielded in a few crystal hits. The light microscopic images of the generated crystals and composition of the corresponding reservoir solutions are presented in Fig. 3.14.

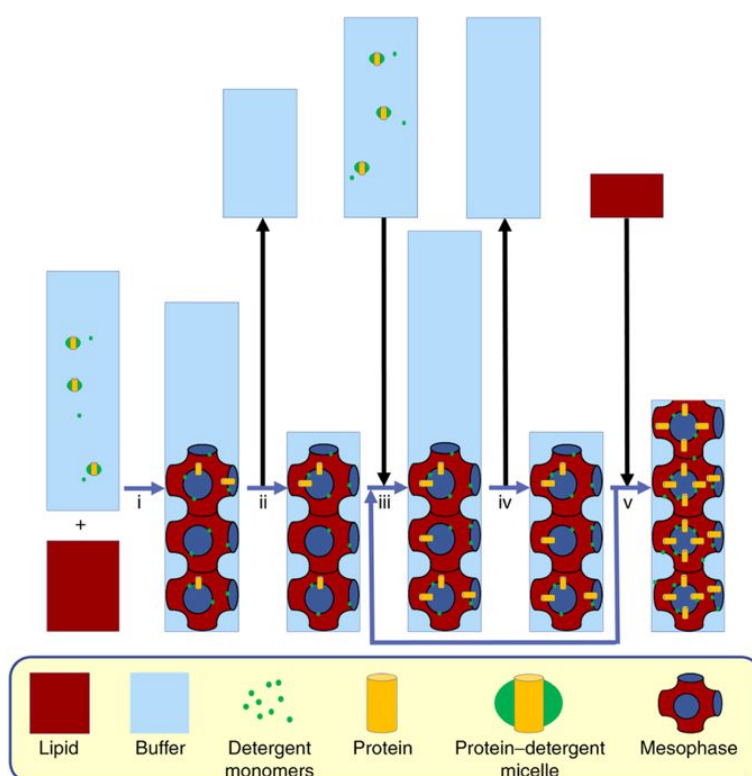


Fig. 3.13. The cubicon method. Step i: Protein solution and lipid are mixed to form the cubic mesophase. Step ii: Most of the excess aqueous solution is removed. Step iii: Fresh protein solution is mixed with the mesophase so that more protein is being reconstituted. Step iv: Most of the excess aqueous solution is removed. Steps iii and iv: multiple rounds of reconstitution and an increase of protein concentration in the mesophase. Step v: a small amount of lipid is added to produce a homogeneous, optically clear cubic phase. Image adapted from¹⁴⁸.

Crystals were harvested and stored in liquid N₂ until the measurements at the synchrotron Swiss Light Source (SLS) in Villigen, Switzerland. The best quality crystal, which grew in the reservoir solution composed of 0.05 M zinc acetate, 0.1 M MES (pH 6.0), 6.6% (w/v) PEG 8000, 6% (v/v) ethylene glycol, produced a few diffraction spots only up to 10 Å. As a

consequence of the cumulative factors, such as low yield of the homologous protein production being insufficient for the extensive crystallization experiments, limited resources for *P. furiosus* cultivation and lack of multiple genetic tools for the genetic manipulation of this archaeon and optimization of gene expression, another strategy was applied to work simultaneously on PfMATE produced heterologously in *E. coli*.

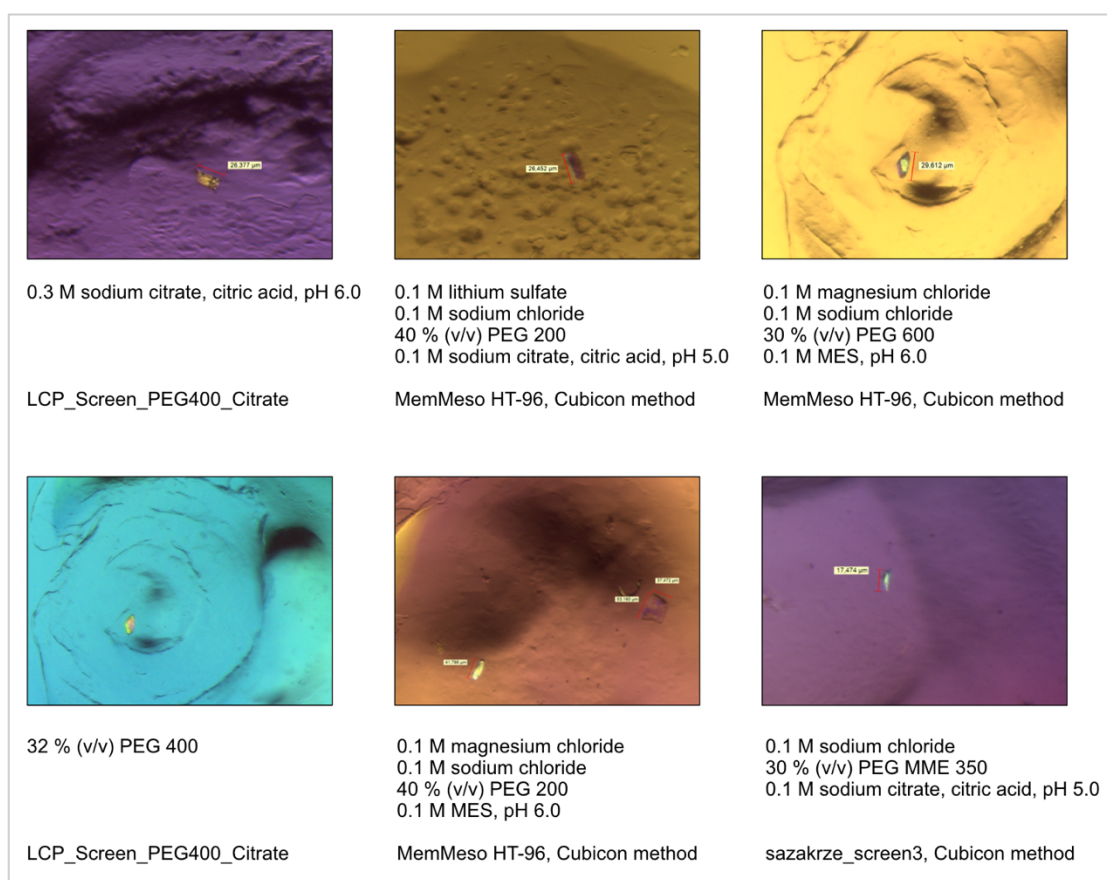


Fig. 3.14. Crystals of PfMATE produced homologously. The crystallization screen and the composition of the reservoir solution in the corresponding crystallization well are listed below each image of the crystal.

3.2. PfMATE produced heterologously

3.2.1. Heterologous expression of *pf0708* in *E. coli*

Generation of the pBAD-PfMATE vector containing the gene *pf0708* coding for PfMATE had been previously performed by Cornelia Münke. Plasmid DNA (pBAD-PfMATE) was transformed into *E. coli* TOP10 competent cells, which were subsequently spread on agar plates containing carbenicillin for the selection of positive clones. After incubation at 37 °C overnight, a single colony was transferred into LB medium and the culture was used as an inoculum for 2-liter of LB medium in 5 liter unbaffled culture flasks. Once an optical density at 600 nm (OD_{600}) reached 0.6-0.8, *pf0708* gene expression was induced by addition of 0.02% (w/v) L-arabinose and the cells were incubated with shaking at 37 °C for 2-4 h.

Due to the presence of the C-terminal HisTag in PfMATE, production of the protein was confirmed by Western Blot using the monoclonal anti-polyhistidine-alkaline phosphatase coupled antibody. As shown in Fig. 3.15, a purple band is observed exclusively in the lane corresponding to the cell lysate containing pBAD-PfMATE after induction of gene expression. Afterwards, *E. coli* TOP10 cells were centrifuged and the cell pellet obtained was stored at -20 °C for subsequent experiments.

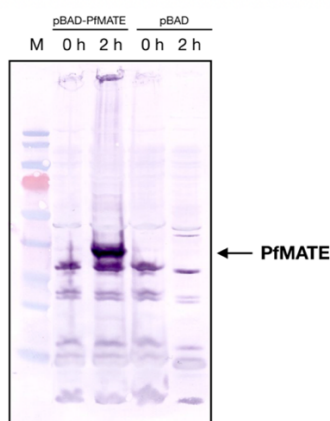


Fig. 3.15. Western Blot of the *E. coli* TOP10 cell lysate after heterologous gene expression. Each lane presents the *E. coli* the TOP10 cell lysate at 0 h or 2 h after induction of gene expression by addition of L-arabinose to a final concentration of 0.02% (w/v). pBAD-PfMATE, the expression vector containing *pf0708*; pBAD, the empty vector without *pf0708*.

3.2.2. Purification of PfMATE produced heterologously

The *E. coli* TOP10 cell pellet was resuspended in a buffer composed of 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 5 mM MgCl₂, pinch of DNaseI and 1 mM PMSF. After cell lysis via three passages through the Microfluidizer, the disrupted cells were centrifuged and the membrane fraction was isolated by ultracentrifugation. The membrane pellet was resuspended in a buffer composed of 20 mM HEPES-NaOH (pH 8.0) and 300 mM NaCl. Afterwards, the BCA assay was used to determine the protein concentration in the membrane fraction, which was aliquoted into 5-10 ml aliquots of 20 mg/ml protein and stored at -80 °C for further experiments. The membrane proteins were solubilized typically with 2% (w/v) β -DDM by incubation at 4 °C in the rotator. Subsequently, PfMATE produced heterologously was purified applying a two-step purification workflow.

Firstly, affinity chromatography using His tag/nickel was performed, followed by size exclusion chromatography. Protein elution was monitored by measuring the absorbance at 280 nm. As shown in Fig. 3.16, the affinity chromatography profile exhibits a peak eluted at 500 mM imidazole. The corresponding protein fractions were pooled, concentrated with a centrifugal filter device and applied on a Superdex 200 10/300 GL column previously equilibrated with buffers composed either of 20 mM HEPES-NaOH (pH 8.0), 150 mM NaCl, 0.03% (w/v) β -DDM or 20 mM HEPES-NaOH (pH 8.0), 15 mM NaCl, 0.06% (w/v) Cymal-6. The final gel filtration step yielded a homogenous protein sample as illustrated by the presence of a single peak in the purification profile (Fig. 3.17).

Protein fractions corresponding to the monodisperse peak were pooled and analyzed by Western Blot and SDS-PAGE. Presence of the His-tagged protein was indicated by a purple band shown on the Western Blot membrane, whereas the electrophoretic analysis revealed a high purity and homogeneity of the PfMATE sample produced heterologously. The protein concentration was quantified by measuring the absorbance at 280 nm using the NanoDrop

spectrophotometer as described in the Materials and Methods section 2.3. The total amount of the heterologously produced protein (1-2 mg after each two-step purification) was sufficient for extensive crystallization trials.

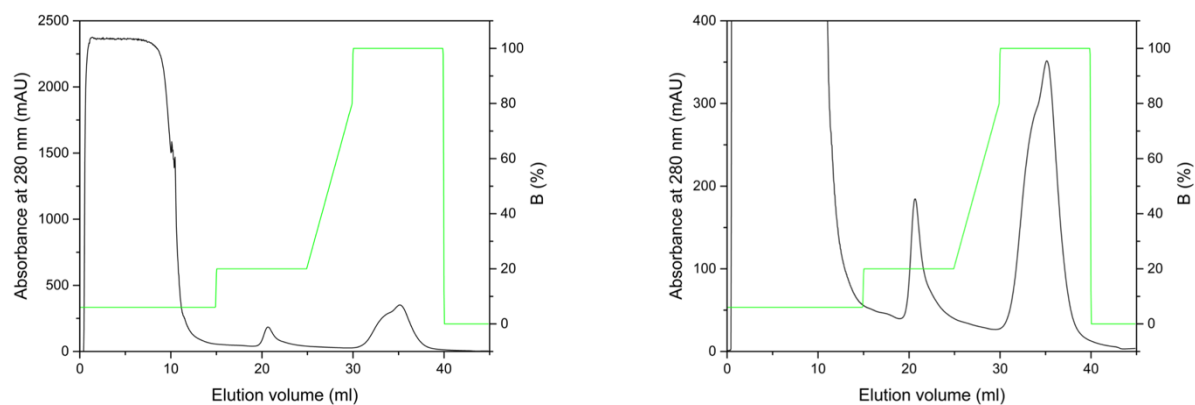


Fig. 3.16. Affinity chromatography column elution profiles of PfMATE produced heterologously. The black line represents the absorbance at 280 nm, whereas the green line shows the concentration gradient of buffer B composed of 20 mM HEPES-NaOH (pH 8.0), 300 mM NaCl, 0.05% (w/v) DDM and 500 mM imidazole. Left, full purification profile. Right, close up of the purification profile between 0-400 mAU, including the elution peak between 30-40 ml.

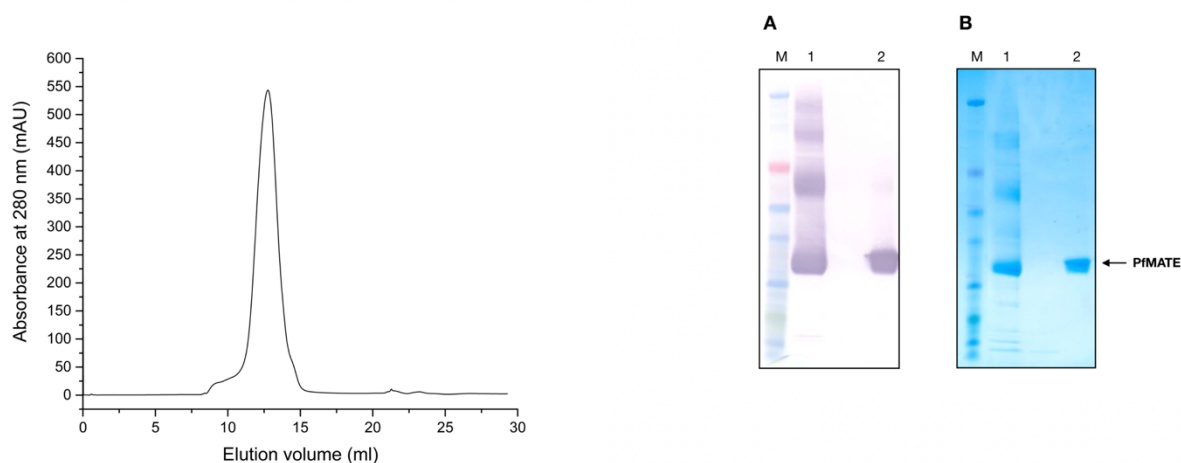


Fig. 3.17. Gel filtration profile of the heterologously produced PfMATE. Left, gel filtration profile of PfMATE in a buffer composed of 20 mM HEPES-NaOH (pH 8.0), 15 mM NaCl, 0.06% (w/v) Cymal-6. (A) Western Blot and (B) SDS-PAGE gel. Lane 1 corresponds to the sample after affinity chromatography, whereas lane 2 represents the sample after gel filtration.

3.2.3. Blue native PAGE results

After size exclusion chromatography, the purified PfMATE produced homologously in *P. furiosus* as well as heterologously in *E. coli* was analyzed by the size of a detergent micelle using blue native PAGE method to assess its molecular size and aggregation state. In this experiment each sample purified in the buffer containing detergents 2-3 times above their CMC (critical micelle concentration), either 0.06% (w/v) Cymal-6 or 0.03% (w/v) β -DDM, was loaded onto the 4-16% gradient native gel. Electrophoresis was performed at 4 °C in tank filled with the anode and cathode buffer.

The molecular weight of the monomeric form of PfMATE calculated by the ExpASY software is 49.2 kDa. Regardless of the expression system, the results of the native PAGE display

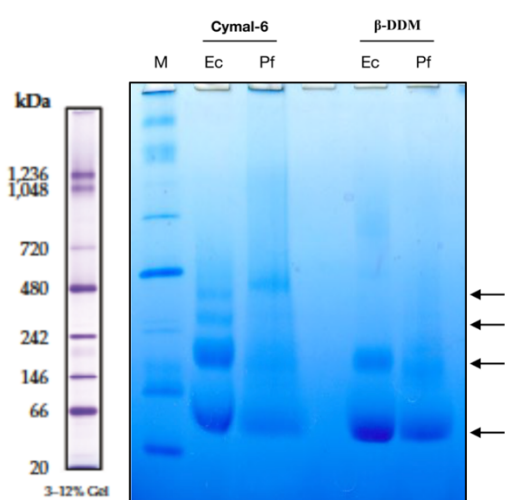


Fig. 3.18. Native PAGE gel of PfMATE after gel filtration. Samples were purified in the presence of either 0.06% (w/v) Cymal-6 or 0.03% (w/v) β -DDM. Ec, PfMATE produced heterologously in *E. coli*; Pf, PfMATE produced homologously in *P. furiosus*. The black arrows indicate bands, which may correspond to different oligomeric states of PfMATE.

multiple aggregation states of PfMATE. For PfMATE purified in the presence of 0.03% (w/v) β -DDM two major bands with an estimated molecular weight at around 100 kDa and 200 kDa were observed. In case of PfMATE in the buffer containing 0.06% (w/v) Cymal-6 two additional bands between the protein ladder bands corresponding to 242 kDa and 480 kDa were present (Fig. 3.18).

These results suggest that different detergents may have an impact on the oligomeric state of this MATE transporter or modulate its electrophoretic migration in the native gel. The amount of the detergent molecules bound to the protein varies depending on the chemical nature, length and binding capacity of the detergent as well as on the protein hydrophobic surface area.

3.2.4. Mass spectrometry based lipidomics of *P. furiosus* lipids

Considering the unique composition of the archaeal lipids, *P. furiosus* WT was cultivated under anaerobic conditions along with the elevated temperature (95 °C), and the cell pellet was subjected to lipid extraction using the MTBE method¹¹⁸. After mass spectrometry based lipidomics, the total lipid extract was used for co-purification and co-crystallization with PfMATE. Lipid extraction and subsequent lipidomics was conducted within a collaboration with the group supervised by PD Dr. Harald Köfeler at the Medical University of Graz, Austria.

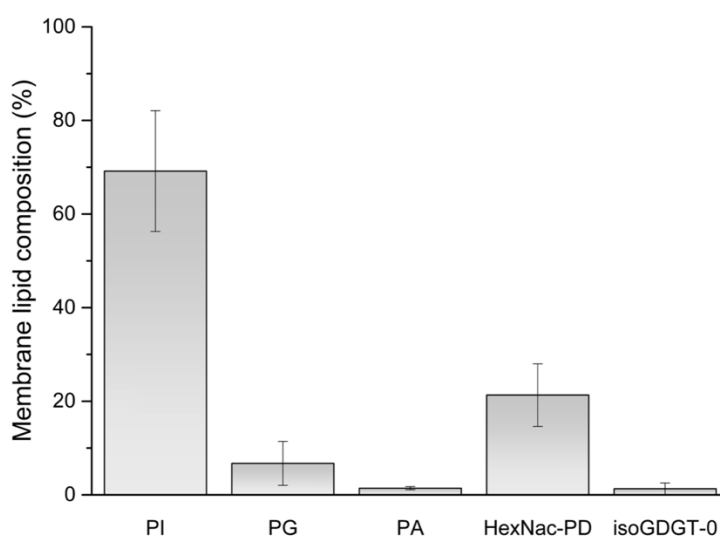


Fig. 3.19. Lipid composition of *P. furiosus* WT analyzed by mass spectrometry based lipidomics. PI, diphytanyl phosphatidyl inositol; HexNac-PD, diphytanyl phosphatidyl *N*-acetyl hexose; PG, diphytanyl phosphatidyl glycerol; PA, diphytanyl phosphatidic acid; isoGDGT-0, isoprenoidal glycerol dialkyl glycerol tetraether.

Lipidomic analysis of the isolated *P. furiosus* lipids led to identification of the following archaeal lipid species: DPI (diphytanyl phosphatidyl inositol), HexNac-PD (diphytanyl phosphatidyl *N*-acetyl hexose), DPG (diphytanyl phosphatidyl glycerol), DPA (diphytanyl phosphatidic acid) and isoGDGT-0 (isoprenoidal glycerol dialkyl glycerol tetraether) (Fig. 3.19). To investigate the mechanistic and structural effects of the native lipids on PfMATE, crystallization experiments in the presence and absence of *P. furiosus* lipid extract were performed.

3.2.5. Crystallization of PfMATE produced heterologously

Heterologously purified PfMATE was crystallized mainly using the LCP method. PfMATE (10-20 mg/ml) in 20 mM HEPES-NaOH (pH 8.0), 15 mM NaCl and 0.06% (w/v) Cymal-6 was filtered through a 0.20 μm filter before being mixed with the molten monoolein with a 2:3 (v/v) protein:monoolein ratio using a coupled syringe mixer. 96-well crystallization trays were set up using the ProCrys Meso lipidic mesophase dispenser. The size of the precipitant and protein-monoolein solution drops were 1.5 μl and 100 nl, respectively. Plates were stored at 295 K in the incubator of the CrystalMation system. Analysis of the images under UV-light taken automatically by the CrystalMation robot was useful only for a limited number of the crystallization drops. Typically, microcrystals were difficult to localize optically, especially the ones embedded deeply in the thick toothpaste-like lipidic cubic phase. Therefore, each well of the plate was inspected manually under the light microscope with polarizing optics. Changing colors during rotation of the polarizing filter lens allows assessment of the birefringent crystals. Usage of the sazkrze_screen7, the LCP_Screen_Citrate_PEG and MemeSo HT-96 crystallization screens resulted in multiple crystal hits. Similar to the crystallization results of the homologously produced PfMATE, the crystals of the heterologously produced protein sample grew with reservoir solutions containing buffer at acidic pH between 5 and 6. After crystals matured to their full size, they were harvested and directly flash-frozen in liquid N_2 without soaking in a cryoprotectant owing to the presence of the cubic phase forming lipid, monoolein, which acts as cryoprotectant. Afterwards, the crystals were measured at the synchrotron SLS. The best diffracting crystals leading to the structure determination of PfMATE in two alternate conformations: an inward and outward-facing state, were obtained using the MemMeSo HT-96 screen. For the inward-facing conformation, the protein sample after the affinity chromatography step was incubated with the total lipid extract from *P. furiosus*, co-purified on a Superdex 200 10/300 GL column (Fig. 3.20) and then crystallized. Crystals of PfMATE in the inward-facing conformation were grown in 0.1 M

sodium chloride, 0.1 M magnesium chloride, 30% (v/v) PEG600 and 0.1 M sodium citrate (pH 5.0). PfMATE in the outward-facing state was captured under almost identical conditions (0.1 M sodium chloride, 30% (v/v) PEG500 DME and 0.1 M sodium citrate (pH 5.0)), however, in absence of the native lipids from *P. furiosus*.

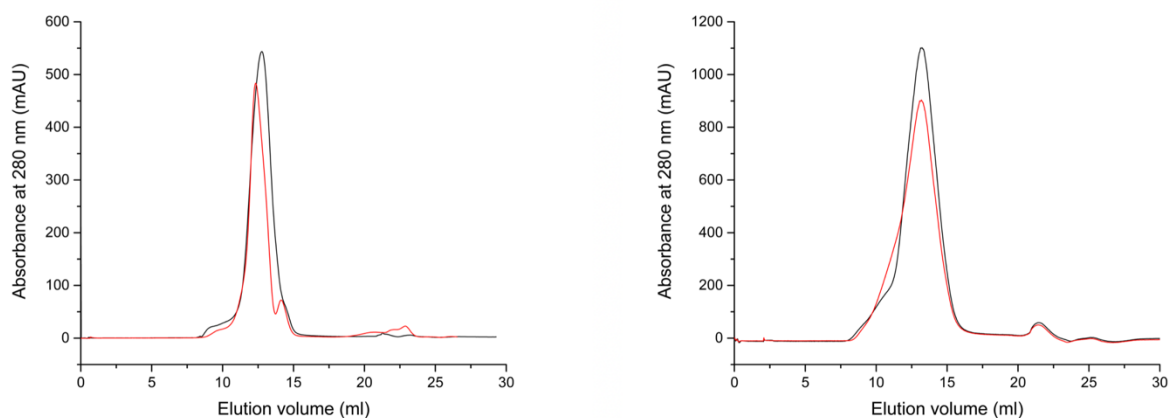


Fig. 3.20. Gel filtration of the heterologously produced PfMATE co-purified with the lipids. The purification profile of PfMATE co-purified with *P. furiosus* total lipid extract on the left side and with NBD-PG on the right side. The black line represents the sample purified without addition of lipids, whereas the red line illustrates the sample purified in the presence of the respective lipid.

Except for the endogenous lipids from *P. furiosus*, the crystallization trials were also performed in the presence of different negatively charged PG (phosphatidylglycerol) lipid species, such as diesters: 16:0-12:0 NBD-PG (7-nitro-2-1,3-benzoxadiazol-4-yl PG), 16:0-18:1 POPG (palmitoyl-2-oleyl PG) and diether 16:0-18:1 PG. Also, NAC6 (*N*-acetyl-D-glucosamine 6-phosphate), whose chemical structure is similar to that of the lipid head group of HexNac-PD (diphytanyl phosphatidyl *N*-acetyl hexose) detected by the mass spectrometry based lipidomics of the *P. furiosus* lipid extract. Each of the lipid species was either mixed with the protein sample after affinity chromatography and co-purified on a Superdex 200 column upon the crystallization trials or mixed after the final gel filtration step in 1:5 or 1:10 protein:lipid molar ratio and then crystallized (Fig. 3.19). Examples of the crystals formed by this approach are depicted in figures 3.21-3.25.

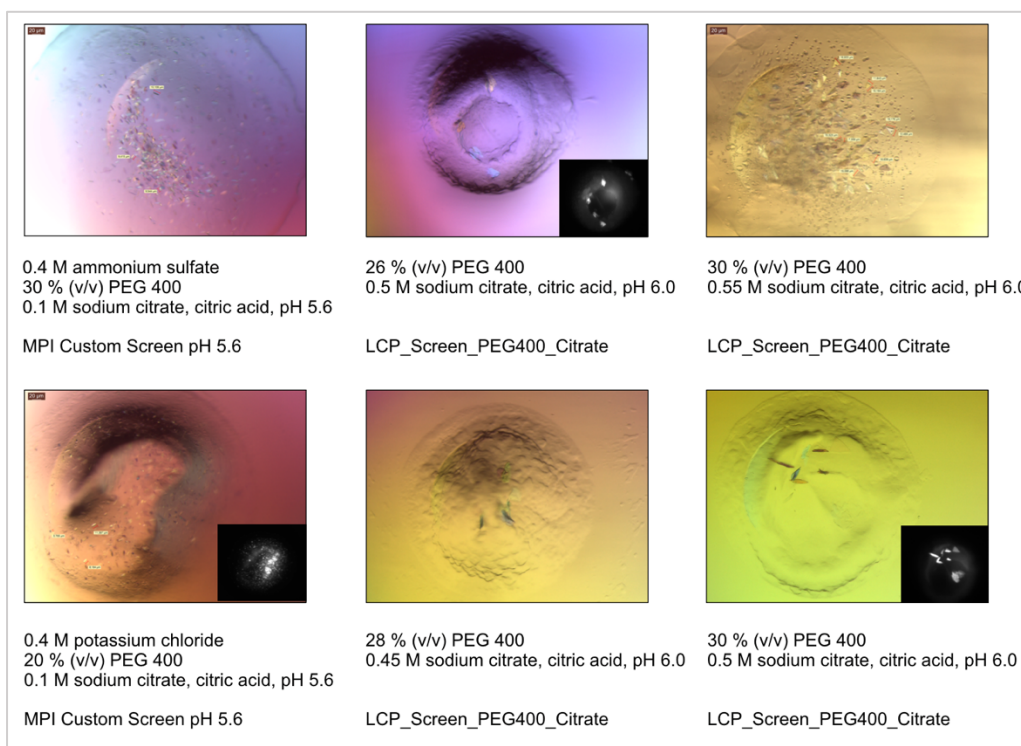


Fig. 3.21. Crystals of PfmMATE produced heterologously after co-crystallization with the total lipid extract from *P. furiosus*. The crystallization screen and composition of the reservoir solution in the corresponding crystallization well are listed below each image of the crystal. Images of the crystals under UV-light are also shown.

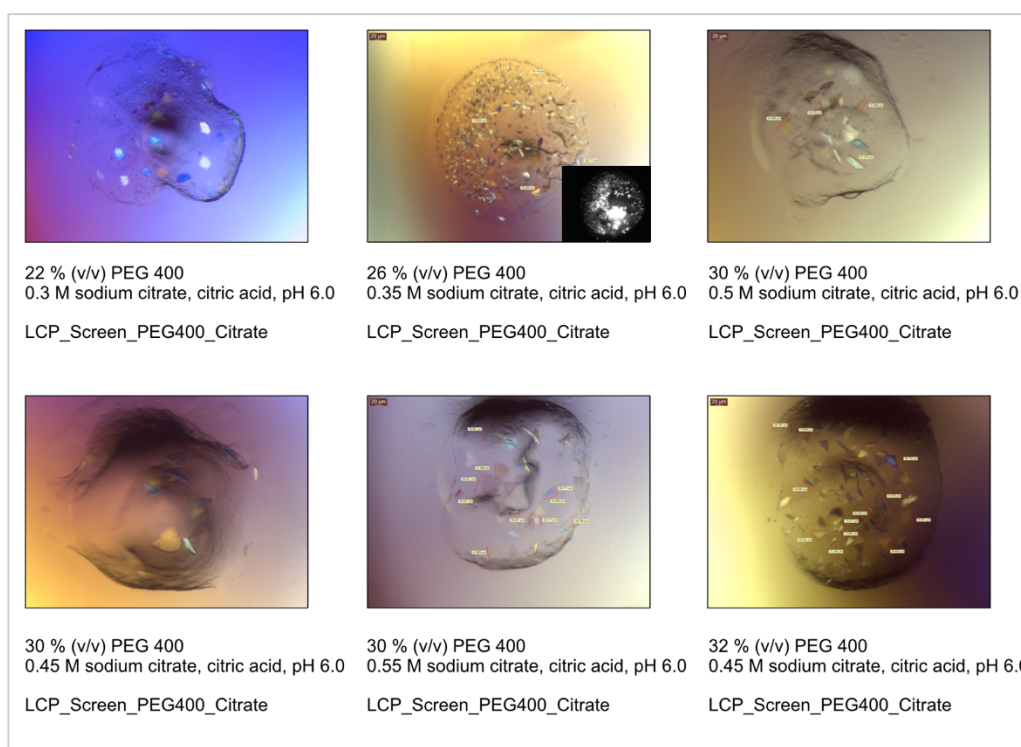


Fig. 3.22. Crystals of PfmMATE produced heterologously after co-crystallization with diether-PG. The crystallization screen and composition of the reservoir solution in the corresponding crystallization well are listed below each image of the crystal. Images of the crystals under UV-light are also shown.

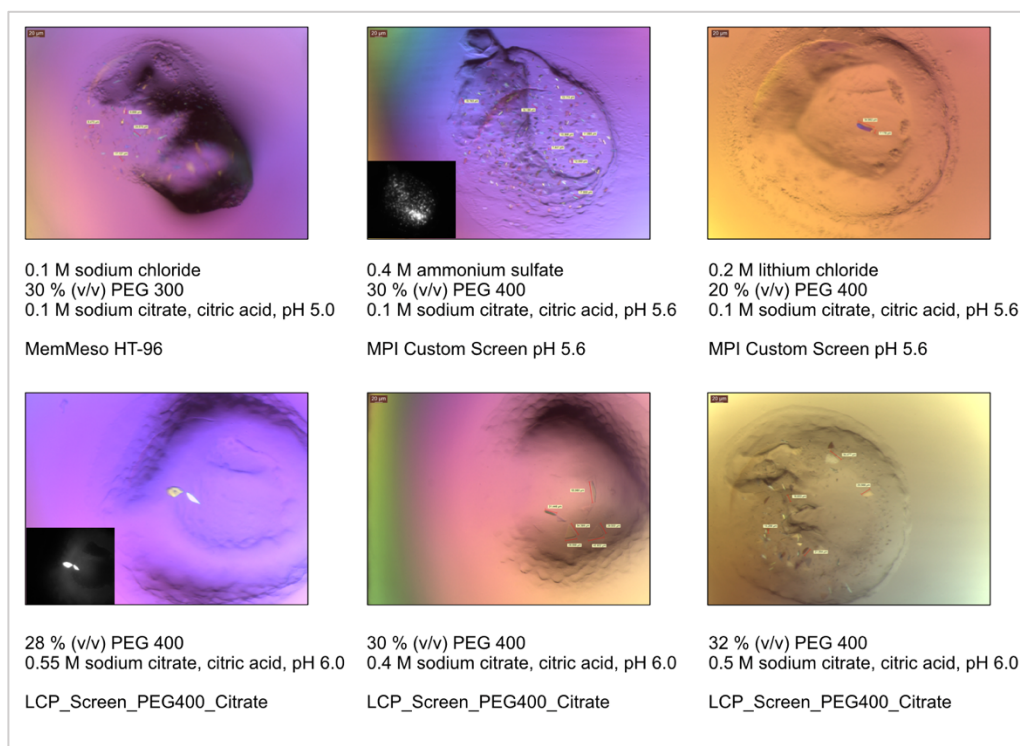


Fig. 3.23. Crystals of PfMATE produced heterologously after co-crystallization with NBD-PG. The crystallization screen and composition of the reservoir solution in the corresponding crystallization well are listed below each image of the crystal. Images of the crystals under UV-light are also shown.

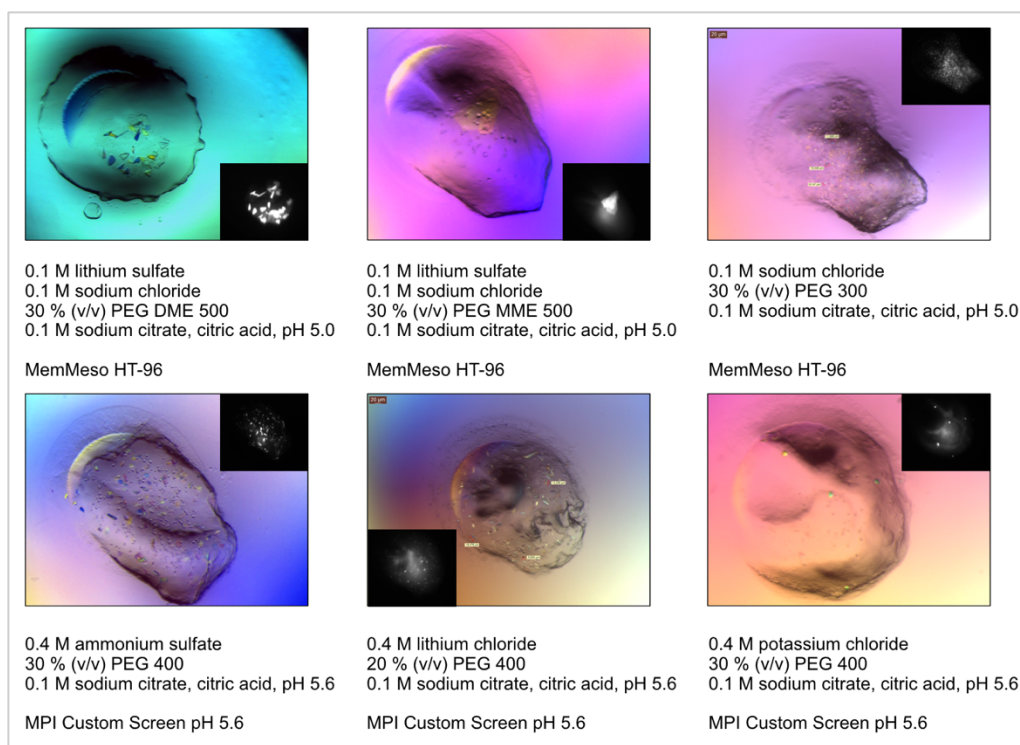


Fig. 3.24. Crystals of PfMATE produced heterologously after co-crystallization with NAC6. The crystallization screen and composition of the reservoir solution in the corresponding crystallization well are listed below each image of the crystal. Images of the crystals under UV-light are also shown.

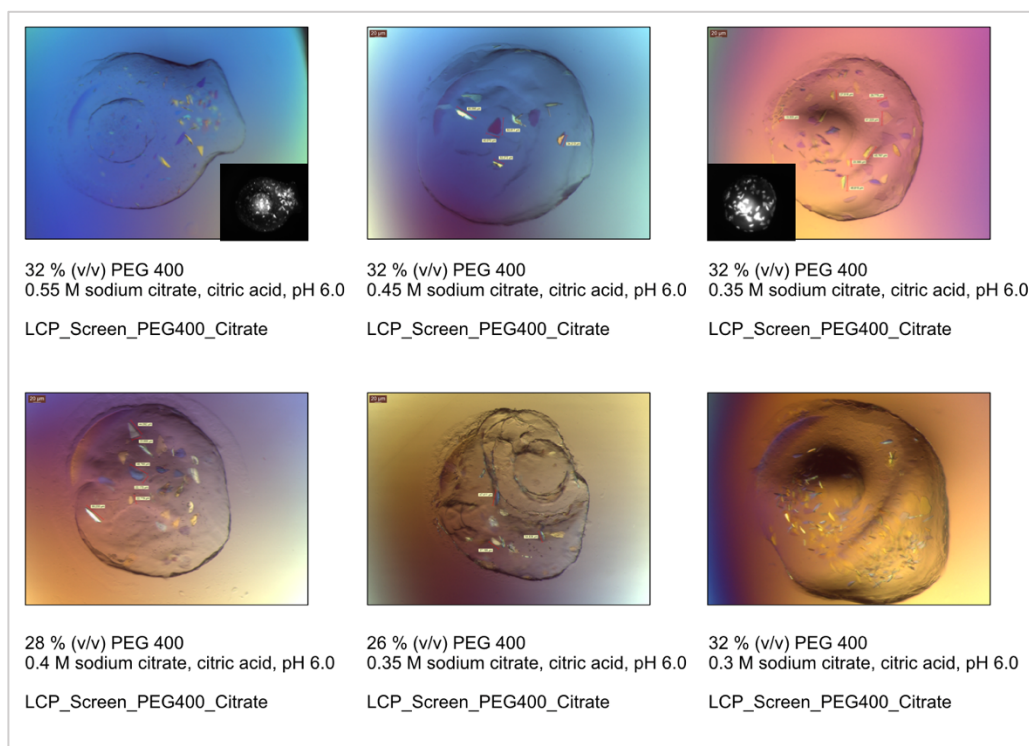


Fig. 3.25. Crystals of PfMATE produced heterologously after co-crystallization with DOPG. C The crystallization screen and composition of the reservoir solution in the corresponding crystallization well are listed below each image of the crystal. Images of the crystals under UV-light are also shown.

Noteworthy, in case of the purification of PfMATE with the fluorescently labeled lipid NBD-PG, the eluted protein fractions were yellow due to the color of the NBD fluorophore, implying that the lipid molecules are co-eluted with the protein. NBD-PG alone elutes with a higher retention volume. Crystals of the protein mixed with the aforementioned specific lipid species were measured at the synchrotron SLS. Some of them resulted in diffraction patterns with resolutions between 3-4 Å. Crystals of PfMATE with NAC6 led to structure determinations of up to 2.6 Å resolution. However, in all crystals PfMATE was trapped in the same outward-facing conformation as the structure obtained in the absence of lipids. Additionally, given the relatively low resolution, possible disorder and high rate of exchange of lipid molecules, no peak in the electron density maps could be unambiguously assigned to a bound lipid molecule. These structures did not provide any unique structural features or new insights into interplay between PfMATE and lipids, and therefore are omitted in this dissertation.

3.2.6. Data collection and structure determination

PfMATE crystals were transported in dewars filled with liquid N₂ and measured at the XII/X10SA beamline at the synchrotron Swiss Light Source (SLS) in Villigen, Switzerland. Due to the fact that the microcrystals were small and embedded in the viscous lipidic cubic phase, it was sometimes difficult to localize them optically. Therefore, rastering by the X-ray beam was applied for screening of the crystals in the loops and centering them (Fig. 3.26). The best diffracting regions of the microcrystals are shown in red in the grid-scan rastering results. The high redundancy data collection strategy was applied to collect typically 1800 diffraction

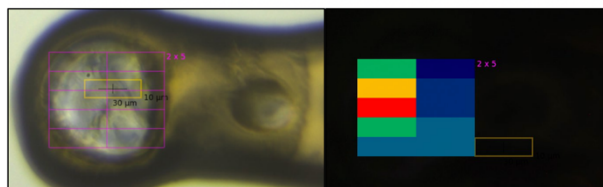


Fig. 3.26. Rastering scan of a PfMATE crystal. The left image shows the loop with the crystal embedded in the lipidic phase, which is covered with the rastering grid. The results of the rastering are depicted on the right side. The best diffracting regions of the crystal are represented in red, whereas the worst ones in blue.

was performed at 100 K using a PILATUS 6M pixel detector (Fig. 3.27). After data collection, the X-ray diffraction datasets were processed using XDS (X-ray detector software) for indexing and integration, XSCALE for merging and scaling, XDSCONV and F2MTZ for converting the reflection files from ahkl to the final mtz file.

The quality of the datasets was analyzed by Xtriage from the Phenix suite. According to the results from the UCLA diffraction anisotropy server, the collected data exhibit mild anisotropy (Fig. 3.28). Therefore, there was no need for an ellipsoidal truncation or anisotropic scaling. Data with freeR flags were generated by Uniqeify software from the CCP4 (Collaborative computational project) suite and used for the next stage of data processing.

The previously deposited PfMATE structure (PDB ID: 4MLB), specifically monomer A from the tetrameric asymmetric unit (ASU), was used as a search model for the phase generation by molecular replacement (MR). This step was performed using PhaserMR from the CCP4 suite.

The final model was subjected to manual adjustment in COOT, and subsequent iterative refinement with phenix.refine from the Phenix suite, which resulted in the structure determination of PfMATE in the outward-facing conformation at 2.8 Å resolution (space group R 3 2, space group number 155). The quality of the electron density map allowed allocation of most of the amino acid residues, with the exception of the loop regions connecting TM3 and TM4 as well as TM9 and TM10.

In case of another dataset, molecular replacement using the monomeric molecule from 4MLB as an ensemble showed that only one half of this model fits into the electron density. Therefore, the initial 4MLB model was split into two halves in the centre, the N-terminal (TMs 1-6) and the C-terminal domain (TMs 7-12), and both of them were used for MR as two ensembles. This approach led to the structure determination at 2.8 Å resolution and the detection of a novel conformational state of a MATE family transporter, an inward-facing state (space group C 2 2₁, space group number 20).

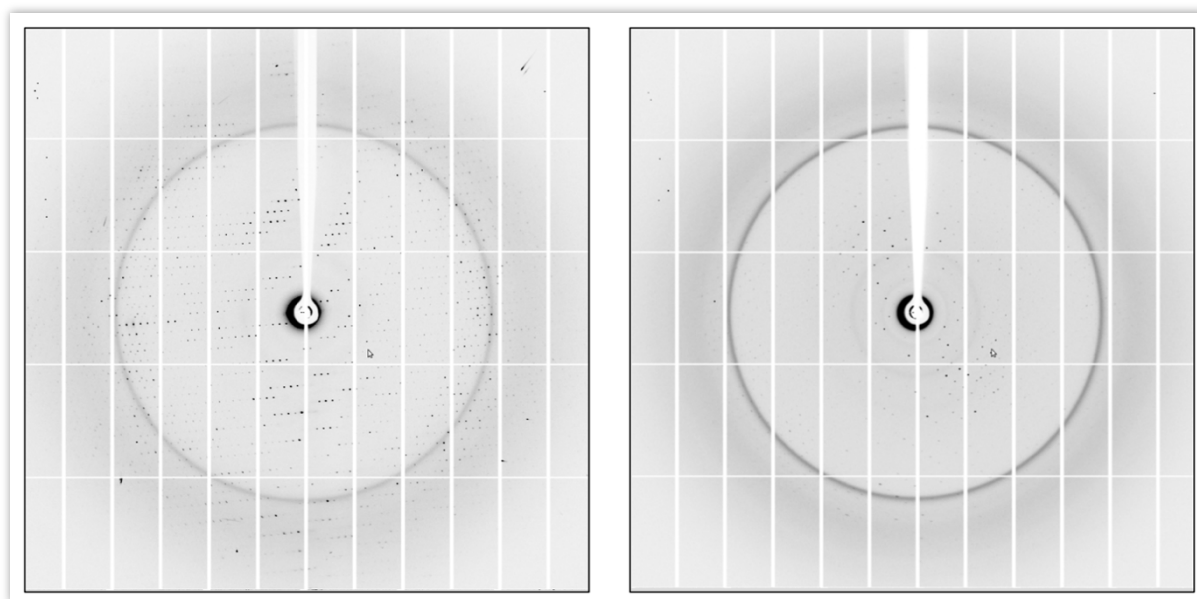
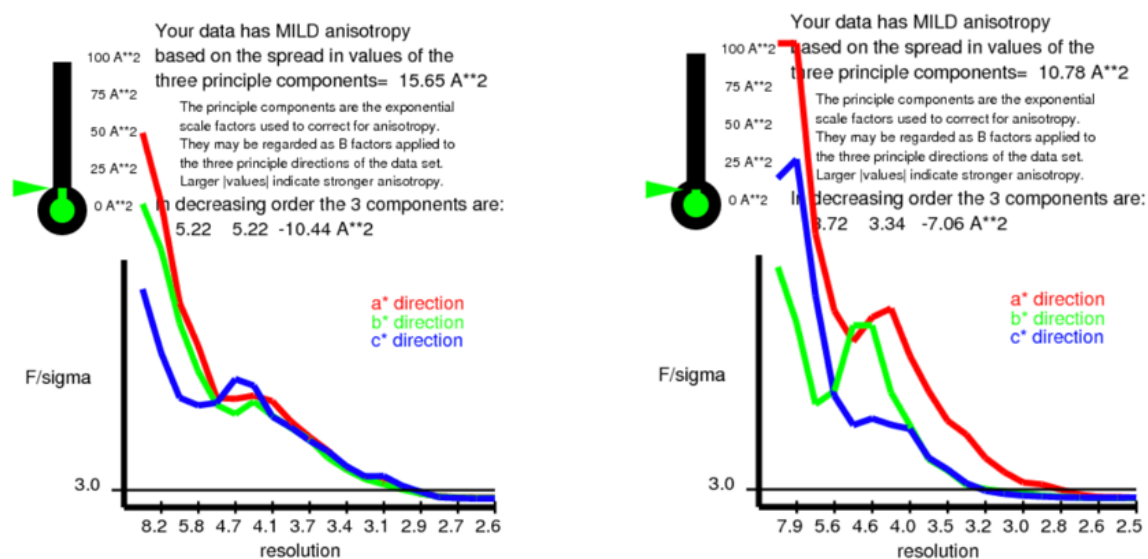


Fig. 3.27. Diffraction pattern. Diffraction pattern of the crystals capturing PfMATE in the outward-facing conformation on the left and the inward-facing conformation on the right. Images were generated using iMosflm by overlapping 20 diffraction images.



The intensity statistics look normal, indicating that the data are not twinned.
Translational NCS does not appear to be present.
Ice rings do not appear to be present.
The fraction of outliers in the data is less than 0.1%.
The data are not significantly anisotropic.
The resolution cutoff appears to be similar in all directions.
The overall completeness in low-resolution shells is at least 90%.
The completeness is 99.41%.

The intensity statistics look normal, indicating that the data are not twinned.
Translational NCS does not appear to be present.
Ice rings do not appear to be present.
The fraction of outliers in the data is less than 0.1%.
The data are not significantly anisotropic.
The resolution cutoff appears to be similar in all directions.
The overall completeness in low-resolution shells is at least 90%.
The completeness is 97.34%.

Fig. 3.28. Diffraction anisotropy plot and Xtriage results. Assessment of the data quality by Xtriage (lower images) and diffraction anisotropy obtained using the UCLA diffraction anisotropy server (upper images) of the outward (left) and inward-facing (right) structure datasets. The plots illustrate intensity falloff with resolution for each direction of the crystal, a*, b*, c*.

Except for the perturbed fragment of TM1, all TMs were assigned unambiguously. After multiple cycles of manual model rebuilding in COOT, iterative refinement was performed with phenix.refine. Crystallographic data processing led to the PfmATE structure determination in two alternate conformations, an outward and an inward-facing state. The data collection and refinement statistics for these structures are summarized in the Tab. 3.1, whereas the crystal packing of PfmATE with one molecule in ASU is shown in Fig. 3.29.

Crystallization of PfmATE in the absence of the *P. furiosus* lipid extract leading to the structure determination in the outward-facing state, and the inward-facing state obtained at almost identical crystallization conditions, but in presence of the native lipids isolated from *P. furiosus*,

hints at a crucial role of the native lipids on the conformational stability of the PfMATE transporter.

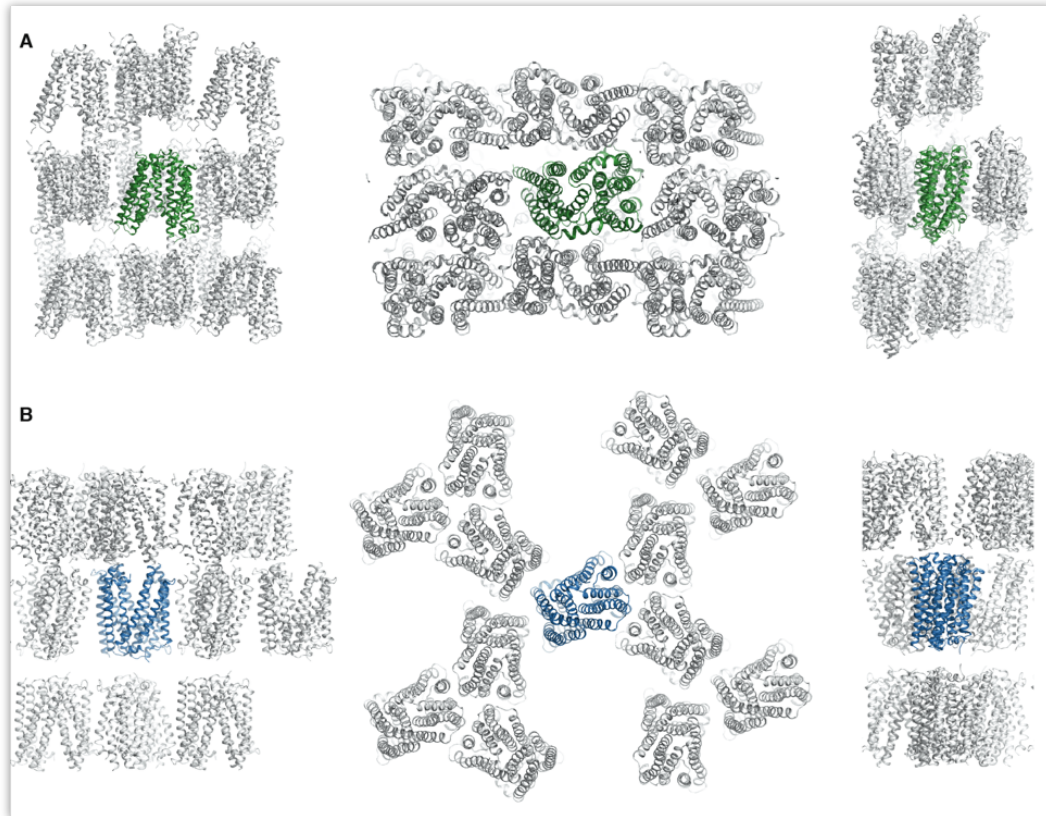


Fig. 3.29. Crystal packing of PfMATE. Packing of PfMATE in the two distinct conformations in their respective crystals. The protein molecules are shown a ribbon models. (A) The inward-facing conformation (PDB ID: 6FHZ); crystals grew in space group $C 2 2 2_1$. (B) The outward-facing conformation with the straight helix TM1 (PDB ID: 6GWH); crystals grew in space group $R 3 2$.

Table 3.1. Data collection and refinement statistics

	PfMATE Outward-facing structure	PfMATE Inward-facing structure
Crystallization method	LCP	LCP
pH of crystallization	5.0	5.0
Data collection		
Space group	R 3 2	C 2 2 2 ₁
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	95.83, 95.83, 369.11	107.12, 134.31, 69.71
α , β , γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 90.00
Resolution (Å)	47.92-2.80 (2.93-2.80)*	43.7-2.80 (2.90-2.80)*
<i>R</i> _{merge}	0.142 (0.973)	0.223 (1.669)
CC _{1/2}	0.995 (0.532)	0.996 (0.296)
<i>I</i> / σ <i>I</i>	7.23 (1.05)	8.80 (1.05)
Completeness (%)	97.70 (89.90)	99.80 (99.80)
Redundancy	4.2 (3.1)	8.3 (5.5)
Refinement		
Resolution (Å)	19.97-2.80 (2.90-2.80)	19.80-2.80 (2.90-2.80)
No. reflections	16161 (1445)	12676 (1263)
<i>R</i> _{work} / <i>R</i> _{free}	0.258/0.280	0.229/0.274
No. atoms		
Protein	3311	3264
Average <i>B</i> -factors (Å ²)	77.0	64.0
R.m.s. deviations		
Bond lengths (Å)	0.001	0.001
Bond angles (°)	0.402	0.360
PDB ID	6GWH	6FHZ

*Values in parentheses represent the highest resolution shell.

3.2.7. Crystal structure of PfMATE in the outward facing conformation

The PfMATE crystal structures in the outward-facing state had been previously determined at high pH (7.0-8.0) and low pH (about 6.0). TM1 in the low pH structure is slightly bent compared to the straight helix TM1 in the high pH structure. To investigate the relationship between pH and TM1 structural alterations, PfMATE was crystallized at an even lower pH (5.0). Consequently, the structure of PfMATE in the outward-facing state was determined at up to 2.8 Å resolution (Fig. 3.30). The overall PfMATE structure is essentially identical to the outward-facing conformation with a straight TM1 obtained previously at the elevated pH, where 12 transmembrane helices are arranged in two bundles of six helices each. This architecture adopts a two-fold pseudo-rotational symmetry between two halves: the N-terminal domain (TMs 1-6) and C-terminal domain (TMs 7-12). Noteworthy, despite of the acidic pH of the crystallization conditions, helix TM1 in the presented below PfMATE structure remained straight and no structural fluctuation was observed. Structural alignment between the low pH structure and the previously presented high pH structure (PDB ID: 3VVN) revealed a high similarity with a global Ca root mean square deviation (r.m.s.d.) of 0.86 Å. Thus, the structural results are consistent with the notion that conformational rearrangement of the helix TM1 is most likely independent of pH. A previously deposited PfMATE structure in the outward-facing state (PDB ID: 4MLB) had been obtained at pH 6.5 and no bending of TM1 has been observed either.

The V-shaped central vestibule of PfMATE formed by the two domains contains four positively charged residues (Arg 83, Arg161, Arg284, Arg402) and the interior of the central cavity is lined with a cluster of methionine residues (Met27, Met31, Met64, Met173, Met256, Met260, Met287, Met381), creating a hydrophobic environment along the membrane plane of the cavity, which could be suitable for accommodation of negatively charged hydrophobic or lipophilic compounds.

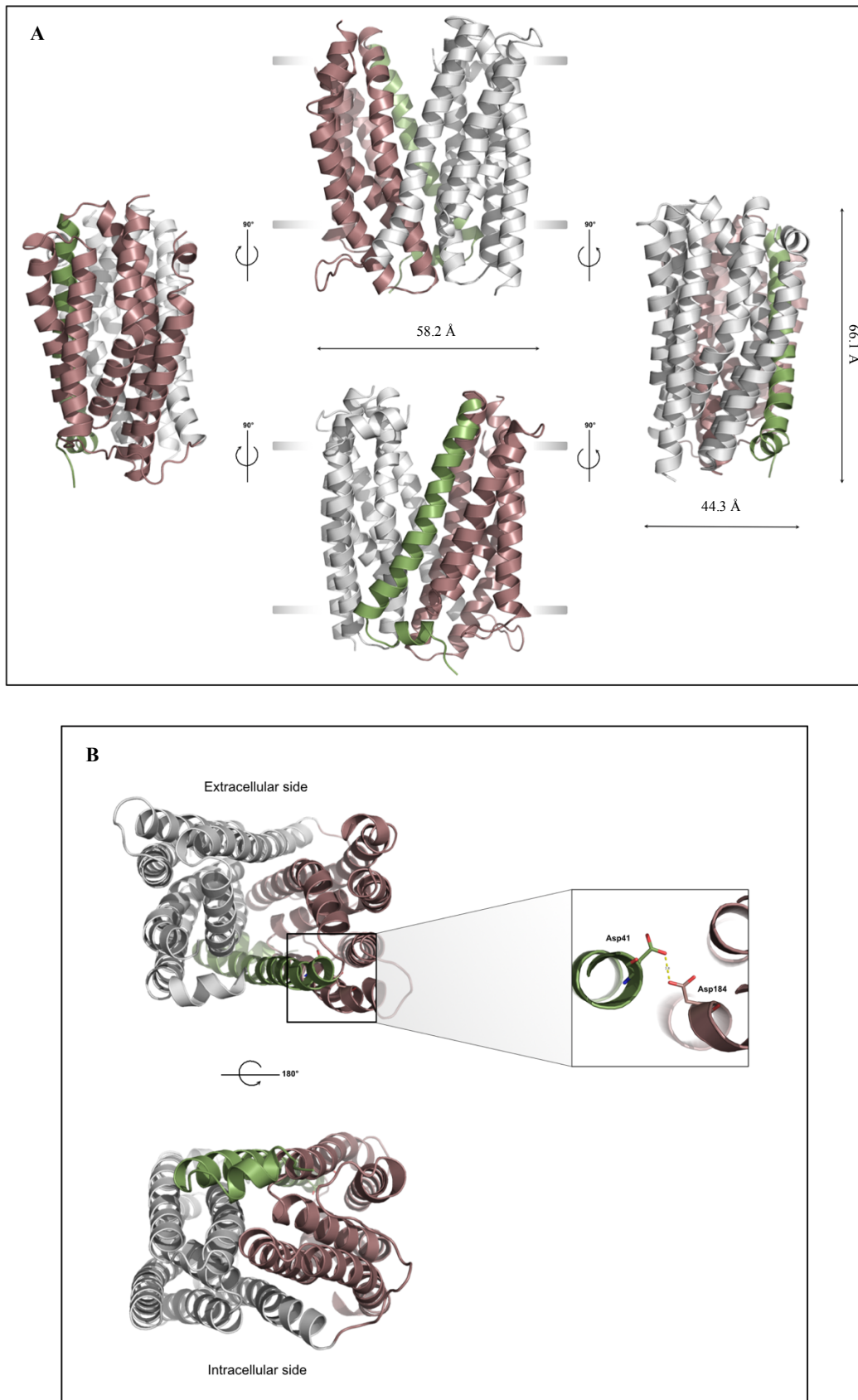


Fig. 3.30. Overall architecture of PfMATE in the outward-facing conformation. (A) Side view. (B) Top view. The N-terminal domain is depicted in purple, the C-terminal domain in silver. TM1 of the N-domain is highlighted in green.

3.2.8. Crystal structure of PfMATE in the inward-facing conformation

Crystallization of PfMATE in the presence of endogenous lipids extracted from the source organism *P. furiosus* led to structure determination of PfMATE in an inward-facing conformation (Fig. 3.31). This structure was determined at up to 2.8 Å resolution from the crystals grown in the LCP at low pH 5.0. Comprised of 12 TMs divided into two domains (TMs 1-6 and TMs 7-12), the inward-facing structure closely resembles the overall fold of the outward-facing state. Yet, structural superimposition of PfMATE in these two distinct states revealed significant conformational differences with a global Ca rmsd of 4.51 Å. The intracellular opening is significantly wider (about 26 Å) than the extracellular one (about 18 Å) in the respective access states.

In contradiction to the outward-facing structure, TM1 in the inward-facing state is significantly bent and undergoes a notable secondary structure rearrangement. A partial unwinding of the TM1 helix segment from Ser32 to Pro26 hints at a high degree of flexibility, suggesting involvement in the conformational transition between outward- and inward-oriented states. A structural superposition between the N-domain and C-domain of the inward-facing structure (global Ca root r.m.s.d = 2.58 Å) demonstrates that the internal twofold pseudosymmetry between these two domains remains well preserved. The conformational rearrangement of TM7 of the C-domain is highly similar to that of its corresponding symmetry related counterpart TM1 of the N-domain, which undergoes substantial bending and is located in close proximity to the C-domain helices.

Native lipids from *P. furiosus* appear to be a key determinant for capturing PfMATE in a novel conformational state, in which the central cavity is accessible from the cytoplasmic side of the membrane.

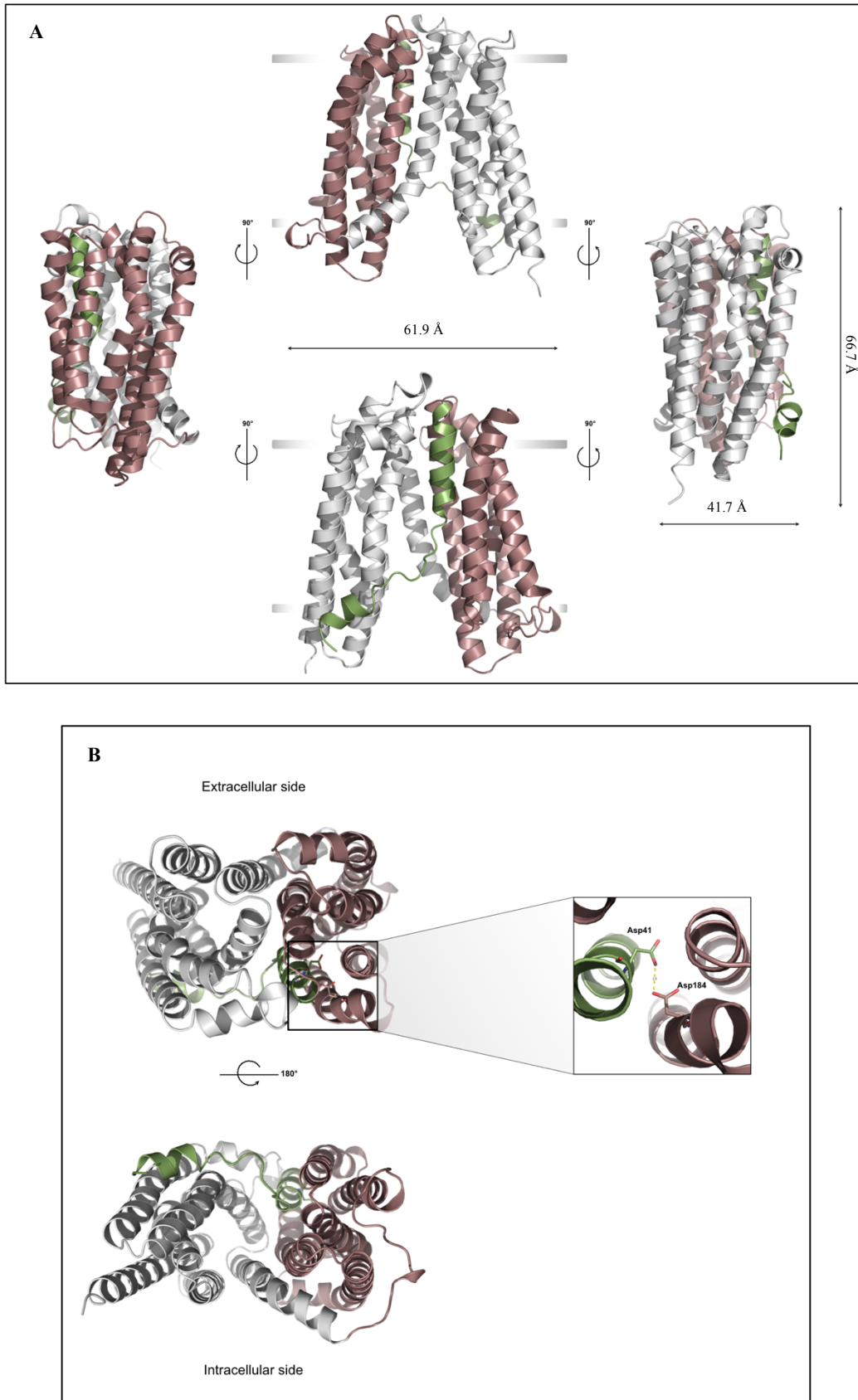


Fig. 3.31. Overall architecture of PfMATE in the inward-facing conformation. (A) Side view. (B) Top view. The N-terminal domain is depicted in purple, the C-terminal domain in silver. TM1 of the N-domain is highlighted in green.

Notably, the positively charged cavity of PfMATE demonstrates a reverse polarity compared to the canonical MATE transporters such as NorM_NG, NorM_VC as well as the eukaryotic MATEs. Among the MOP superfamily, the PfMATE vestibule rather resembles that of the MurJ negatively charged lipid II flippase from the MVF family, which is predominantly cationic⁵⁰ (Fig. 3.32) and indicates that the central cavity of PfMATE might be suitable to accommodate negatively charged hydrophobic compounds, like lipids.

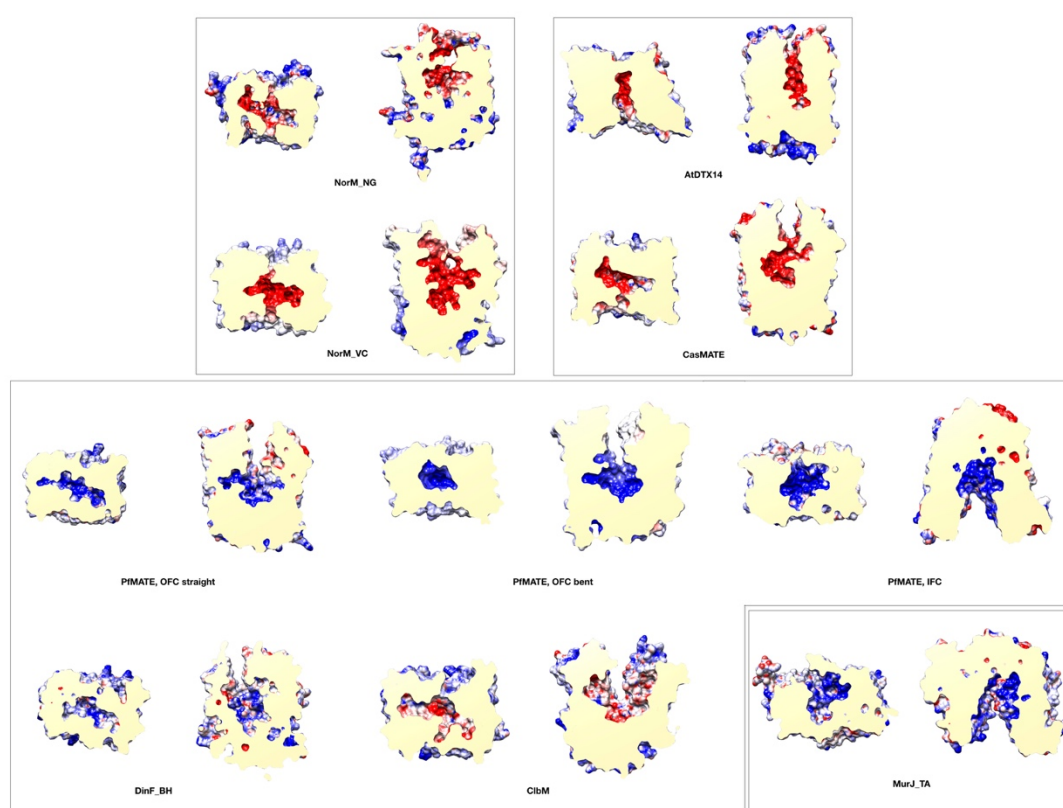


Fig. 3.32. Electrostatic surface potential of the central cavity of MOP superfamily proteins. Surface electrostatics representation of the MOP superfamily transporters, including MATEs as well as an MVF family lipid II flippase. Members of the NorM subfamily: NorM_NG (4HUK) and NorM_VC (3MKT); eukaryotic MATE subfamily: CasMATE (5XJJ) and AtDTX14 (5Y50); DinF subfamily: PfMATE, IFC (6FHZ); PfMATE, OFC straight (6GWH); PfMATE, OFC bent (3VVO), DinF_BH (4LZ6) and ClbM_EC (4Z3N); MVF family: MurJ_TA (5T77) were used for the comparison. Electrostatic surface potentials were calculated by APBS and shown from -10 kT/e (red, most anionic) to +10 kT/e (blue, most cationic), uncharged (white). PfMATE central vestibule shows a high similarity to the MurJ structure, which is strongly cationic.

3.3. Functional studies on PfMATE

3.3.1. Drug susceptibility test in *E. coli*

Based on the functional studies performed on PfMATE by a former PhD student, Viveka Nand Malviya, the minimum inhibitory concentration (MIC) values¹⁴⁹ were determined for several antimicrobial compounds including antibiotics and DNA binding dyes in the *E. coli* KAM32 strain ($\Delta acrB$, $\Delta ydhE$, $hsd\Delta 5$), which lack the genes encoding the major multidrug efflux pumps AcrB and YdhE and the restriction system hsd^{150} , and three variants of the BW25113 strain, namely BW25113A ($\Delta acrAB$, $\Delta tolC$, $\Delta ydhE$, $\Delta mdfA$, $\Delta emrE$), BW25113B ($\Delta acrAB$, $\Delta ydhE$, $\Delta emrE$) and BW25113C ($\Delta acrB$). The MIC values for different compounds were same for the test cells containing the plasmid with the gene coding for PfMATE (pBAD-*PfMATE*) and the control cells containing the empty vector (pBAD). Overexpression of the gene coding for PfMATE did not provide any additional resistance against any of the tested drugs and toxic compounds. These results imply that PfMATE may be more specialized than the classical MATE family transporters. During my PhD studies MIC tests were performed in the following strains: *E. coli* KAM32 (Fig. 3.33), *E. coli* C43(DE3) $\Delta acrAB$ (Fig. 3.34) (kindly provided by Prof. Ben Luisi), *E. coli* BL21(DE3) $\Delta acrAB$ (Fig. 3.35), *E. coli* C41(DE3) $\Delta acrAB$ (Fig. 3.36) (kindly provided by Prof. Klaas Martinus Pos). To determine the MIC value, a series of agar plates were prepared with increase of the concentrations of norfloxacin, identified previously as a substrate for PfMATE⁷³. The MIC value is defined as the lowest concentration of an antimicrobial agent that inhibits the visible growth of cells after overnight incubation. Control cells containing an empty plasmid (pBAD) were also screened along with the test cells containing expression plasmid (pBAD-*PfMATE*). All experiments were repeated at least three times independently. Briefly, ten-fold diluted solutions were spotted on the LB agar plates supplemented with 50 μ g/ml Carbenicillin, tested compound and an expression inducer (0.02 % arabinose). After 16 h of incubation at 37 °C, bacterial growth was analyzed visually.

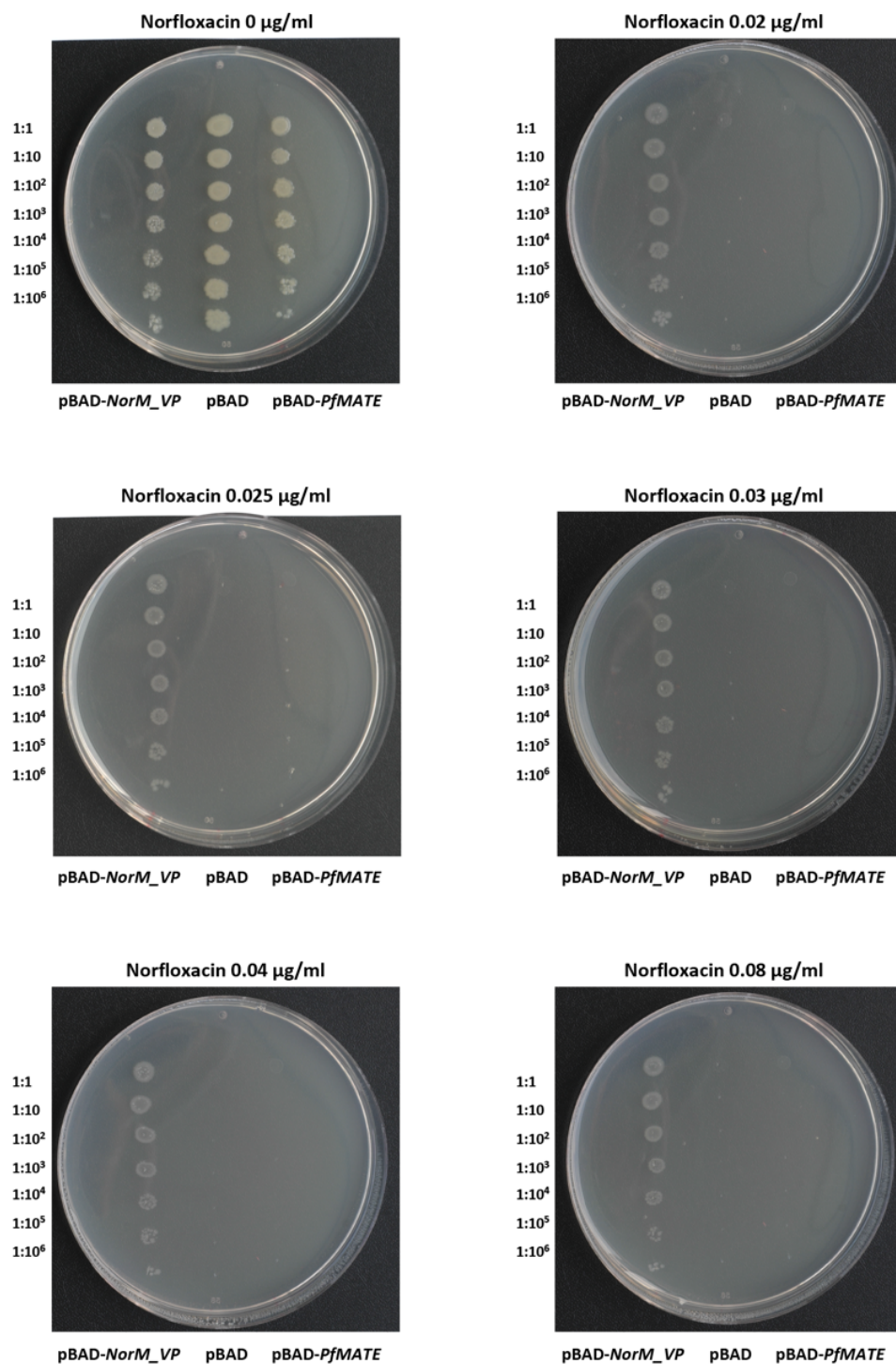


Fig. 3.33. MIC test in *E. coli* KAM32. Drug susceptibility test in the presence of different concentrations of norfloxacin. *E. coli* KAM32 cells with an empty plasmid (pBAD) (negative control), with pBAD vector containing the *PfMATE* gene (pBAD-*PfMATE*) and with pBAD vector containing the *NorM_VP* gene (pBAD-*NorM_VP*) (positive control) were used. Ten-fold serial diluted inoculum was spotted on LB agar plates.

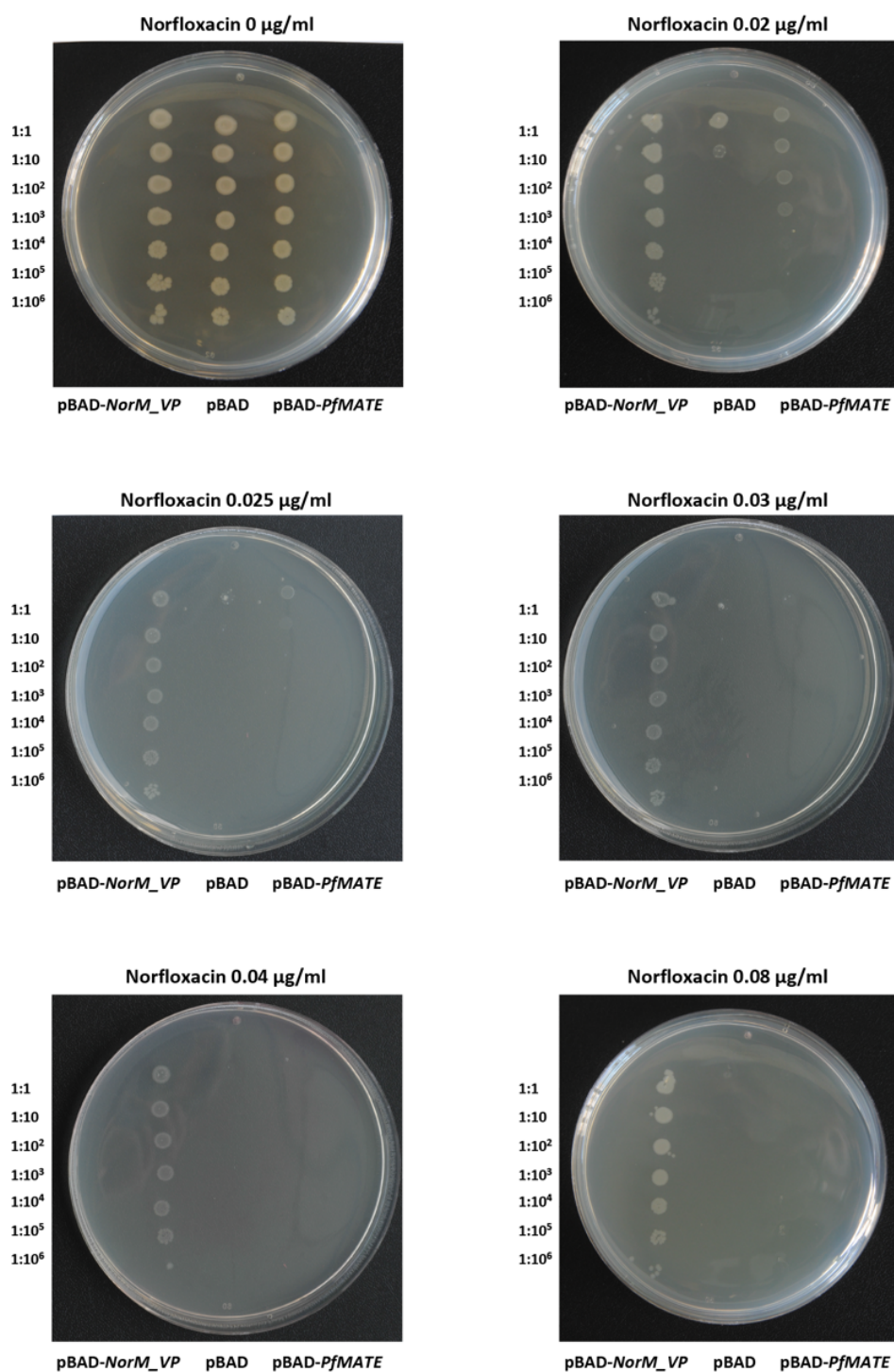


Fig. 3.34. MIC test in *E. coli* C43(DE3) Δ *acrAB*. Drug susceptibility test in the presence of different concentrations of norfloxacin. *E. coli* C43(DE3) Δ *acrAB* cells with an empty plasmid (pBAD) (negative control), with pBAD vector containing the *PfMATE* gene (pBAD-*PfMATE*) and with pBAD vector containing the *NorM_VP* gene (pBAD-*NorM_VP*) (positive control) were used. Ten-fold serial diluted inoculum was spotted on LB agar plates.

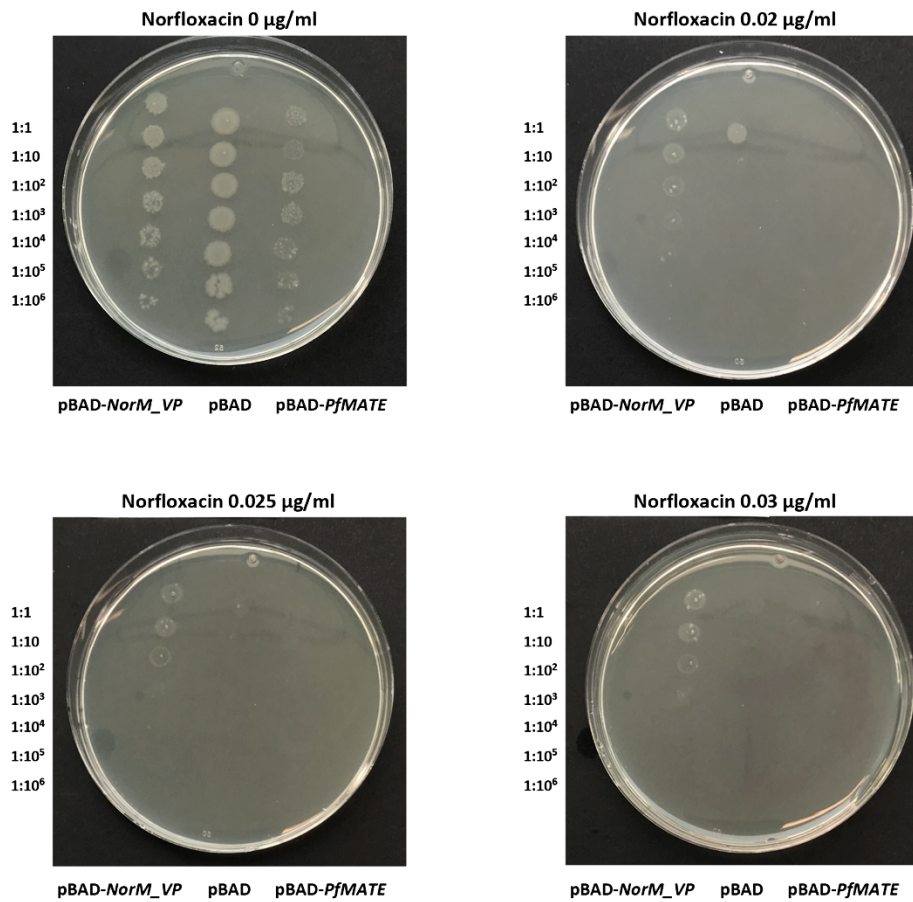


Fig. 3.35. MIC test in *E. coli* BL21(DE3) Δ *acrAB*. Drug susceptibility test in the presence of different concentrations of norfloxacin. *E. coli* BL21(DE3) Δ *acrAB* cells with an empty plasmid (pBAD) (negative control), with pBAD vector containing the *PfMATE* gene (pBAD-*PfMATE*) and with pBAD vector containing the *NorM_VP* gene (pBAD-*NorM_VP*) (positive control) were used. Ten-fold serial diluted inoculum was spotted on LB agar plates.

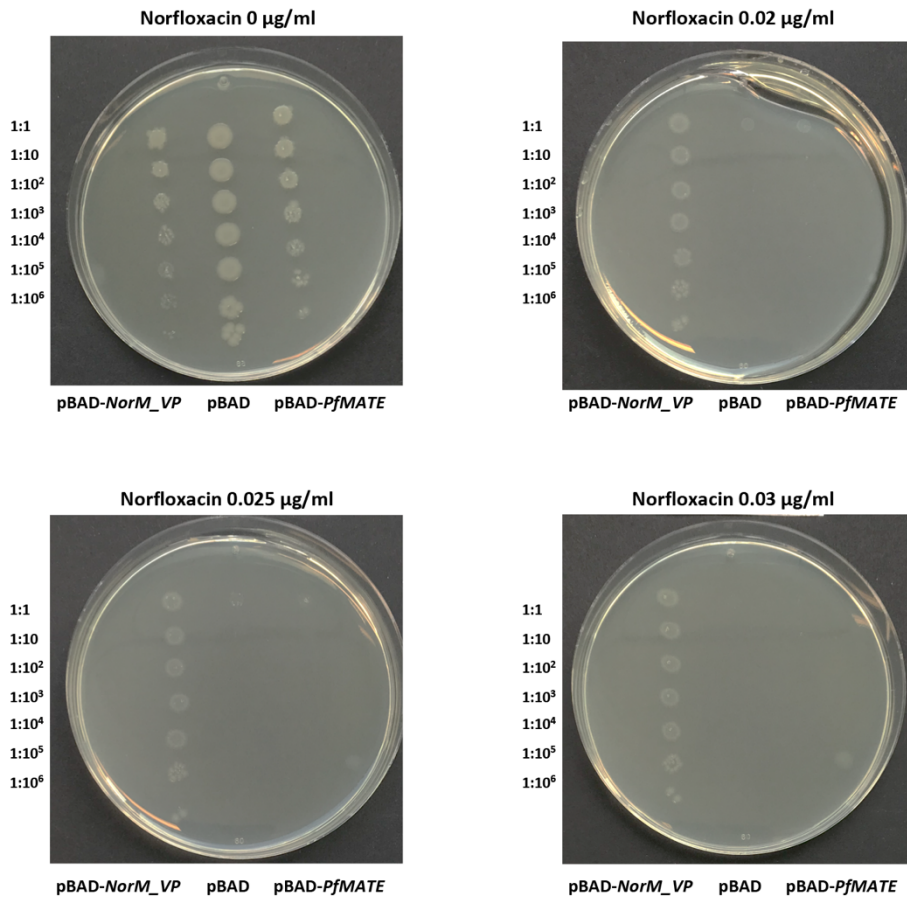


Fig. 3.36. MIC test in *E. coli* C41(DE3) Δ *acrAB*. Drug susceptibility test in the presence of different concentrations of norfloxacin. *E. coli* C41(DE3) Δ *acrAB* cells with an empty plasmid (pBAD) (negative control), with pBAD vector containing the *PfMATE* gene (pBAD-*PfMATE*) and with pBAD vector containing the *NorM_VP* gene (pBAD-*NorM_VP*) (positive control) were used. Ten-fold serial diluted inoculum was spotted on LB agar plates.

3.3.2. Drug susceptibility test in *P. furiosus*

In vivo functional assay has been also performed in the native organism *Pyrococcus furiosus*. Due to the elevated growth temperature of this archaeon, two chemical compounds with high thermal stability were tested: norfloxacin and ethidium bromide (EtBr). This experiment was carried out in *P. furiosus* WT containing the *pf0708* gene, and the *pf0708* knock-out strain (Fig. 3.37) under anaerobic conditions in ½ SME medium supplemented with the respective toxic compound. The knock-out strain (KO) was generated using homologous recombination by Franziska Maria Zörndlein under supervision of PD Dr. Winfried Hausner from the University of Regensburg. Deletion of the gene *pf0708* coding for PfmATE was confirmed at the DNA level (PCR using extracted genomic DNA as a template), and the RNA level (PCR using cDNA after reverse transcription of the extracted RNA as a template) (Fig. 3.37).

For the drug susceptibility assay in *P. furiosus*, firstly two strains were incubated anaerobically at 95 °C overnight. Then, fresh ½ SME medium was inoculated with the overnight preculture at the initial cell density $\sim 10^6$ - 10^7 cells/ml, followed by addition of a toxic compound. It was previously reported that the *pf0708* gene expression is up-regulated by so called cold-shock at 72 °C¹⁵¹. Therefore, the functional assay was performed by incubating the cells with shaking at 95 °C or 75 °C. The cell growth was monitored under a microscope by counting the cells using a Thoma chamber (Fig. 3.38). Initially, high concentrations of a chemical compound were tested: 80 µg/ml of norfloxacin and 30 µg/ml of EtBr. After overnight incubation, EtBr inhibited the growth of both strains, whereas the WT and KO cells were still able to survive in the presence of norfloxacin. This observation leads to the conclusion that both *P. furiosus* strains are resistant to norfloxacin at high concentrations 80 µg/ml, in contradiction to *E. coli* cells with and without the plasmid containing the *pf0708* gene, which survive in the presence of norfloxacin only at 0.02 µg/ml. Drug susceptibility test was also carried out in the presence

of decreasing concentrations of EtBr: 20, 10, 5, 3, 1 $\mu\text{g/ml}$. The cells were still able to survive in the presence of 1 and 3 $\mu\text{g/ml}$, but their growth was inhibited in the presence of 5 $\mu\text{g/ml}$ of EtBr. The growth behavior was similar for the cells with and without the gene *pf0708* coding for PfMATE.

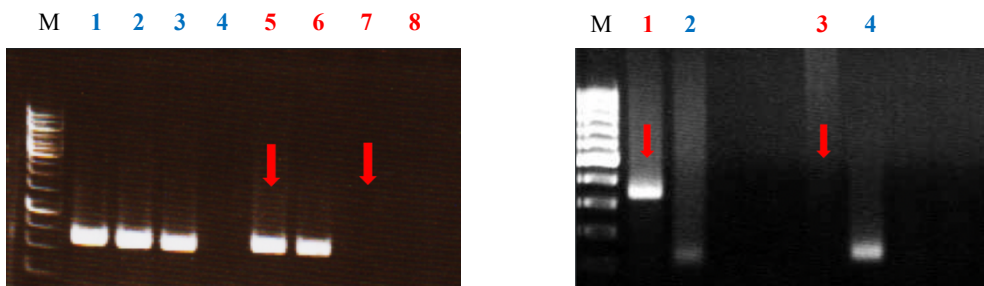
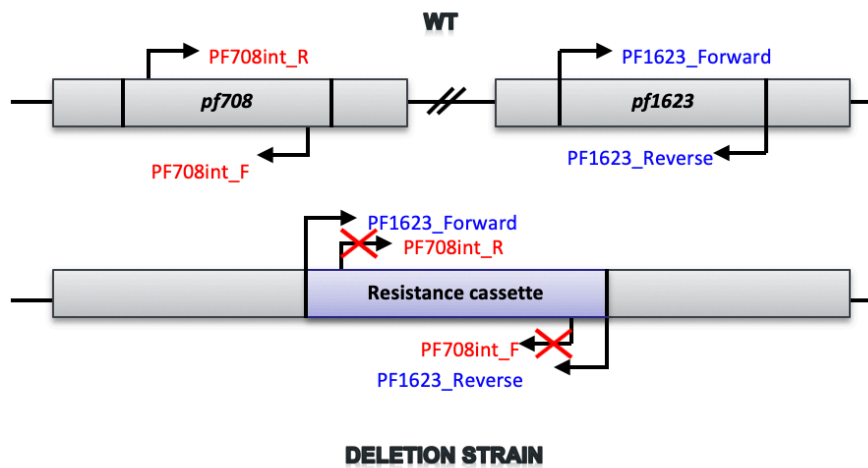


Fig. 3.37. *P. furiosus* WT and the *pf0708* deletion mutant. Schematic presentation of the gene arrangement in *P. furiosus* WT and the *pf0708* deletion strain with the oligonucleotides (PF1623_forward and PF1623_reverse, blue; PF708int_F and PF708int_R, red) used for PCR-based gene identification. Lower panel left, PCR products using genomic DNA as a template; 1,5, WT; 2,6, positive control; 3,7, deletion mutant, 4,8, negative control. Lower panel right, PCR products using cDNA generated by reverse transcription on an RNA as a template; 1,2, WT; 3,4, deletion mutant. Red arrows indicate differences between WT and the knock-out strain.

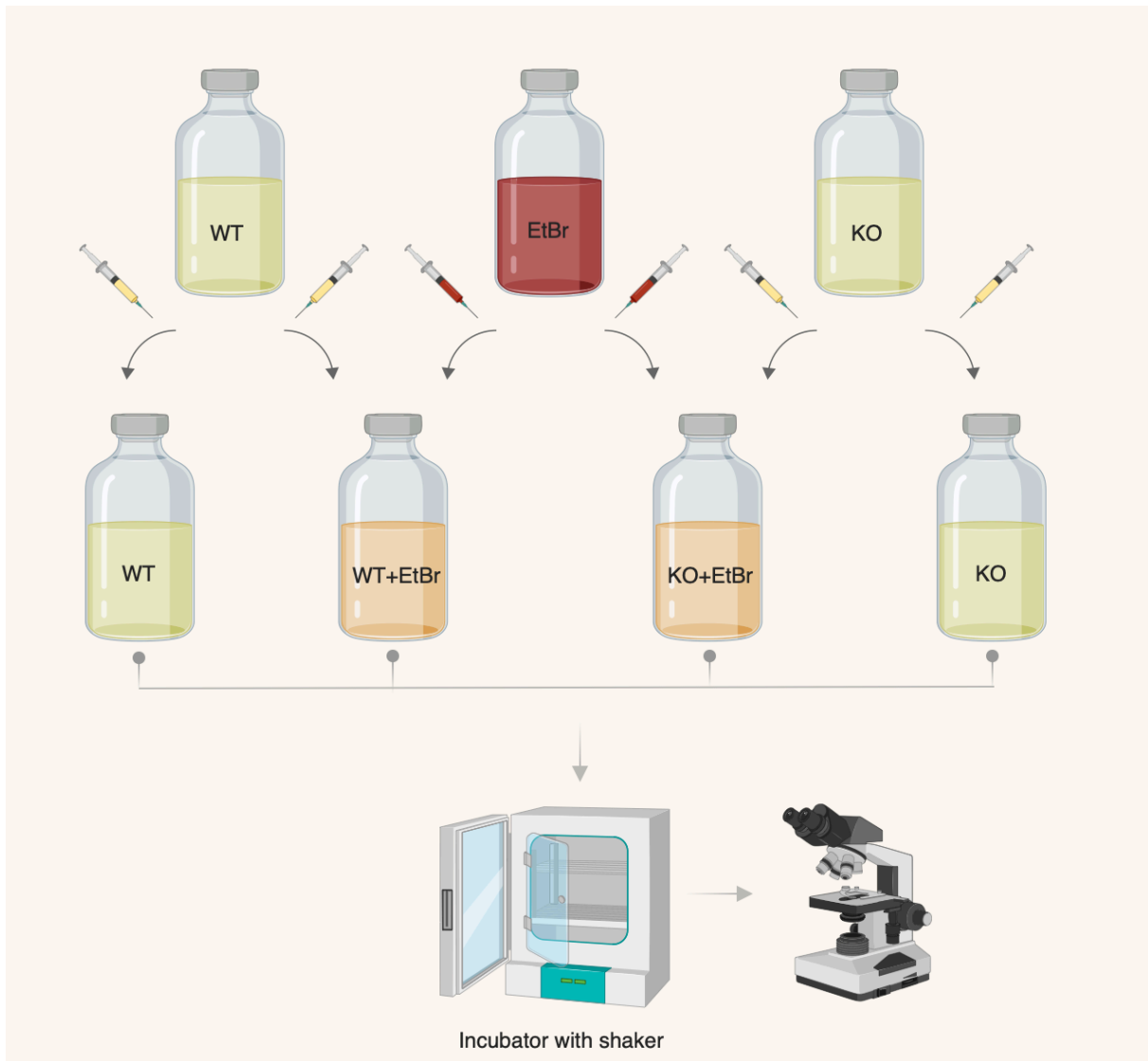


Fig. 3.38. Drug susceptibility assay in *P. furiosus*. Functional assays were performed in glass bottles pressurized with N_2 and filled with 20 ml $\frac{1}{2}$ SME medium. Each bottle corresponds to one strain in the presence or absence of a tested compound. The cells were counted under a microscope using a Thoma counting chamber. WT, wild type; KO, knock-out strain; EtBr, ethidium bromide. Upper image created using BioRender.com.

Assessment of the *in vivo* functional assay in the hyperthermophilic and anaerobic archaeon needs considerable care. There are many factors, which could influence the final output, such as growth phase, initial cell counts or syringe-based transfer of drugs and cells through robber stoppers in the glass bottles. Although, all precautions were taken, there is always a risk of some variations among the cultures, or inaccurate volumes added while handling bottles and syringes at high pressure and temperature. So far, there is no automated technique for monitoring the *P. furiosus* cell growth in a complex growth medium. The cells have to be counted manually under a microscope, which may not be as precise as measuring an optical density by spectrophotometer, typically used for mesophilic bacteria growth analysis. Nevertheless, at the current stage of this newly developed assay, no significant difference between WT and KO was observed. PfMATE did not provide any additional resistance against tested compounds, neither at 95 °C nor at 75 °C.

It is important to note that *P. furiosus* genome contains four genes annotated as MATE family members, namely *pf0708* (coding for PfMATE), *pf0823*, *pf0824* and *pf1850* as well as genes encoding other multidrug transporters from different families (RND, ABC, MFS). Results of the drug susceptibility assay suggest that PfMATE may not be involved in active efflux of the tested compounds or this function may be taken over by other multidrug transporters, which were still present in the cells.

4. DISCUSSION

4.1. Homologous and heterologous production of PfMATE

In order to provide a native environment, PfMATE was initially produced in the source organism, a hyperthermophilic and anaerobic archaeon, *Pyrococcus furiosus*. The protein has been successfully produced at the elevated temperature in a 100-l bioreactor pressurized with N₂, and homogeneously purified applying a two-step protocol consisting of affinity chromatography followed by gel filtration. Purified samples were subsequently used for the crystallization trials by the vapor diffusion as well as the LCP method, yielding a few crystal hits. To the best of my knowledge, the homologous production of PfMATE is the first example of a successful membrane protein production in *P. furiosus*. Taken together, expression under control of the *gdh* promoter resulted in a slightly higher amount of PfMATE compared with the expression under control of the *pf1623* promoter. The presence of an additional gene, *pf0709*, coding for the putative transcriptional regulator, has not improved the final protein yield and purity. Despite the generation of different constructs with two different promoters and different arrangement of the genes, including or excluding the gene *pf0709*, the *pf0708* expression level has not been improved and the final purified PfMATE quantities were insufficient for the extensive optimization of the crystallization conditions and further structural studies.

Due to the hurdles associated with the low amount of the protein produced homogeneously and limited tools for the optimization of expression in the hyperthermophilic and anaerobic archaeon, heterologous expression of the gene coding for PfMATE in the mesophilic bacterium *Escherichia coli* was carried out. The heterologous expression system of multiple secondary active transporters has been established in our department as described previously^{152–155}. The strategy of using pBAD system in *E. coli* TOP10 cells resulted in the significantly higher amounts of the homogeneously purified protein sample (in the mg range), being sufficient for

the crystallization experiments in the presence and absence of native lipids from *P. furiosus* leading to the PfMATE structure determination in two alternate conformations, inward and outward facing states, respectively (Fig. 4.1 and Fig. 4.2).

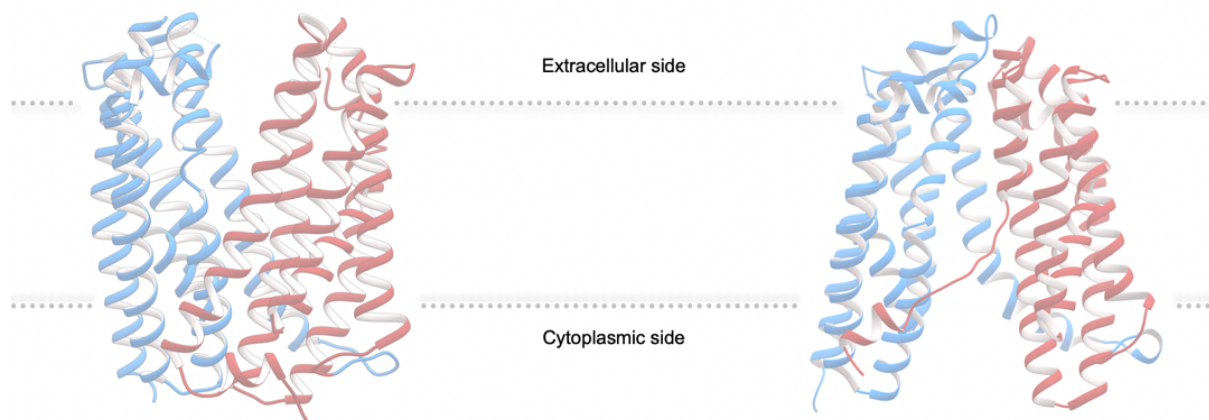


Fig. 4.1. Side view of PfMATE structures. Side view of PfMATE representing an outward and inward-facing conformation; ribbon models of the N-terminal domain (TMs 1-6) and the C-terminal domain (TMs 7-12) are shown in red and blue, respectively. The dashed lines represent the borders of the lipid bilayer.

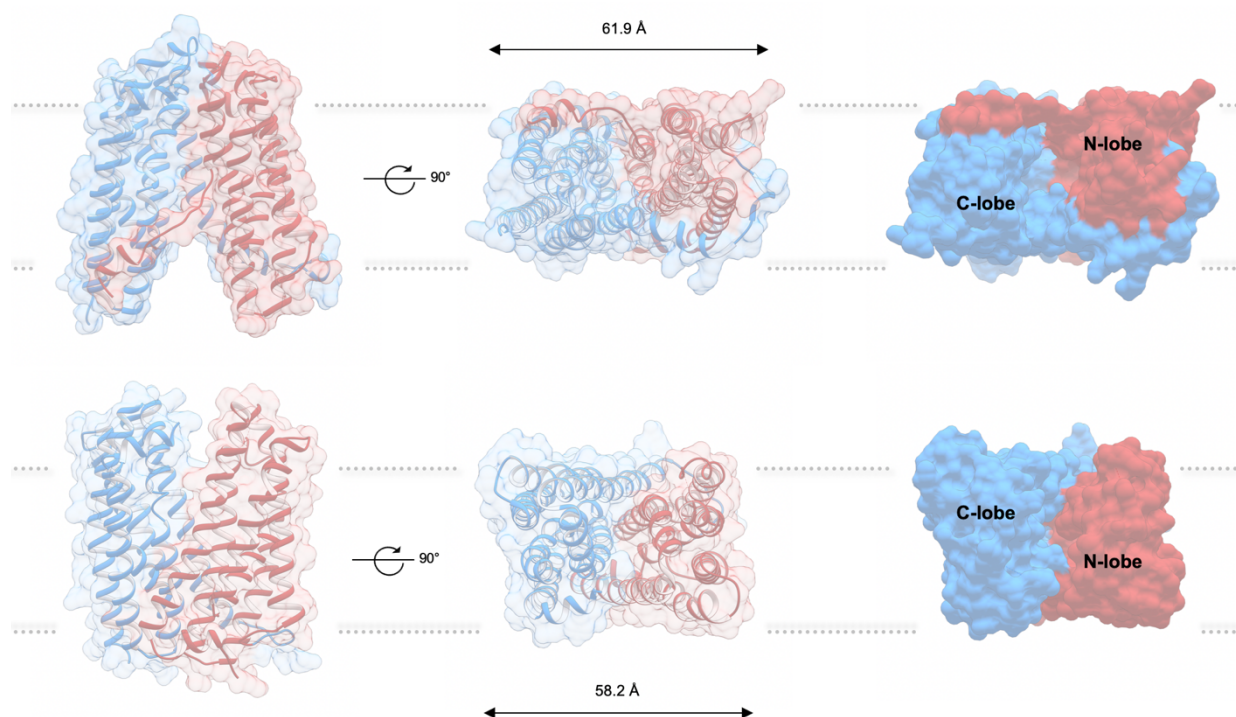


Fig. 4.2. Schematic presentation of PfMATE structures. Surface representation of the OFC and IFC, showing the central binding cavity from the top view and the side view. The N-terminal domain is depicted in red, the C-terminal domain in blue.

4.2. Structural and functional implications of the TM1 bending

Due to the weak electron density in the intramembrane region of TM1 of the inward-facing PfMATE structure, the chain tracing was validated by 2Fo-Fc simulated annealing composite omit maps¹²⁹ and feature enhanced maps (FEM)¹³⁰ generated by PHENIX (Fig. 4.3). TM1 is not involved in crystal packing and thus its bending is unlikely to be a crystallographic artifact.

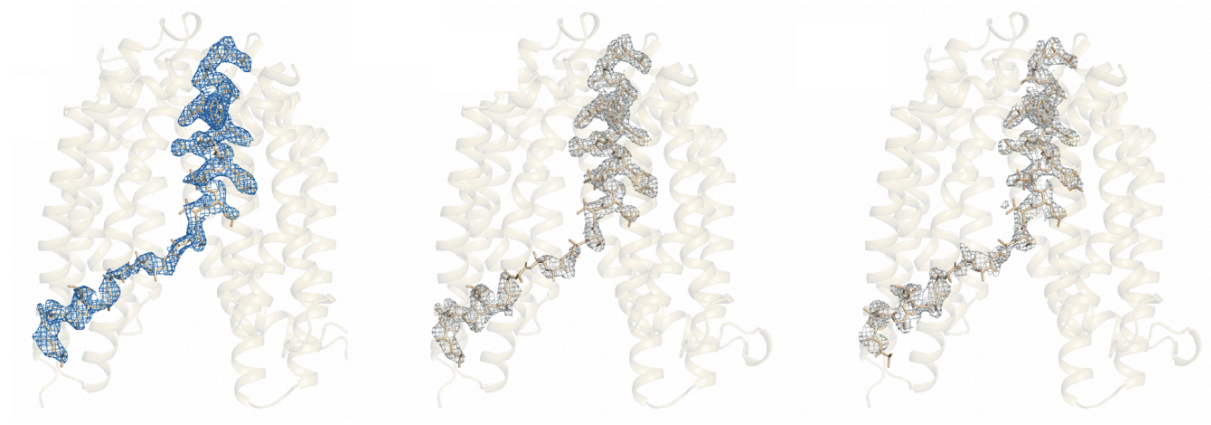


Fig. 4.3. TM1 bending. Partial unwinding of the intrahelical region and substantial bending of TM1 in the inward-facing conformation. TM1 is shown in a stick representation, whereas remaining helices TMs 2-12 are depicted as a ribbon model. (Left) 2Fo-Fc electron density map corresponding to TM1, shown as blue mesh. Arrangement of TM1 in the inward-facing structure was validated by the unbiased electron density maps. (Middle) 2Fo-Fc feature enhanced map (FEM) and (Right) 2Fo-Fc simulated annealing composite OMIT map for TM1. Both maps are shown in grey mesh.

According to a previously proposed model of the PfMATE transport mechanism^{73,156}, the conserved aspartate Asp41 constitutes the key residue associated with the proton coupling. It has been proposed that protonation of Asp41 induces a conformational change of TM1 switching from the straight to a bent state. This conclusion was based on two PfMATE crystal structures obtained at high pH (7.0-8.0) and low pH (6.0-6.5) conditions, respectively (Fig. 4.4). The aforementioned interpretation was discussed controversially, leading to the hypothesis

about TM1 bending to be rather affected by exogenous lipids (monoolein) present in the crystallization conditions as described previously by Radchenko *et al.*⁶⁹ and Lu¹¹⁶. Based on this discrepancy, the relationship between pH and structural alteration of TM1 was evaluated. Therefore, the outward-facing structure of PfMATE was determined from crystals generated at low pH conditions (pH 5.0) by the LCP method. Comparative analysis of this structure and previously presented high pH outward-facing structure (3VVN) reveals that TM1 remains straight irrespective of the pH of the crystallization conditions. There are no major structural differences as reflected by a very low Ca r.m.s.d.=0.86 Å of the structure obtained in the presence of the exogenous lipids (6GWH) and 0.69 Å, 0.71 Å, 1.08 Å, 0.64 Å of chain A, B, C and D, respectively (4MLB) and 0.74 Å, 0.77 Å, 1.10 Å, 0.65 Å of chain A, B, C and D, respectively (6HFB) of the structures generated in the absence of monoolein molecules. Hence, these results imply that the conformational rearrangement of the helix TM1 is most likely independent of pH.

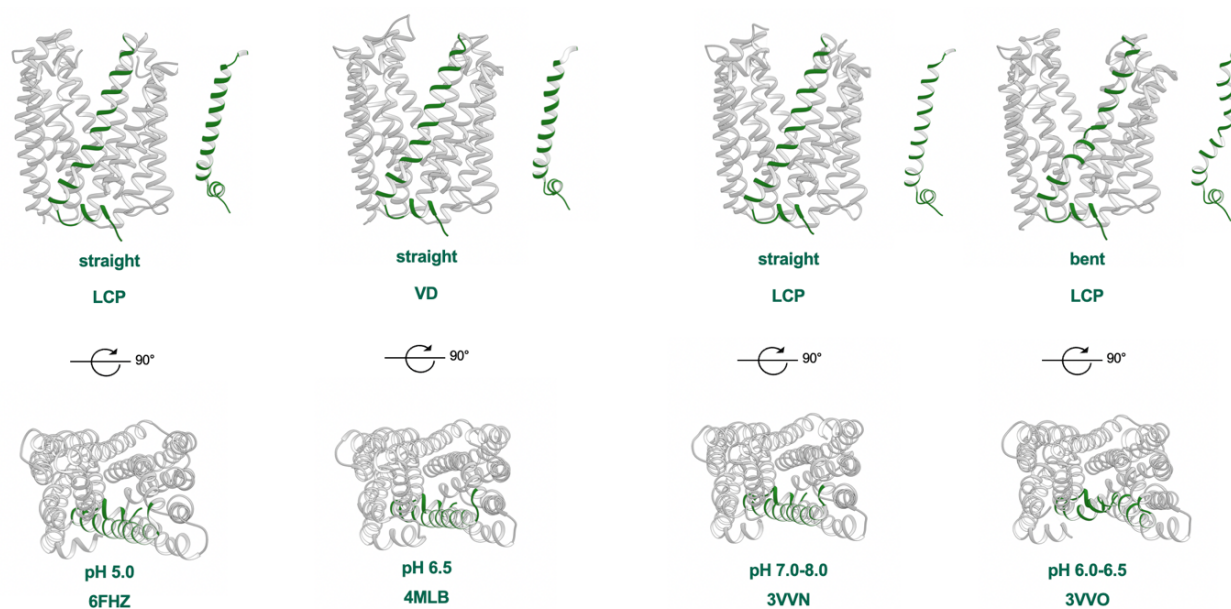


Fig. 4.4. PfMATE structures in the outward-facing conformation. Schematic representation of PfMATE structures in the OFC straight and bent form. TM1 is highlighted in green. PDB ID, crystallization method and crystallization pH conditions are shown.

Nevertheless, the largest deviation between the PfMATE outward and inward-facing conformations is found within TM1. A pivot point at Gly30 allows the large hinge-like movement within the N-terminal segment of TM1 to adopt an inward-facing conformation. The unfolded intramembrane region in TM1 exhibits a remarkably high flexibility (as featured by large B-factors in Fig. 4.5), suitable for coordination of substrates or ions. In case of the outward-facing structure, the high B-factors are found for the extracellular loop regions; whereas in the inward-facing structure, the highest B-factors are observed for the intracellular half of TM1. The B-factors are also called temperature factors or displacements parameter. High B-factors may indicate the atomic displacement from an average value as a result of temperature-dependent atomic vibrations, static disorder in a crystal lattice or dynamic disorder, when parts of the molecule have multiple alternate conformations¹⁵⁷.

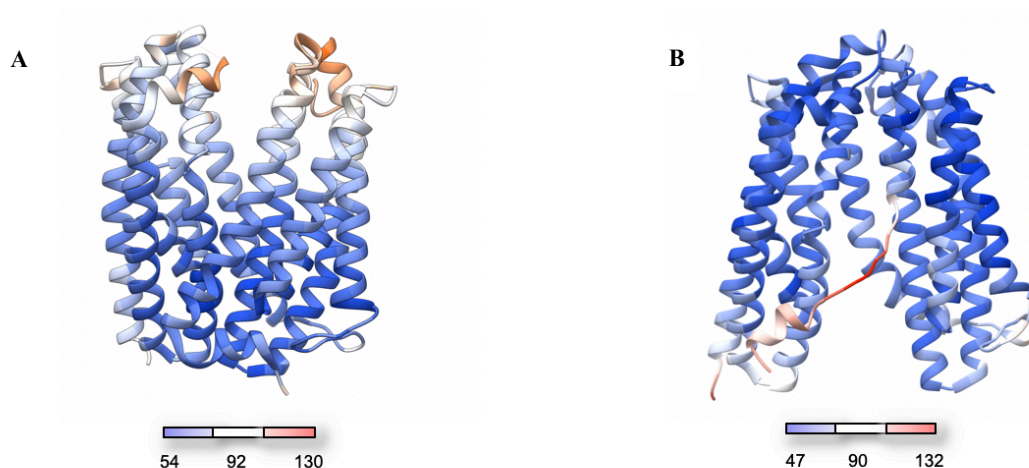


Fig. 4.5. Average B-factors. Average B-factors in the (A) outward and (B) inward-facing conformations. Cyan represents low B-factors (54.3525 \AA^2 and 47.4514 \AA^2 for the outward and inward-facing state, respectively), white moderate B-factors (92.05 \AA^2 and 89.6057 \AA^2 for the outward and inward-facing state, respectively), whereas maroon high B-factors (129.748 \AA^2 and 131.76 \AA^2 for the outward and inward-facing state, respectively).

During transition from the outward to inward-facing state, two segments of TM1, one from Ala17 to Ala31 and the second one from Ser32 to Val45, are tilting relative to each other. This structural alteration could be divided into two movements described by the following angles. The first one defines the movement of TM1 as the gatekeeper of the lateral opening in the YZ plane (about 27°). The second angle describes the pronounced movement in the XZ plane (about 42°), showing that the intra-membrane half of TM1 acts as a plug to seal the lateral opening during the transition to the outward-facing conformation. In the inward-facing structure the side chain of Phe60 (TM2) is flipped towards Val56 (TM2), creating a hydrophobic contact with Gly42 (TM1), which leads to a helical kink. In the outward-facing structure, these residues are shifted apart and TM1 remains straight. The bending of TM1 in the inward-facing state is further enhanced by the interaction between Asn154 (TM10) and Gln34 (TM1), which can correspond to either a hydrogen bond or an amide bridge. The energy of an amide interaction is in the range -65 to -70 kJ/mol, therefore it is significantly stronger than a hydrogen bond (8-30 kJ/mol). A hydrogen bond between the side chain carbonyl group of Asp41 (TM1) and the side chain OH of Tyr139 (TM4), as well as a hydrophobic interaction between Pro26 (TM1) and Ala166 (TM5), are most likely involved in the formation of the kinked state of TM1. The structural analysis of two distinct PfMATE conformations reveals that the intracellular entrance is wider than the extracellular gate. It may be predominantly caused by the substantial movement of TM1 towards the C-domain bundle in the inward-facing state, thereby providing a wider entrance for a substrate from the internal surrounding. The structural alteration of TM1 in the inward-facing structure displays a highly dynamic nature and flexibility of this helix, which may play a functional role in ion and substrate gating.

A structural superposition of the N-domains (Ca r.m.s.d = 2.02 Å) and C-domains (Ca r.m.s.d = 1.84 Å) of the PfMATE outward and inward-facing conformations reveals that the intracellular and extracellular barriers are formed by local and global conformational changes

including rigid-body movements within two domains, a hinge-like motion of the N-terminal segment of TM1, and unwinding of TM1 halfway across the membrane (Fig. 4.6). Alternating access to the substrate binding site in PfMATE appears to be enabled by significant movements of TM1 and TM7, while the remaining transmembrane helices of the N-domain (TMs 2-6) and the C-domain (TMs 8-12) stay together and undergo a rigid body movement (Fig. 4.7).

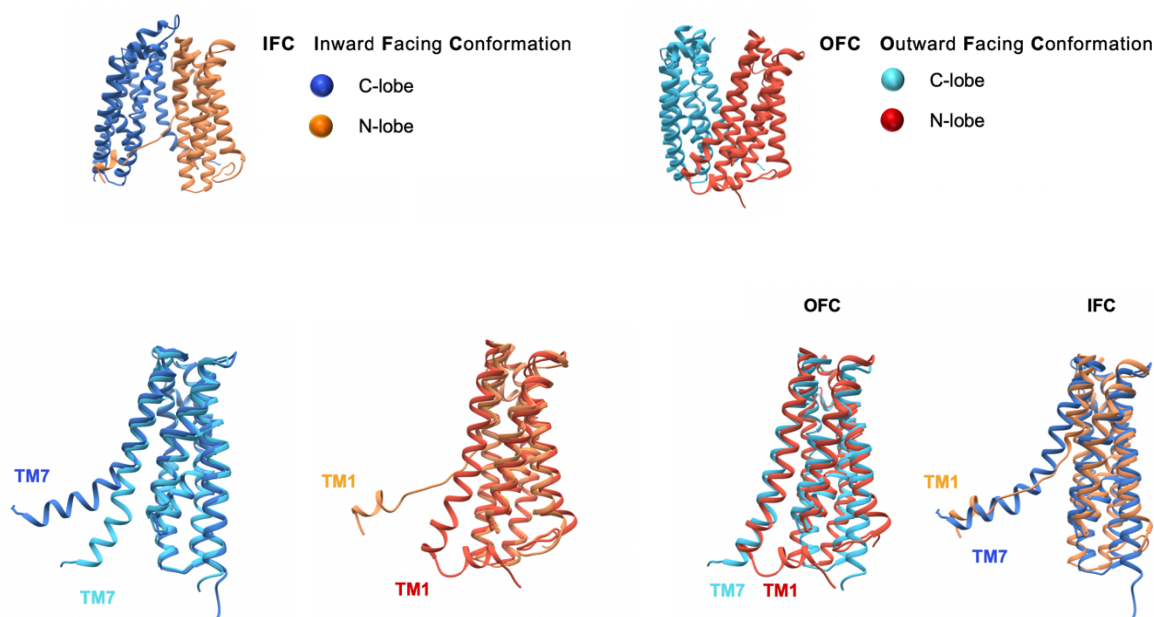


Fig. 4.6. Structural superposition of the N- and C-domains of the two distinct PfMATE states. Superposition of the N- and C-domains of the outward (OFC) and inward-facing (IFC) conformation indicates a substantial shift of TM1 in the N-domain by 42° and TM7 in the C-domain by 40°, whereas the remaining helices undergo a rigid body movement. On the right side, structural alignments of the N- and C-domain of the IFC and OFC show their symmetrical arrangement. The alternating access mechanism of PfMATE is based on the rigid body movements of TMs 2-5 and 8-12, bending of TM7, and bending and unwinding of TM1.

Conformational changes occurring during transition between the inward to outward-facing state were also analyzed by the targeted-MD simulations performed by Dr. Ahmad Reza Mehdipour from the Theoretical Biophysics Department of the Max Planck Institute of Biophysics (Section 2.24). Selected snapshots representing different stages of a 50-ns run are depicted in Fig. 4.8. These simulations demonstrate unwinding and flexible nature of TM1.

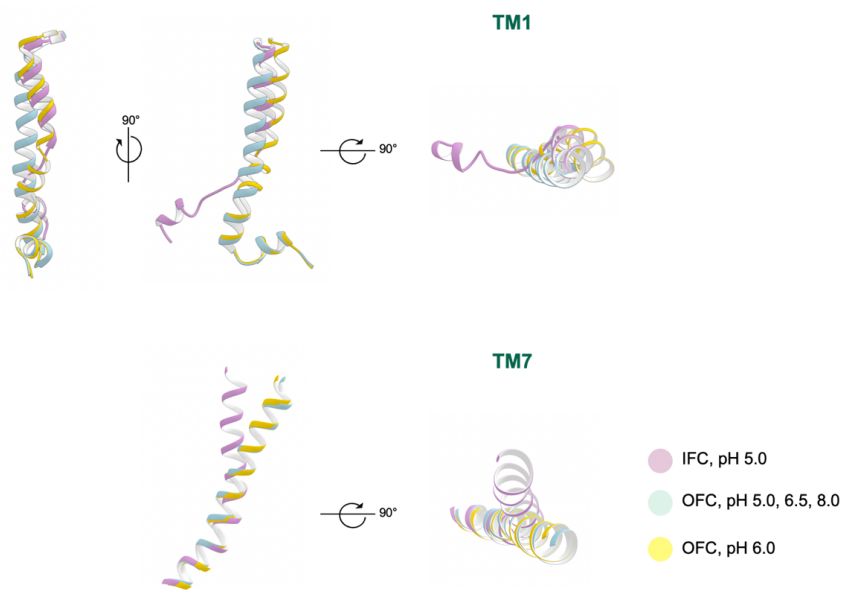


Fig. 4.7. Fluctuations within the helices TM1 and TM7. Structural superposition of TM1 and TM7 of the inward and outward facing conformations.

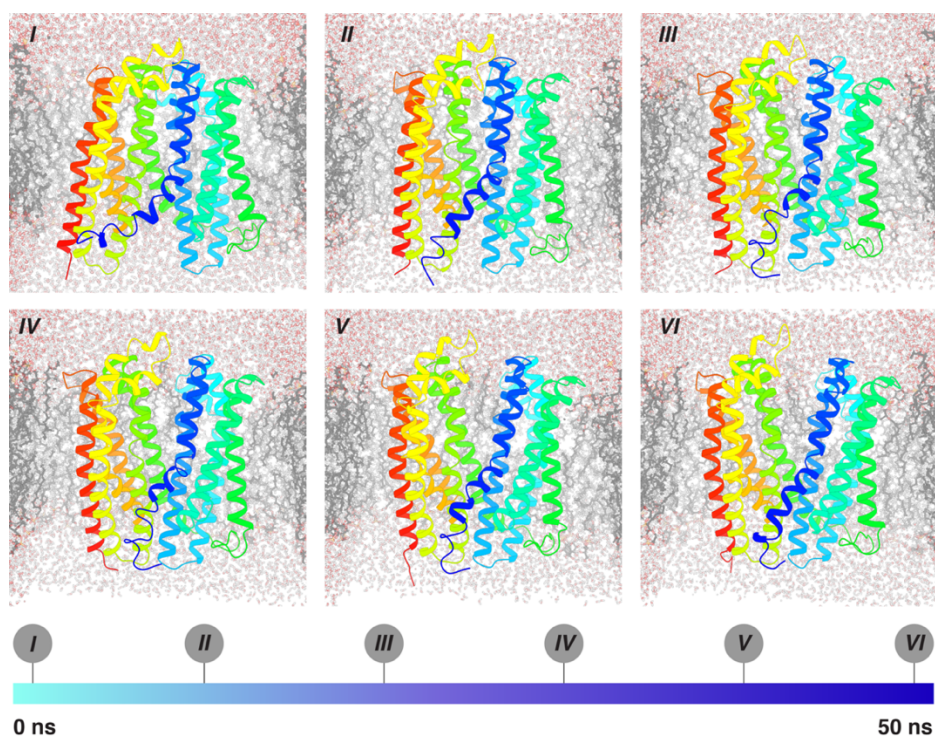


Fig. 4.8. Conformational changes of PfMATE. Snapshots from a 50-ns targeted MD simulation show intermediate states along a transition from the inward-facing to an outward-facing conformation.

4.3. Structural basis of the extracellular barrier formation

Secondary active transporters use electrochemical gradients of either protons or sodium ions in order to enable the uphill movement of substrates across the biological membrane. According to the basic principles of the alternating access model⁷⁷, MATE transporters sequentially expose their central cavity to either side of the membrane during a transport cycle by internal or external barrier formation. These barriers maintain the accessibility of the substrate and ion binding sites to one respective surrounding at a time.

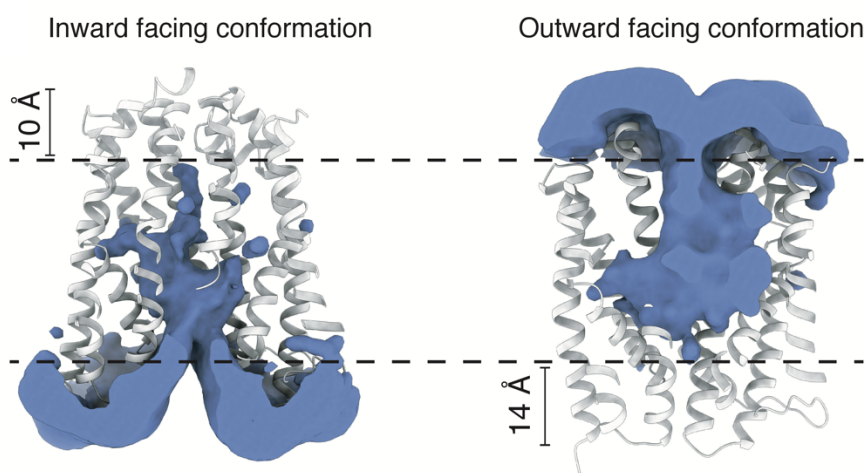


Fig. 4.9. Water accessibility. Water occupancy of the binding site cavity (blue surface) in simulations of the inward and outward-facing conformations of PfMATE. Water access to the cavity is restricted to the inside and outside, respectively.

In the PfMATE inward-facing conformation, the extracellular barrier, which blocks access to the substrate binding site from the extracellular side of the membrane, predominantly involves hydrophobic and aromatic interactions between residues located at the extracellular apex of the N- and C-domains. This tightly closed barrier of about 10 Å in thickness (Fig. 4.9), is mediated by hydrophobic interactions between Pro50 (TM2), Leu53 (TM2), Ala54 (TM2), Ala277 (TM8), Ala355 (loop connecting TM9 and TM10), Val357 (TM10) and Ile358 (TM10). The

hydrophobic interaction between Met64 (TM2) and Met260 (TM7) is associated with the formation of a hydrophobic groove within the C-domain. A network of interactions between several residues from the C-domain create this groove: Leu263 (TM7), Phe279 (TM8), Trp283 (TM8), Met287 (TM8), Leu369 (TM10), Phe372 (TM10), Ala411 (TM11), Val425 (TM12), Trp426 (TM12) and Ile429 (TM12). Additionally, the extracellular barrier is stabilized by a hydrogen bond between the carbonyl oxygen atom of Ile43 (TM1), the side chain OH group of Tyr351 (loop between TM9 and TM10) and the side chain OH group of Ser46 (TM1). The van der Waals interactions between Phe349 (TM9) and Ile43 (TM1), Gly49 (TM1), Ile285 (TM8) serve as another component for the closure of the extracellular gate (Fig. 4.10). Structural superposition of two alternate states of PfMATE reveals that helices TM7 and TM8 in the inward-facing structure are displaced at the extracellular apex of the C-domain compared to their positions in the outward-facing structure, closely approaching the helices of the N-domain. This alteration is most likely caused by the hydrophobic interactions formed between Leu58 (TM2), Val62 (TM2), Leu122 (TM3), Met126 (TM3), Ala128 (loop between TM3 and TM4), Phe261 (TM7), Ile268 (TM7) and Val276 (TM8).

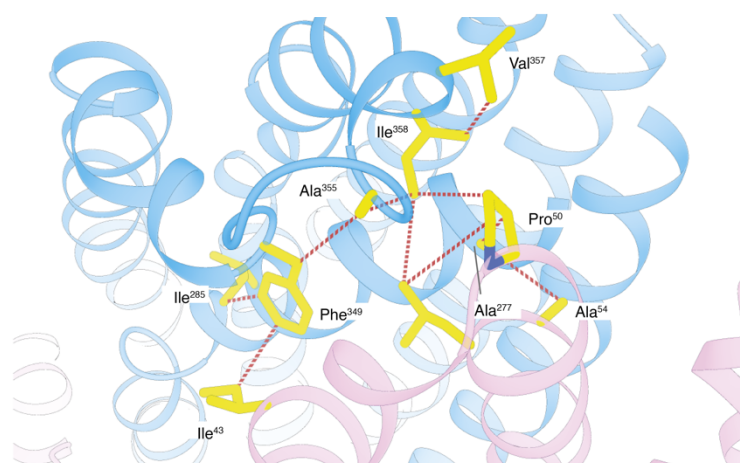


Fig. 4.10. Extracellular barrier in PfMATE inward-facing conformation. Close-up views of the extracellular barrier in the inward-facing structure. Selected residues are shown as sticks and labeled. Salt bridges and van der Waals interactions are displayed as black and red dashed lines, respectively.

4.4. Structural basis of the intracellular barrier formation

During PfMATE transition from the inward- to the outward-facing state, the interaction network contributing to the closure of the extracellular barrier has to be disrupted. Formation of the intracellular barrier obstructs solvent accessibility from the cytoplasm, while the extracellular barrier opens. Thus, the substrate and ion binding sites are accessible via an aqueous cavity, which is exposed to the extracellular side of the membrane. The thicker barrier between the substrate and the internal surrounding (about 14 Å) (Fig. 4.9) may be correlated with the substantial movement of the intracellular half of TM1.

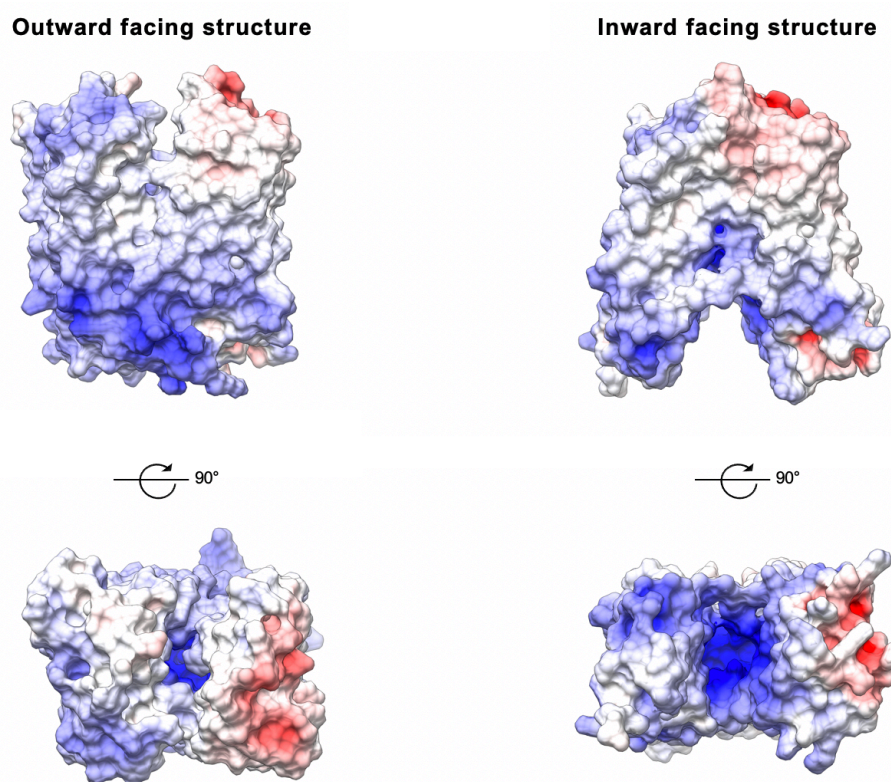


Fig. 4.11. Charged residues in the inward and outward-facing structure of PfMATE. According to the positive inside rule, the intracellular part of the membrane protein contains more positive charges compared to the extracellular region. Blue indicates positive charges of Arg and Lys, whereas red negative ones of Asp and Glu.

The intracellular part of PfMATE contains more positively charged residues (Arg and Lys) compared to the extracellular region (Fig. 4.11), which is consistent with the positive inside rule¹⁵⁸. Therefore, not only the hydrophobic interactions, predominantly stabilizing the extracellular barrier, but also ionic interactions, such as the salt bridge between the side chains of Arg244 (TM7) and Glu393 (TM11) contribute to the formation of the intracellular barrier. The ionic interactions between the side chains of Glu310 (TM8), Arg13 (TM1) and Arg88 (TM2) most likely drive the movement of TM8 and TM9 on the intracellular side, thereby closing the intracellular substrate access pathway. In the outward-facing structure, the side chain of Arg244 (TM7) is flipped away from Asp241 (TM7) and forms an ionic interaction with the side chain of Glu393 (TM11), which consequently bring TM7 and TM11 together. Interestingly, the side chains of Arg13 (TM1) and Arg88 (TM2) located at the apex of the intracellular barrier are in close proximity (3.6 Å) and stacked in an antiparallel manner at the apex of the intracellular barrier (Fig. 4.12). This coordination of two positively charged residues may imply the existence of a repulsive electrostatic gate, which could preclude a flux of protons and other cations, thereby forming a tight barrier and preventing ion leakage during the conformational transition. The hydrophobic interaction between Val9 (TM1) and Ile85 (TM2) strengthen the intracellular barrier, which stabilizes the position of the short helical stretch of the N-terminus in the outward-facing structure. An aqueous path to the intracellular side is also closed through a hydrophobic interaction network between Ala82, Gly86 (TM2) and Ala306 (TM8).

The outward and inward-facing conformations of PfMATE illustrate how these two distinct barriers are formed by a network of mostly hydrophobic residues at the closely interacting domain interfaces in each state.

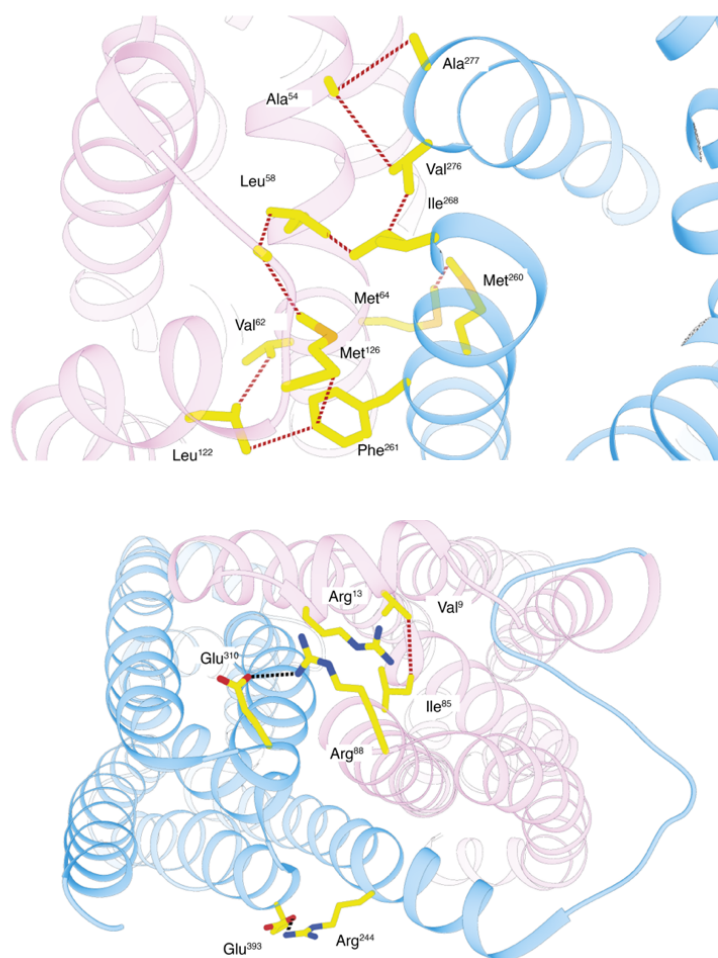


Fig. 4.12. Intracellular barrier in PfMATE inward-facing conformation. Close-up views of the intracellular barrier in the outward-facing structure. Selected residues are shown as sticks and labeled. Salt bridges and van der Waals interactions are displayed as black and red dashed lines, respectively.

4.5. Ion binding sites

The ion binding site in the outward-facing conformation of PfMATE is located near Asp41, Asn180, Asp184 and Thr202 in a water accessible surrounding, whereas in the inward-facing conformation it appears to be occluded. Although sodium ion and water molecule are not straightforwardly distinguishable through X-ray crystallography, Tanaka *et al.*⁷³ interpreted an electron density peak near this site as a bound water molecule in the outward-facing state of PfMATE (3VVO and 3VVN). A highly similar ion binding site formed by Asp35, Asn174, Asp178, Ala192 and Thr196 (corresponding to the homologous residues Asp41, Asn180, Asp184, Ala198 and Thr202 in PfMATE) was identified in the crystal structure of another MATE family transporter from *Vibrio cholerae*, VcmN⁷⁰. The electron density at this site was also assigned to a water molecule by analogy with the aforementioned studies on PfMATE.

The CsCl heavy atom derivative dataset of the outward-facing structure (6HFB) from the SAD experiments using anomalous dispersion of Cs⁺, (performed previously by a former member of our department, Tsuyoshi Nonaka), shows an anomalous difference electron density for a cesium ion, adjacent to the carboxyl oxygen atom Oδ1 of Asp41 (Fig. 4.13C). Cesium ions do not occupy exactly the same position as Na⁺, probably due to the steric hindrance and larger size compared to sodium ions. The anomalous signal from a Cs⁺ demonstrates the negative electrostatic potential of this site, which is suited for binding monovalent cations. Also, recently published MD simulations⁶⁵ of the outward-facing bent state of PfMATE (3VVO) reveal a highly coordinated Na⁺ binding site at Asp41. This sodium ion-occluded bent state appears to be an intermediate in the transport cycle stabilized by ion binding before transition into the inward-facing state.

A prominent spherical electron density signal was observed in the Fo-Fc map of the inward-facing structure (6FHZ) in a close proximity to residues Asp41, Asp184, Asn180 and Thr202 (Fig. 4.13A), consistent with a coordination number of six. This observation raises the

possibility that the electron density in the center is caused by the presence of a sodium ion. The distances between the sodium ion and oxygen atoms of the coordinating residues in this structure: O δ 1 of Asp41 (2.42 Å), O δ 1 of Asn180 (2.45 Å), O δ 1 of Asp184 (2.60 Å), carbonyl oxygen atom of Ala198 (2.08 Å) and O γ of Thr202 (2.38 Å), are too short for hydrogen bonds and indicate a Na⁺ bound state.

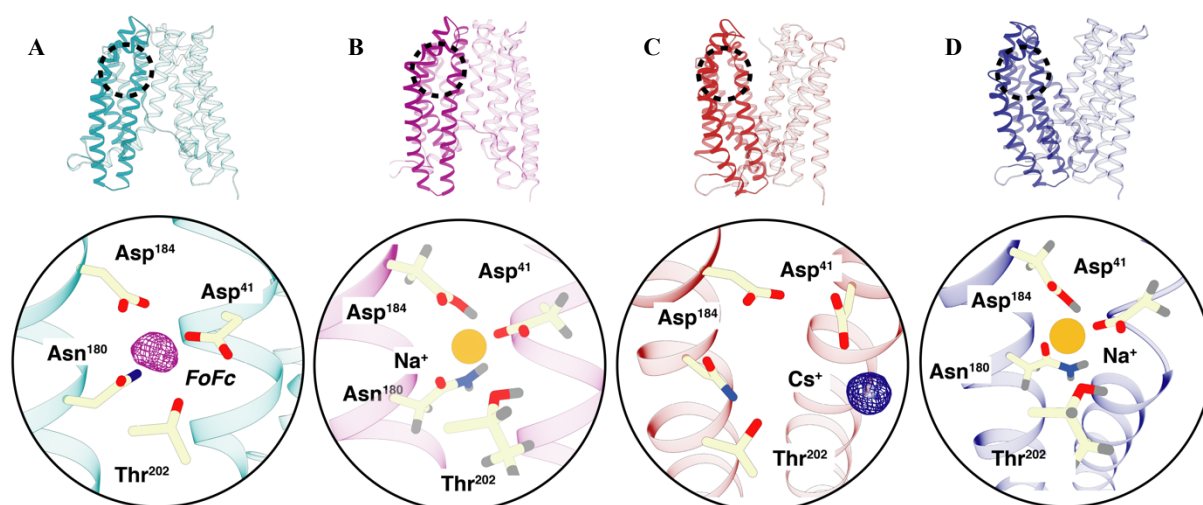


Fig. 4.13. Sodium ion binding site in PfMATE. (A-B) Sodium ion binding site in the inward-facing state based on the crystal structure (PDB ID: 6FHZ) (A) and the representative MD snapshot (B). The residues labeled and shown by a stick model coordinate the bound Na⁺, depicted as an orange sphere. The purple mesh represents the Fo-Fc electron density peak, which is tentatively assigned to the Na⁺ ion. (C-D) Sodium ion binding site in the outward-facing state based on the crystal structure (PDB ID: 6HFB) (C) and an MD snapshot of the fully coordinated sodium ion (taken from the simulation at 700 ns) (D). The coordinating residues are shown in stick representation, whereas the blue mesh shows the anomalous peak for Cs⁺.

The presence of a monovalent cation binding site in the N-domain of PfMATE was furthermore supported by our MD simulations (within a collaboration with Dr. Ahmad Reza Mehdipour from the Theoretical Biophysics Department of the Max Planck Institute of Biophysics), which show a Na⁺ binding site near Asp41 (pK_a=5.1) and Asp184 (pK_a>14) in the outward-facing

structure (Fig. 4.13B). The calculated pK_a values for the inward-facing structure strongly imply that in the absence of sodium ions both Asp41 (pK_a=13.7) and Asp184 (pK_a>14) are protonated, while in the presence of sodium ions, Asp41 (pK_a=0) is most likely charged. Results of the MD simulations in the presence of sodium ions indicate a well-coordinated binding site, while in the absence of sodium ions the protonation of both residues, Asp41 and Asp184, is essential for the stability of the binding site. In light of the high-calculated pK_a of Asp41 in the inward-facing structure (pK_a=13.7 in the absence of sodium ions), Na⁺ release could be coupled to Asp41 protonation. Structuring of the unwound TM1 segment could be another relevant factor.

The MD runs of the outward-facing straight structure (6GWH) showed spontaneous Na⁺ binding to the corresponding Cs⁺ site of the crystal structure solved by SAD (6HFB) (Fig. 4.13D). After ion recognition, the binding site did not immediately close and the sodium ion created four interactions with the protein (two with Asp41 and one each with Asn180 and Thr202), but retained two water molecules. Later, the missing interactions with Asp184 and Ala198 were formed, replacing the two water molecules. However, the interactions of bound Na⁺ with Asp184 and Ala198 are transient. These simulations resulted in a coordination structure as in the inward-facing state (6FHZ) as well as in the previously published simulation of the outward-facing structure (3VVO). Interestingly, in the fully coordinated state, TM1 is slightly bent, adopting an intermediate conformation between 3VVO and 3VVN. Combination of the crystallographic data with the results of the MD simulations implies a sodium ion dependent transport mechanism of PfMATE. However, we cannot exclude a dual specificity for sodium ions and protons, which was proposed in the recent publication demonstrating a dual ion coupling behavior for NorM_VC⁶⁵.

In light of ion binding events, spontaneous binding of a chloride ion into the cavity in the inward-facing structure was also observed during the MD simulations. The chloride interacts with Arg284, Thr35, Asn38, Trp283, and Met287 as shown in Fig. 4.14.

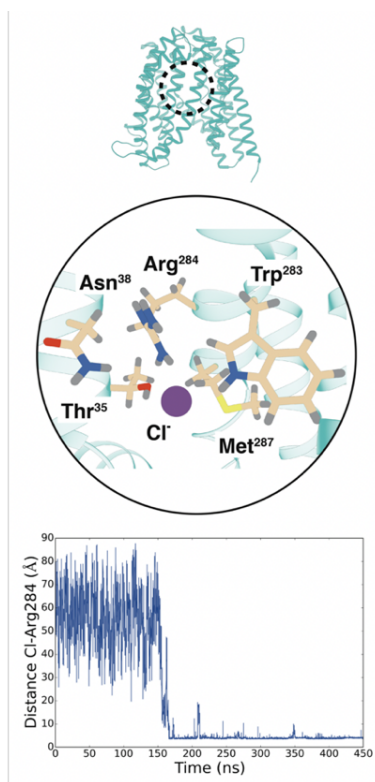


Fig. 4.14. Chloride ion binding site. Graphs represent the distance between chloride ion and chloride binding site in the spontaneous Cl⁻ binding during MD simulations of the inward-facing PfMATE state. The coordinating residues are labeled and shown as a stick representation.

Interestingly, a chloride ion was observed near Arg24 and Arg255 in the MurJ inward-facing structure, which corresponds to a similar position of PfMATE after superposition of these two structures (Fig. 4.15). Crystal structure of MurJ from the bacterium *Thermosipho africanus* represents an inward-facing conformation of the MVF family lipid II flippase from the MOP superfamily. In contradiction to PfMATE, MurJ contains two additional helices in the C-domain, TMs 13 and 14, which form a hydrophobic groove. Superposition of PfMATE and MurJ reveals that both structures share the same orientation of the binding cavity, which is exposed to the intracellular space. Currently, the functional importance of the chloride ion for PfMATE and MurJ remains unclear. Nevertheless, it may be possible that the bound chloride is released to the external side during the transition to the outward facing state in a way that

MATE (or MOP transporters in general) transporters are actually [substrate-chloride]/sodium ion antiporters. A co-transport of a chloride ion would be energetically beneficial by contributing driving force to the transport. In order to elucidate a chloride ion binding site experimentally, purification of PfMATE in the presence of NaBr instead of NaCl in purification buffers and subsequent crystallization trials were initiated at the final stage of my PhD studies. The aim of this experiment is to detect an anomalous signal of a bromide, which would correspond to a position of the chloride binding site in PfMATE.

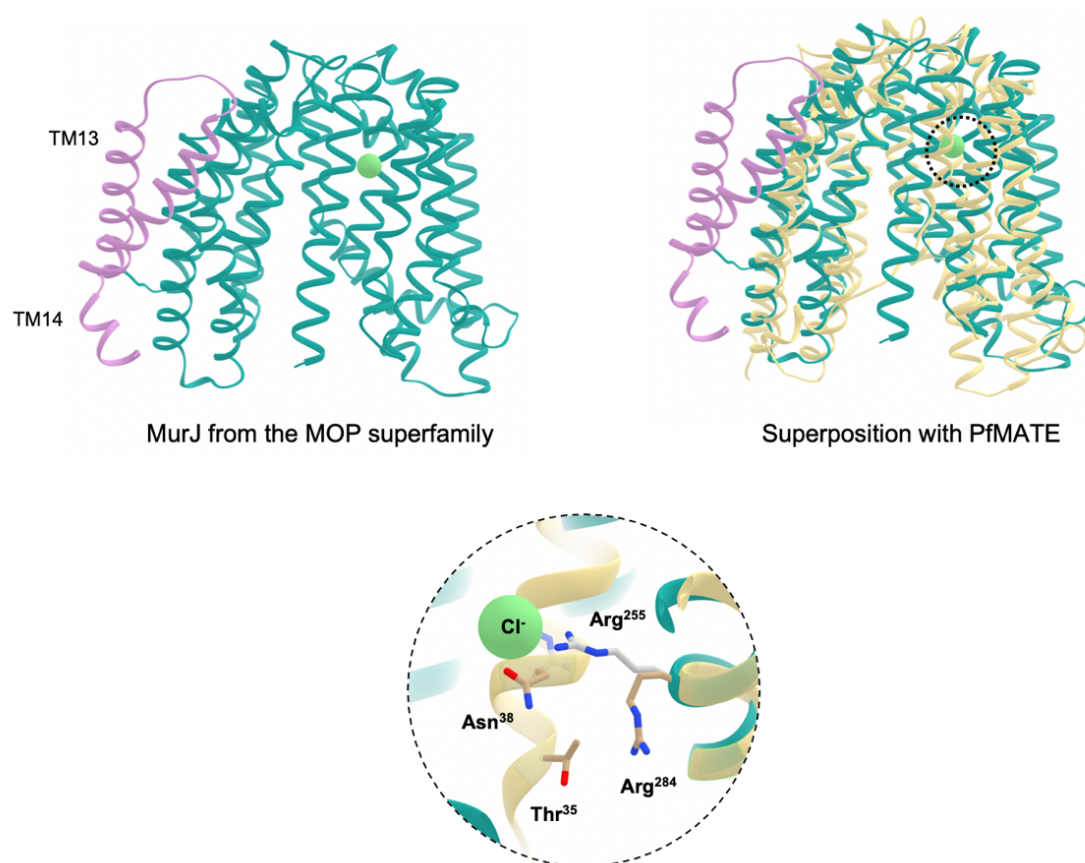


Fig. 4.15. Structural superposition of MurJ and PfMATE inward-facing structures. Ribbon model of PfMATE is depicted in pink, whereas MurJ is shown in dark green. Light green sphere represents chloride ion in MurJ crystal structure near Arg24 and Arg255. According to the MD simulations on PfMATE, Cl⁻ is located near Arg284 and Asn38, which are shown in a stick representation.

4.6. Interaction of PfMATE with the archaeal lipids

Lengthy electron densities were observed in the map of the inward-facing conformation near the intramembrane region of TM1 and helices TM8 and TM9, which may originate from lipid molecules. At the present resolution, these densities cannot be unambiguously assigned to lipids, thus, they were not included during the structure refinement. It is widely known that lipids are very often poorly resolved and/or difficult to distinguish from other molecules present in purification buffers or crystallization media, such as detergents. The local environment for each lipid can influence the conformation of the lipid tails, their mobility and disorder. Due to their flexibility or multiple alternative conformations, lipids can result in weak signals in the electron density maps. Therefore, structural information on the lipid-protein interaction is challenging to obtain. At the current stage, specific lipid binding site in PfMATE cannot be empirically identified, and it still remains unknown how exactly the addition of the native lipids leads to crystallization of PfMATE in the inward-facing state. Nevertheless, having established that native lipids act like conformational modulators of PfMATE, MD simulations of this protein embedded in a lipid bilayer of archaeal lipids were carried out to assess the protein-lipid interactions.

P. furiosus, is a strictly anaerobic and hyperthermophilic archaeon. To cope with the elevated temperature, pressure, and lack of oxygen, this microorganism has evolved certain molecular mechanisms, which allow surviving at extreme conditions. One of the fundamental thermal adaptations is an unusual membrane lipid composition. The mass spectrometry based lipidomics reveals that *P. furiosus* membranes consists of DPI (diphytanyl phosphatidyl inositol), HexNac-PD (diphytanyl phosphatidyl *N*-acetyl hexose), DPG (diphytanyl phosphatidyl glycerol), DPA (diphytanyl phosphatidic acid) and isoGDGT-0 (isoprenoidal glycerol dialkyl glycerol tetraether) (Fig. 3.18). Based on these results, MD simulations were performed to study PfMATE and lipids interplay under native-like conditions at elevated

temperature (100 °C). PfMATE was embedded in the archaeal-type lipid bilayer, consisting of DPI, DPG, and DPA lipid species (45:20:35). The MD simulation results for the outward-facing state show the lipid molecules moving into the cavity and being accommodated by the positively charged pocket. The lipid access pathway based on these simulations is shown in Fig. 4.16.

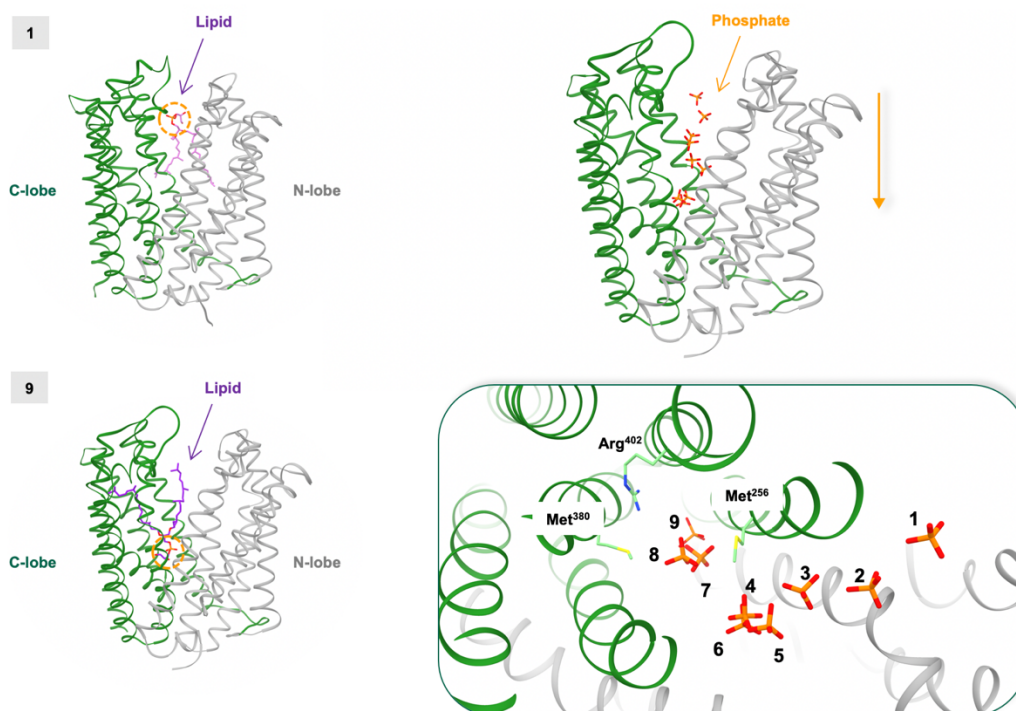


Fig. 4.16. Lipid access pathway. MD simulations snapshots of the lipid access pathway into the cavity of PfMATE. Selected snapshots are numbered from 1 to 9. For clarity, DPG lipid tails are removed in the right panel.

During several different MD simulation runs, a complete reorientation of negatively charged lipids (mostly PG species) inside the cavity with the head group moiety pointing towards the bottom of the cavity can be observed (Fig. 4.17). In these runs, the lipid entered the cavity from the outer leaflet. Initially, the head group of lipid intruded into the cavity several times and flipped back, and interacted mainly with Arg284 and Trp283. Subsequently, predominantly electrostatic interaction between the head group and Arg402 resulted in a further movement of the head group inside the cavity, which finally triggered the full reorientation of the lipid. In

the flipped orientation, the head group interacts with several residues at the bottom of the cavity including Arg161, Thr399, Arg402, Asn253, and Gln387. Furthermore, a cluster of methionine residues, Met31, Met64, Met256, Met260, and Met287, coordinates the aliphatic tails of the lipid. The lipid remained bound into the cavity for the rest of the simulation (up to 1.5 μ s) with its headgroup stabilized by hydrogen bonds and electrostatic interactions and its tail by abundant hydrophobic interactions with the protein cavity. Lipid binding was observed only for the outward-facing structure. In the inward-facing structure, closed lateral gates block lipid access to the cavity. These results demonstrate the existence of a pathway for lipid access into the central cavity of PfMATE and the possibility of lipid flipping.

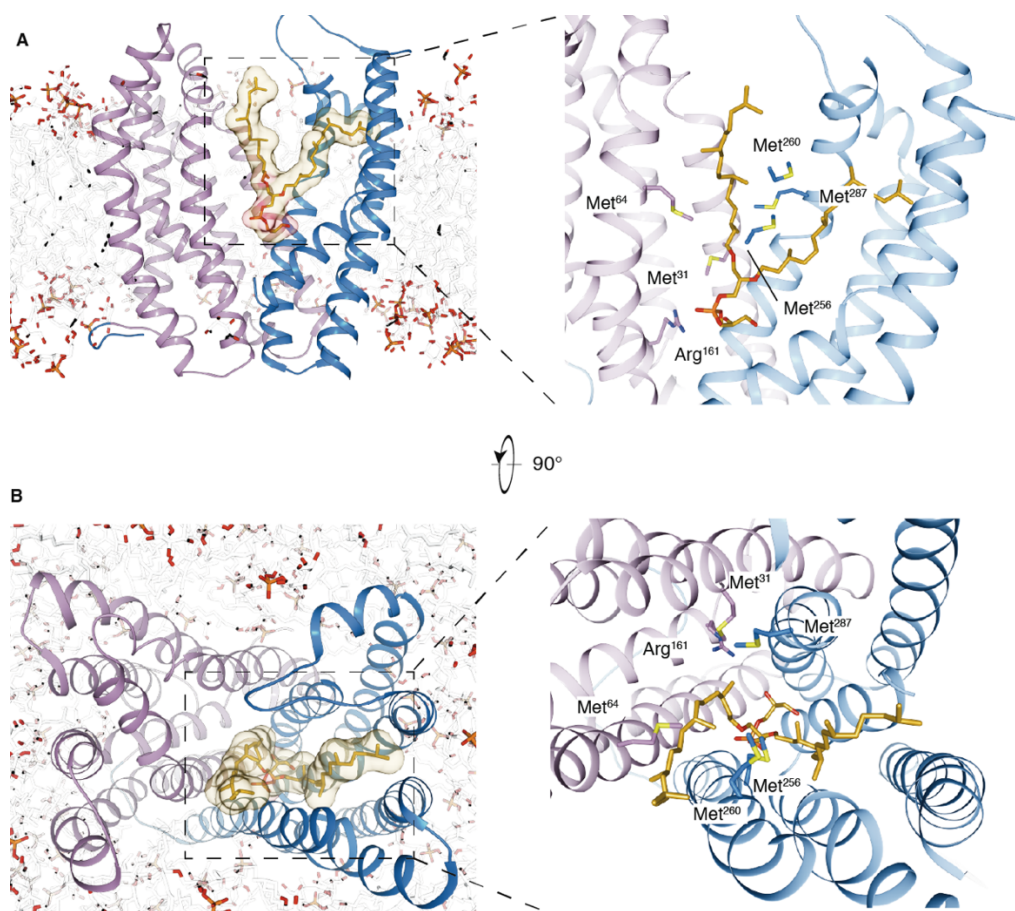


Fig. 4.17. Spontaneous binding and flipping of an archaeal lipid in molecular dynamics simulations of PfMATE. (A) Side view and (B) top view of of the representative MD snapshot showing the interaction of PfMATE (N-domain in pink and C-domain in blue) with the archaeal-type PG lipid that entered the cavity from the outside leaflet and spontaneously flipped its orientation. Zoom-in on the right highlights residues contacting the lipid.

Differential scanning fluorimetry (DSF) experiment (Section 2.25) using label-free nanoDSF Prometheus device is based on the fluorescence change at emission wavelength 330 nm and 350 nm of Trp and Tyr residues in a protein upon heating from 20 °C to 95 °C. In order to assess thermal stability of PfMATE, two different strategies were applied:

1. The purified PfMATE samples were measured in the absence or presence of different concentrations of total lipid extract from *P. furiosus*.
2. The PfMATE sample after affinity chromatography was firstly incubated with 2% (v/v) OGNG (Octyl Glucose Neopentyl Glycol) at 4 °C overnight. This detergent is commonly used during sample preparation for native mass spectrometry in order to remove lipids surrounding membrane proteins. The OGNG-treated sample was purified by gel filtration in a buffer consisting of 200 mM ammonium acetate (pH 8.0) and 0.15% (v/v) OGNG. Finally, delipidated PfMATE samples were analyzed with or without addition of the native lipids.

As shown in Fig. 4.18., the T_m (melting point) value of PfMATE corresponding to half of the protein molecules being unfolded is increased in the presence of the lipid extract from *P. furiosus*. Measurements of the OGNG-treated samples reveal that removal of lipids destabilizes PfMATE dramatically, reducing its T_m up to 9 °C in comparison with the non-treated sample (Fig. 4.19). When delipidated PfMATE was supplemented with the native lipids, the thermal stability of PfMATE was recovered. The transition midpoint shifted in response to addition of lipids in a concentration dependent manner.

DSF results of the delipidated and lipid-supplemented samples clearly show that native lipids have an effect on the stability of PfMATE, which points to the lipid binding events occurring. The stabilizing effect may be also caused by lipid acting as substrates for PfMATE.

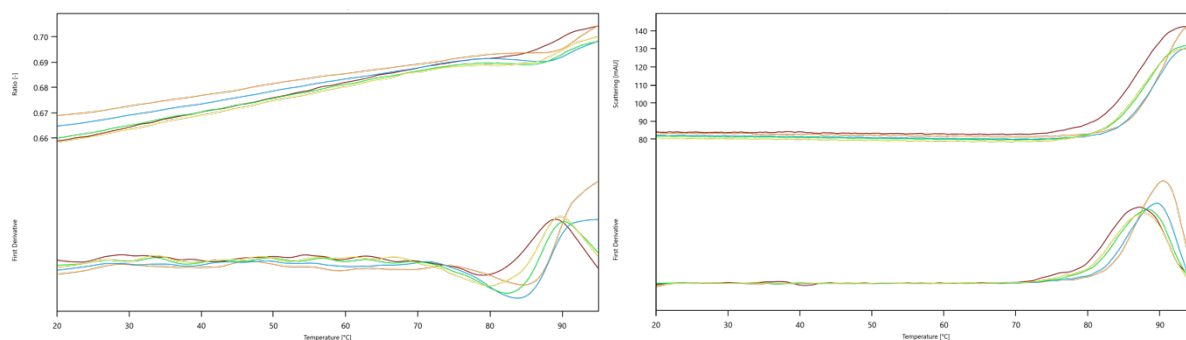


Fig. 4.18. Stability curves from DSF results of PfMATE with and without the native lipid extract. Left, F350/F330 ratio; right, scattering. PfMATE without addition of the lipid extract (red) and in the presence of the lipid extract 1:16 (orange), 1:32 (blue), 1:64 (green), 1:128 (yellow).

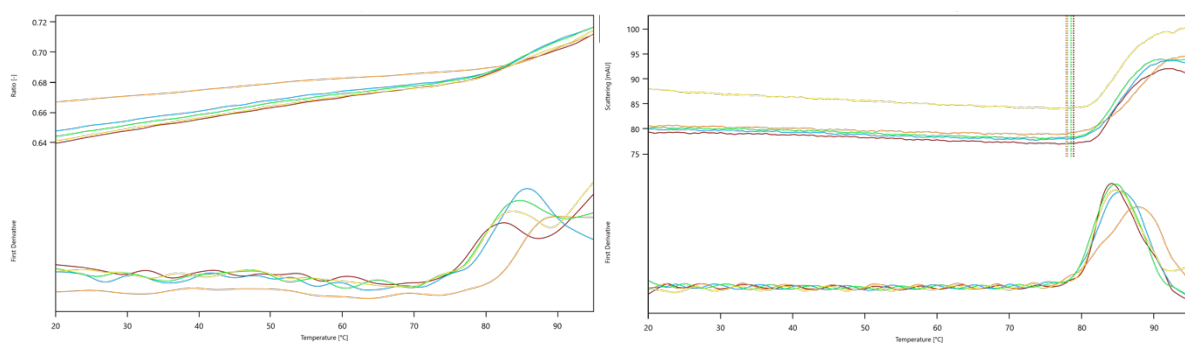


Fig. 4.19. Stability curves from DSF results of the OGNG-treated PfMATE with and without the native lipid extract. Left, F350/F330 ratio; right, scattering. PfMATE without addition of the lipid extract (red) and in the presence of the lipid extract 1:16 (orange), 1:32 (blue), 1:64 (green), 1:128 (yellow).

Another experimental approach taken in the course of this dissertation is native mass spectrometry, allowing determination of the oligomeric state and identification of tightly bound lipids to the protein molecules^{159–162}. The lipid binding to PfMATE was assessed by this technique within a collaboration with Dr. Jani Reddy Bolla from the University of Oxford. The idea behind the native MS experiment is based on removal of the shell lipids (annular lipids), which surround the membrane protein and its hydrophobic surface while remaining the specific lipid molecules which are bound with high affinity to the protein and may be essential for its

function. If the interaction between the membrane protein and the lipid molecule is strong, then this lipid molecule has slow-off rate compared to bulk lipids. Firstly, the delipidation protocol was optimized so that it was possible to apply relatively mild MS conditions while still obtaining well-resolved spectra. Also, detergent screen was performed to select one, LDAO (N,N-dimethyl-N-dodecylamine-N-oxide), which yielded resolvable spectra of PfMATE. The protein was released from the detergent micelles and then carried into the gas phase, which is unsuitable to retain non-covalent interactions.

For the native MS experiments, two different strategies were applied:

1. The MS measurements of PfMATE co-purified by size exclusion chromatography with the native lipid extract from *P. furiosus* and without addition of lipids. Protein samples were prepared in the same manner as for the crystallization trials leading to PfMATE structure determination in the outward and inward-facing states (section 3.2.2).
2. In the second approach, prior to the MS measurements the apo-PfMATE after SEC purification step was incubated with increasing amounts of the *P. furiosus* lipid extract.

The theoretical mass of a monomeric PfMATE is 52176.62 Da. Well-resolved charge states were observed with the anticipated masses corresponding to the dilipidated protein (Fig. 4.20). The mass spectra of PfMATE indicate that this protein is monomeric in solution, consistent with the crystal structures presented in the subsections 3.2.7 and 3.2.8. In case of the protein sample purified without addition of the native lipids, a complex formation with cardiolipin as well as lipopolysaccharide (LPS) was detected, which means that they have relatively high affinities for PfMATE. These lipid molecules are likely to remain bound to PfMATE during the purification process and originate from *E. coli* used as expression host.

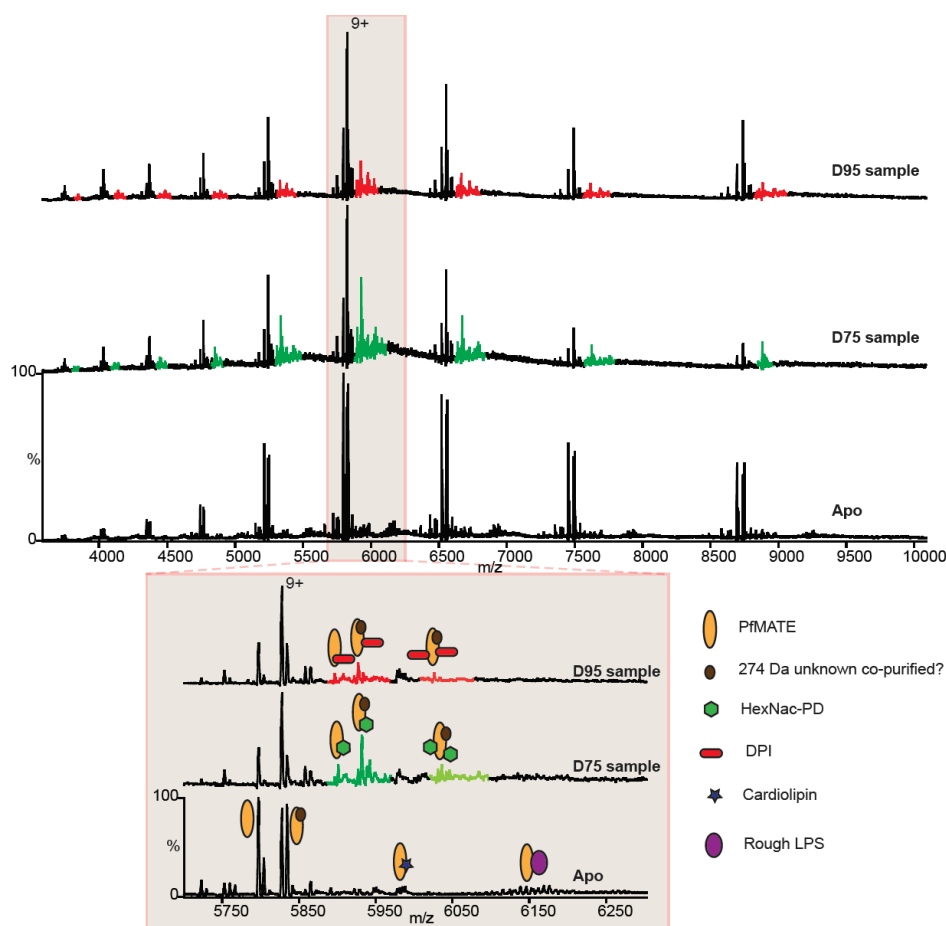


Fig. 4.20. Mass spectra of PfMATE in the presence and absence of archaeal lipids. Charge state series consistent with the mass of the monomeric protein bound to different lipid species. Addition of archaeal lipids D75 and D95 leads to formation of a complex with PfMATE.

The mass spectra of the sample co-purified with lipid extract D95 composed of DPI: DPG: DPA: HexNac-PD (93.3%:1.68%:0.21%:4.81%), and lipid extract D75 composed of DPI: DPG: DPA: HexNacPD (3.4%:4.9%:8.2%:81%) indicate binding of up to two lipid molecules to PfMATE. Theoretical masses of these lipids are as follows: DPI (894.7 MW), DPG (806.7 MW), DPA (732.6 MW) and HexNacPD (935.7 MW). The additional peaks were assigned to adducts of masses corresponding to DPI and Hex-Nac-PD in case of the PfMATE sample co-purified with D95 and D75, respectively. These lipids remained tightly bound to PfMATE, have survived the purification and remained associated to the protein in the gas phase. No *E. coli* lipids binding was evident anymore as previously observed for the apo-PfMATE. These

observations points to the reasonable conclusion that lipids bind tightly to this transporter, where the lipids from the native source bind with a higher affinity than the co-purified ones from the expression organism.

Another experiment shows titration of PfMATE with D95_2 (DPI: DPG: DPA: HexNac-PD (59.6%:3.9%:2.6%:28.5%)) in a saturable manner, which also reveals specific lipid binding events to PfMATE (Fig. 4.21). Additional peaks with a low relative intensity compared to the unbound form were observed with a mass difference corresponding to lipid DPI bound to PfMATE. At 1.25 μl of the added lipid extract, the lipid-bound species predominate over the unbound form of the protein.

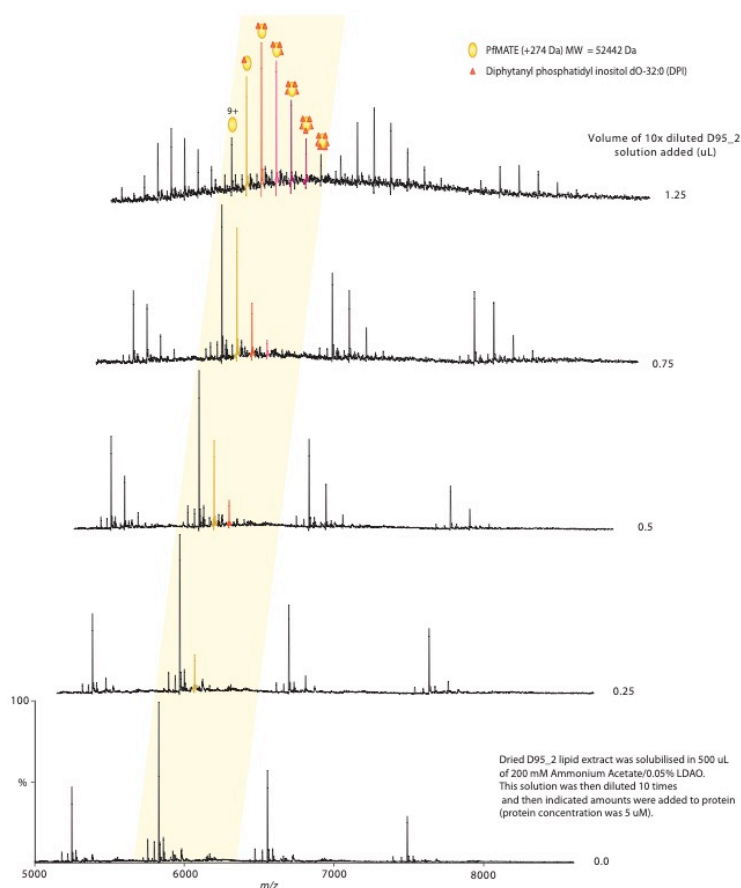


Fig. 4.21. Mass spectra of PfMATE with increasing concentrations of archaeal lipids. A charge series is observed (yellow), corresponding to bound lipids, which increases in its intensity as the lipid concentrations increase. A second lipid binding peak (orange) emerges at lipid amount above 0.5 μl , a third peak (pink) above 0.75 μl and two more (purple and red) above 1.25 μl .

The mass spectra recorded for solutions of PfMATE co-purified with archaeal lipids provide a solid support that the protein sample applied for the crystallization trials leading to the PfMATE structure determination in the inward-facing state contains tightly bound lipid molecules. The protein-lipid complex survives additional purification steps confirming that it remains associated in solution. In light of these observations, lipids could either stabilize the overall architecture of PfMATE or have an effect on its function. The native MS results are supported by the MD simulation data, which demonstrate specific lipid binding events to the residues in the central binding cavity of PfMATE. Based on the native MS data, it is apparent that the lipids are tightly bound despite extensive delipidating procedures, however, at the current stage it could not be determined experimentally where exactly the lipids bind. For clarifying this point, future experiments described in the Chapter 5 would be useful to identify the specific lipid binding site in PfMATE.

4.2. Functional studies on PfMATE

In the previous publication by Tanaka *et al.* two outward-open structures of PfMATE, with a straight or bent TM1 form, were presented. These studies also include functional data that PfMATE confers norfloxacin resistance to the *Escherichia coli* BW25113 Δ *acrB* strain. Interestingly, the lipid or monoolein molecules were observed in the straight and bent outward-facing structures of PfMATE (3VVO and 3VVN, respectively). The binding site of two monoolein molecules observed exclusively in the N-domain of their straight structure (3VVN) corresponds to the drug recognition site in the Br-Norfloxacin bound structure (3VVP). The authors suggested that these molecules may mimic the hydrophobic substrates.

As described in the Section 3.3, the MIC tests were performed in the following strains: *E. coli* KAM32, *E. coli* C43(DE3) Δ *acrAB*, *E. coli* BL21(DE3) Δ *acrAB*, *E. coli* C41(DE3) Δ *acrAB*. Interestingly, only *E. coli* C43(DE3) Δ *acrAB* strain containing pBAD-PfMATE showed some increased norfloxacin resistance compared to the cells with the empty vector, however, the difference is not as significant as compared to the cells containing the gene coding for the MATE transporter NorM_VP from *Vibrio parahaemolyticus* used as a positive control. These results are also similar to those of functional studies on ClbM transporter from the MATE family, which did not provide resistance to ethidium bromide in *E. coli* KAM32 cells, however a resistance was observed in *E. coli* C43(DE3) Δ *acrAB* cells. Despite the high structural resemblance to the other MATE family transporters, in particular to PfMATE, ClbM has an additional unique function of precolibactin transport.

C41(DE3) Δ *acrAB* and C43(DE3) Δ *acrAB* are mutant strains derived from the parent *E. coli* BL21 (DE3) Δ *acrAB* strain (Fig. 4.22). These two strains are commonly used to overcome the problem of plasmid instability and toxicity associated with overexpression of recombinant proteins. In order to clarify if the lipid composition affect the function of PfMATE, the cell pellet of each strain was recently sent for the lipidomics analysis. These results will provide the

answer whether there is any difference in the lipid content among the *E. coli* strains used for the MIC tests and how is it related to the PfMATE activity.

It is interesting to note that based on the phylogenetic analysis of the MOP superfamily transporters, the MATE family transporters are most closely related to the prokaryotic MVF family¹⁶³, which is involved in lipid transport according to the alternate access mechanism¹⁶⁴. It is possible that MATE transporters may be involved in similar events and be capable of transporting lipids apart from extruding drugs.

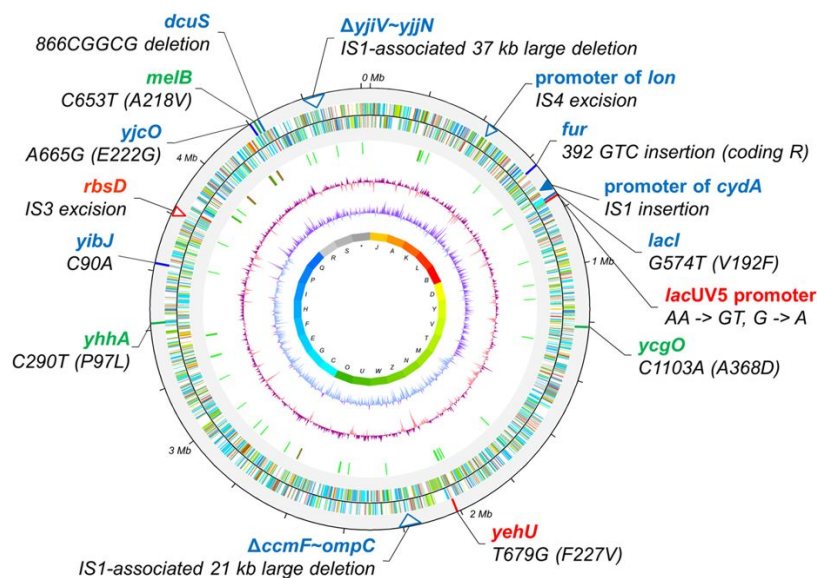


Fig. 4.22. Genetic changes in *E. coli* C41(DE3) and C43(DE3) comparative with BL21(DE3). Mutations found in both C41(DE3) and C43(DE3) are shown in red, C41(DE3) in green, and C43(DE3) in blue. Small-scale changes like SNPs or small DIPs are indicated with solid lines, and IS element-mediated large-scale insertions or deletions with triangles. Picture adapted from Kwon *et al.*¹⁶⁵

5. CONCLUDING REMARKS

The main highlight of the aforementioned results is the PfMATE structure determination of the inward-facing conformation. This novel state of the MATE transport cycle provides the structural evidence for the postulated alternating access model, in which rocking of the N and C-terminal bundles closes and opens the intracellular and extracellular barriers. The structural analysis reveals that TM1 is remarkably bent towards helices in the C-domain and undergoes helical unwinding of the N-terminal end of the transmembrane region. This alteration displays a highly dynamic nature and conformational flexibility of this segment, which most likely plays a crucial role for the substrate and ion gating. For the first time, structure determination of PfMATE in an inward-facing conformation has delivered valuable structural insights into a novel intermediate of the MATE family transporter. Based on the combination of crystallographic and MD simulations data of two distinct conformational states of PfMATE, the plausible transport mechanism was discussed.

Protonation and deprotonation of the conserved residue Asp41 has been previously postulated to be the main trigger for TM1 transition between the straight and bent outward-open conformation. However, TM1 remains straight in the crystal structure of the outward-facing conformation obtained from the crystals grown under low pH conditions (5.0) in the absence of native lipids from *P. furiosus*. Co-crystallization of PfMATE with native lipids at the same acidic pH led to the structure determination in an inward-oriented state. Hence, these observations implicate a plausible rationale that lipid extract from *P. furiosus* acts as a key contributor for the mechanistic rearrangement between the outward and inward-facing state. With regard to the ion specificity, the structural and computational results lead to the identification of a sodium ion binding site at Asp41, supporting a sodium ion coupled transport mechanism of PfMATE. In the light of the aforementioned functional and structural studies on

PfMATE as well as the native MS and MD simulation data, we cannot exclude that the primary function of PfMATE may be lipid transport. The interaction of lipid molecules with PfMATE opens an alternative view on the mechanism of action for the MATE family members, in which specific lipid species could either modulate the transition between the outward and inward-facing state, or even act as potential substrates themselves, expanding the substrate spectrum for the MATE proteins. The interplay between membrane proteins and lipids can affect membrane protein topology, structure and function.

A detailed analysis of two completely distinct states provides mechanistic details of the interactions involved in the extracellular and intracellular barrier formation, which is relevant from the clinical point of view. A comprehensive knowledge about different intermediates of MATE transporters creates a strong framework for a rational and structure-aided drug design in order to tackle global health challenges such as bacterial multidrug resistance. The aforementioned approach of using physiologically relevant lipids, allowed trapping and visualization of the novel conformational state of the MATE family protein. The insights from the structural and computational analysis pave the way towards better understanding of the underlying steps of the PfMATE transport cycle and strongly favor the conclusion that native lipids from the corresponding source organism may be essential for function and structural dynamics of the MATE (or MOP) transporters in general.

Although this work provides new insights and expands our knowledge of the MATE family transporter, PfMATE from *Pyrococcus furiosus*, there are still open questions, which can be addressed by performing future experiments.

1. The native MS experiments should be performed with different variants of the PfMATE, which may be useful for the identification of critical amino acid residues responsible for the binding of lipids.

2. Further MD simulations on the conformational change between the inward and outward facing conformation can provide more insights into the mechanism of action, including ion-coupling, structural alterations and lipid protein interplay. These experiments can be performed by embedding PfMATE in the lipid bilayer composed of different lipids species in different ratios.
3. Co-crystallization with various lipid species can be further elaborated to determine lipid specificity and selectivity in the light of PfMATE conformational modulation as well as substrate recognition. Continuation of the structural studies on the PfMATE produced homologically can deliver valuable information about the physiological state of this transporter. The cubicon method may be particularly useful for the low concentrated samples. Although lower quantity of the final protein sample would not be a hurdle for the cryoEM method, the molecular weight of around 50 kDa is still a limiting factor for analyzing PfMATE by this technique. Taking into consideration the resolution revolution¹⁶⁶ and technological advances in the field of cryoEM¹⁶⁷, including development of detectors or Volta phase plates¹⁶⁸, this method may be an alternative approach in the future to capture more distinct intermediates representing missing steps underlying the PfMATE transport cycle. The size of protein particles can be increased, for example by using a recently published Megabody technology. This approach allowed less preferential orientation of protein particles and structure determination of GABA_A receptor by using a scaffold protein enlarging the nanobody against this receptor¹⁶⁹.
4. Noteworthy, Martens *et al.* recently proposed a model for the lipid-induced conformational equilibrium of secondary active transporters based on hydrogen-deuterium exchange (HDX) mass spectrometry and MD simulation experiments. This

work particularly described how lipid binding to conserved networks of charged residues induces the conformational transition towards an inward-facing conformation¹⁷⁰. The initial HDX experiments were performed, however, due to a one-year waiting time for the data, it was impossible to continue this approach within my PhD timeline.

5. Taking into consideration that the majority of the lipids used for the co-purification and co-crystallization with PfMATE are phospholipids, the anomalous scattering of crystals obtained in the presence of lipids at the wavelength corresponding to the phosphorous edge, can allow identification of phosphorous atoms, indicating the location of the bound phospholipid headgroups. This work has been already initiated at the high wavelength beamline I23¹⁷¹ within a collaboration with Dr. Ramona Duman from the Diamond Light Source (DLS) in Oxford. The higher wavelength, the lower energy and the data quality becomes greatly reduced. The first attempts at 3.5 Å resulted in the crystals diffracting up to 3.7 Å. However, in order to differentiate between S and P atoms ($\lambda = 5.015$ and 5.779 Å, respectively), the data should be collected at the wavelength below the absorption edge of sulfur (Fig. 5.1), which is not feasible at present.

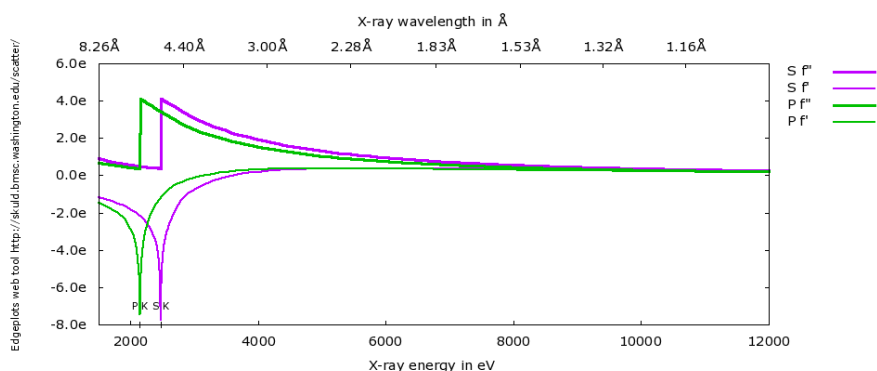


Fig. 5.1. X-ray anomalous scattering coefficients of phosphorous (P) and sulfur (S).

6. COLLABORATIVE WORK

Except where stated otherwise by reference or acknowledgment, the work presented was generated by myself under the supervision of my advisors during my doctoral studies. All contributions from colleagues are explicitly referenced in the thesis. The material listed below was obtained in the context of collaborative research:

Fig. 3.19: Lipid composition of *P. furiosus* WT analyzed by mass spectrometry based lipidomics

PD Dr. Harald Köfeler (Medical University of Graz): MS-based lipidomics;

my own contribution: cultivation of *P. furiosus* cells

Fig. 4.8: Conformational changes of PfMATE

Fig. 4.14: Chloride ion binding site

Fig. 4.16: Lipid access pathway

Fig. 4.17: Spontaneous binding and flipping of an archaeal lipid in molecular dynamics simulations of PfMATE

Dr. Ahmad Reza Mehdipour (Max Planck Institute of Biophysics): MD simulations;

my own contribution: coordinates of the PfMATE structures

Fig. 4.20: Mass spectra of PfMATE in the presence and absence of the archaeal lipids

Fig. 4.21: Mass spectra of PfMATE with increasing concentrations of the archaeal lipids

Dr. Jani Reddy Bolla (University of Oxford): native MS measurements;

my own contribution: PfMATE samples preparation

Whenever a figure, table or text is identical to a previous publication, it is stated explicitly in the thesis that copyright permission and/or co-author agreement has been obtained.

The following parts of the thesis have been previously published:

Figures “1.15-1.18, 2.1, 2.3, 2.4, 3.13, 4.22”

7. REFERENCES

1. Cournia, Z. *et al.* Membrane Protein Structure, Function, and Dynamics: a Perspective from Experiments and Theory. *Journal of Membrane Biology* (2015). doi:10.1007/s00232-015-9802-0
2. Kubicek, J., Block, H., Maertens, B., Spriestersbach, A. & Labahn, J. Expression and purification of membrane proteins. in *Methods in Enzymology* (2014). doi:10.1016/B978-0-12-420119-4.00010-0
3. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* (80-.). (2015). doi:10.1126/science.1260419
4. Yin, H. & Flynn, A. D. Drugging Membrane Protein Interactions. *Annu. Rev. Biomed. Eng.* (2016). doi:10.1146/annurev-bioeng-092115-025322
5. Al, L. *Molecular Cell Biology (5th edition)*. *Biochemistry and Molecular Biology Education* (2001). doi:10.1016/S1470-8175(01)00023-6
6. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* (2006). doi:10.1038/nrd2199
7. Organization, W. H. *Antimicrobial resistance: global report on surveillance*. *WHO Report* (2014). doi:1.4.2014
8. Li, X. Z., Plésiat, P. & Nikaido, H. The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin. Microbiol. Rev.* (2015). doi:10.1128/CMR.00117-14
9. Levy, S. B. Active efflux mechanisms for antimicrobial resistance. *Antimicrob. Agents Chemother.* (2012). doi:10.1128/aac.36.4.695
10. Poole, K. Efflux-mediated antimicrobial resistance. in *Antibiotic Discovery and Development* (2014). doi:10.1007/978-1-4614-1400-1_10
11. Alekshun, M. N. & Levy, S. B. Molecular mechanisms of antibacterial multidrug resistance. *Cell* (2007). doi:10.1016/j.cell.2007.03.004
12. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* (2003).
13. Zgurskaya, H. I. & Nikaido, H. Multidrug resistance mechanisms: Drug efflux across two membranes. *Molecular Microbiology* (2000). doi:10.1046/j.1365-2958.2000.01926.x
14. Masi, M., Réfregiers, M., Pos, K. M. & Pagès, J. M. Mechanisms of envelope permeability and antibiotic influx and efflux in Gram-negative bacteria. *Nat. Microbiol.*

- (2017). doi:10.1038/nmicrobiol.2017.1
15. Li, X. Z. & Nikaido, H. Efflux-mediated drug resistance in bacteria: An update. *Drugs* (2009). doi:10.2165/11317030-000000000-00000
 16. Ahmad, I. *et al.* Bacterial Multidrug Efflux Proteins: A Major Mechanism of Antimicrobial Resistance. *Curr. Drug Targets* (2018). doi:10.2174/1389450119666180426103300
 17. Locher, K. P. Structure and mechanism of ABC transporters. *Current Opinion in Structural Biology* (2004). doi:10.1016/j.sbi.2004.06.005
 18. ter Beek, J., Guskov, A. & Slotboom, D. J. Structural diversity of ABC transporters. *J. Gen. Physiol.* (2014). doi:10.1085/jgp.201411164
 19. Locher, K. P. Review. Structure and mechanism of ATP-binding cassette transporters. *Philos. Trans. R. Soc. Biol. Sci.* (2009). doi:10.1098/rstb.2008.0125
 20. Pao, S. S., Paulsen, I. T. & Saier, M. H. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* (1998).
 21. Zomot, E. *et al.* A New Critical Conformational Determinant of Multidrug Efflux by an MFS Transporter. *J. Mol. Biol.* (2018). doi:10.1016/j.jmb.2018.02.026
 22. Yan, N. Structural Biology of the Major Facilitator Superfamily Transporters. *Annu. Rev. Biophys.* (2015). doi:10.1146/annurev-biophys-060414-033901
 23. Daury, L. *et al.* Tripartite assembly of RND multidrug efflux pumps. *Nat. Commun.* (2016). doi:10.1038/ncomms10731
 24. Nikaido, H. & Takatsuka, Y. Mechanisms of RND multidrug efflux pumps. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2009). doi:10.1016/j.bbapap.2008.10.004
 25. Ruggerone, P., Murakami, S., M. Pos, K. & Vargiu, A. RND Efflux Pumps: Structural Information Translated into Function and Inhibition Mechanisms. *Curr. Top. Med. Chem.* (2013). doi:10.2174/15680266113136660220
 26. Brown, M. H., Paulsen, I. T. & Skurray, R. A. The multidrug efflux protein NorM is a prototype of a new family of transporters [2]. *Molecular Microbiology* (1999). doi:10.1046/j.1365-2958.1999.01162.x
 27. Chen, J. *et al.* VmrA, a member of a novel class of Na⁺-coupled multidrug efflux pumps from *Vibrio parahaemolyticus*. *J. Bacteriol.* (2002). doi:10.1128/JB.184.2.572-576.2002
 28. Kuroda, T. & Tsuchiya, T. Multidrug efflux transporters in the MATE family. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2009). doi:10.1016/j.bbapap.2008.11.012
 29. Staud, F., Cervený, L., Ahmadimoghaddam, D. & Ceckova, M. Multidrug and toxin

- extrusion proteins (MATE/SLC47); role in pharmacokinetics. *International Journal of Biochemistry and Cell Biology* (2013). doi:10.1016/j.biocel.2013.06.022
30. Bay, D. C. & Turner, R. J. Small Multidrug Resistance Efflux Pumps. in *Efflux-Mediated Antimicrobial Resistance in Bacteria* (2016). doi:10.1007/978-3-319-39658-3_3
31. Bay, D. C. & Turner, R. J. Diversity and evolution of the small multidrug resistance protein family. *BMC Evol. Biol.* (2009). doi:10.1186/1471-2148-9-140
32. Hassan, K. A. *et al.* Pacing across the membrane: the novel PACE family of efflux pumps is widespread in Gram-negative pathogens. *Res. Microbiol.* (2018). doi:10.1016/j.resmic.2018.01.001
33. Delmar, J. A. & Yu, E. W. The AbgT family: A novel class of antimetabolite transporters. *Protein Science* (2016). doi:10.1002/pro.2820
34. Du, D. *et al.* Multidrug efflux pumps: structure, function and regulation. *Nature Reviews Microbiology* (2018). doi:10.1038/s41579-018-0048-6
35. Vinothkumar, K. R. & Henderson, R. Structures of membrane proteins. *Q. Rev. Biophys.* (2010). doi:10.1017/S0033583510000041
36. Szakács, G., Paterson, J. K., Ludwig, J. A., Booth-Genthe, C. & Gottesman, M. M. Targeting multidrug resistance in cancer. *Nature Reviews Drug Discovery* (2006). doi:10.1038/nrd1984
37. Liu, Y. H., Di, Y. M., Zhou, Z. W., Mo, S. L. & Zhou, S. F. Multidrug resistance-associated proteins and implications in drug development. in *Clinical and Experimental Pharmacology and Physiology* (2010). doi:10.1111/j.1440-1681.2009.05252.x
38. Choi, Y. H. & Yu, A.-M. ABC transporters in multidrug resistance and pharmacokinetics, and strategies for drug development. *Curr. Pharm. Des.* (2014).
39. Saier, M. H. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res.* (2005). doi:10.1093/nar/gkj001
40. Hvorup, R. N. *et al.* The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. *Eur. J. Biochem.* **270**, 799–813 (2003).
41. Brown, M. H., Paulsen, I. T. & Skurray, R. A. The multidrug efflux protein NorM is a prototype of a new family of transporters. *Mol. Microbiol.* **31**, 394–5 (1999).
42. Paulsen, I. T., Beness, A. M. & Saier, M. H. Computer-based analyses of the protein constituents of transport systems catalysing export of complex carbohydrates in bacteria. *Microbiology* (1997). doi:10.1099/00221287-143-8-2685
43. Helenius, J. *et al.* Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature* (2002). doi:10.1038/415447a

-
44. Rudnick, P. A., Arcondéguy, T., Kennedy, C. K. & Kahn, D. *glnD* and *mviN* are genes of an essential operon in *Sinorhizobium meliloti*. *J. Bacteriol.* (2001). doi:10.1128/JB.183.8.2682-2685.2001
 45. Kim, J.-G. *et al.* Bases of biocontrol: sequence predicts synthesis and mode of action of agrocin 84, the Trojan horse antibiotic that controls crown gall. *Proc. Natl. Acad. Sci. U. S. A.* (2006). doi:10.1073/pnas.0602965103
 46. Ho, A. M., Johnson, M. D. & Kingsley, D. M. Role of the mouse *ank* gene in control of tissue calcification and arthritis. *Science* (80-). (2000). doi:10.1126/science.289.5477.265
 47. Reader, J. S. *et al.* Virology: Major biocontrol of plant tumors targets tRNA synthetase. *Science* (80-). (2005). doi:10.1126/science.1116841
 48. Ling, J. M. ., Moore, R. A., Surette, M. G. & Woods, D. E. The *mviN* homolog in *Burkholderia pseudomallei* is essential for viability and virulence . *Can. J. Microbiol.* (2006). doi:10.1139/w06-042
 49. Zheng, S. *et al.* Structure and mutagenic analysis of the lipid II flippase MurJ from *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* (2018). doi:10.1073/pnas.1802192115
 50. Kuk, A. C. Y., Mashalidis, E. H. & Lee, S.-Y. Crystal structure of the MOP flippase MurJ in an inward-facing conformation. *Nat. Struct. Mol. Biol.* **24**, 171–176 (2017).
 51. Kim, H. J., Minashima, T., McCarthy, E. F., Winkles, J. A. & Kirsch, T. Progressive ankylosis protein (*ank*) in osteoblasts and osteoclasts controls bone formation and bone remodeling. *J. Bone Miner. Res.* (2010). doi:10.1002/jbmr.60
 52. Kuroda, T. & Tsuchiya, T. Multidrug efflux transporters in the MATE family. *Biochim. Biophys. Acta - Proteins Proteomics* **1794**, 763–768 (2009).
 53. Radchenko, M., Symersky, J., Nie, R. & Lu, M. Structural basis for the blockade of MATE multidrug efflux pumps. *Nat. Commun.* **6**, 7995 (2015).
 54. Omote, H., Hiasa, M., Matsumoto, T., Otsuka, M. & Moriyama, Y. The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends Pharmacol. Sci.* (2006). doi:10.1016/j.tips.2006.09.001
 55. Long, F., Rouquette-Loughlin, C., Shafer, W. M. & Yu, E. W. Functional cloning and characterization of the multidrug efflux pumps NorM from *Neisseria gonorrhoeae* and YdhE from *Escherichia coli*. *Antimicrob. Agents Chemother.* (2008). doi:10.1128/AAC.00475-08
 56. Miyamae, S. *et al.* A MATE family multidrug efflux transporter pumps out fluoroquinolones in *Bacteroides thetaiotaomicron*. *Antimicrob. Agents Chemother.*
-

- (2001). doi:10.1128/AAC.45.12.3341-3346.2001
57. Tocci, N. *et al.* Functional analysis of pneumococcal drug efflux pumps associates the MATE DinF transporter with quinolone susceptibility. *Antimicrob. Agents Chemother.* (2013). doi:10.1128/AAC.01298-12
58. Morita, M. *et al.* Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in *Nicotiana tabacum*. *Proc. Natl. Acad. Sci.* (2009). doi:10.1073/pnas.0812512106
59. Remy, E. & Duque, P. Beyond cellular detoxification: A plethora of physiological roles for MDR transporter homologs in plants. *Frontiers in Physiology* (2014). doi:10.3389/fphys.2014.00201
60. Yonezawa, A., Masuda, S., Yokoo, S., Katsura, T. & Inui, K. -i. Cisplatin and Oxaliplatin, but Not Carboplatin and Nedaplatin, Are Substrates for Human Organic Cation Transporters (SLC22A1-3 and Multidrug and Toxin Extrusion Family). *J. Pharmacol. Exp. Ther.* (2006). doi:10.1124/jpet.106.110346
61. Komatsu, T. *et al.* Characterization of the human MATE2 proton-coupled polyspecific organic cation exporter. *Int. J. Biochem. Cell Biol.* (2011). doi:10.1016/j.biocel.2011.03.005
62. Morita, Y., Kataoka, A., Shiota, S., Mizushima, T. & Tsuchiya, T. NorM of *Vibrio parahaemolyticus* is an Na(+)-driven multidrug efflux pump. *J. Bacteriol.* **182**, 6694–7 (2000).
63. He, G.-X. *et al.* An H(+)-coupled multidrug efflux pump, PmpM, a member of the MATE family of transporters, from *Pseudomonas aeruginosa*. *J. Bacteriol.* **186**, 262–5 (2004).
64. Jin, Y., Nair, A. & Van Veen, H. W. Multidrug transport protein NorM from *Vibrio cholerae* simultaneously couples to sodium- and proton-motive force. *J. Biol. Chem.* (2014). doi:10.1074/jbc.M113.546770
65. Ficici, E., Zhou, W., Castellano, S. & Faraldo-Gómez, J. D. Broadly conserved Na⁺-binding site in the N-lobe of prokaryotic multidrug MATE transporters. *Proc. Natl. Acad. Sci.* (2018). doi:10.1073/pnas.1802080115
66. Ogawa, W. *et al.* Characterization of MATE-type multidrug efflux pumps from *Klebsiella pneumoniae* MGH78578. *PLoS One* (2015). doi:10.1371/journal.pone.0121619
67. He, X. *et al.* Structure of a cation-bound multidrug and toxic compound extrusion transporter. *Nature* (2010). doi:10.1038/nature09408

68. Lu, M. *et al.* Structures of a Na⁺-coupled, substrate-bound MATE multidrug transporter. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2099–104 (2013).
69. Radchenko, M., Symersky, J., Nie, R. & Lu, M. Structural basis for the blockade of MATE multidrug efflux pumps. *Nat. Commun.* (2015). doi:10.1038/ncomms8995
70. Kusakizako, T. *et al.* Structural Basis of H⁺-Dependent Conformational Change in a Bacterial MATE Transporter. *Structure* **27**, 293-301.e3 (2019).
71. Lu, M., Radchenko, M., Symersky, J., Nie, R. & Guo, Y. Structural insights into H⁺-coupled multidrug extrusion by a MATE transporter. *Nat. Struct. Mol. Biol.* (2013). doi:10.1038/nsmb.2687
72. Mousa, J. J. *et al.* MATE transport of the E. coli-derived genotoxin colibactin. *Nat. Microbiol.* **1**, 15009 (2016).
73. Tanaka, Y. *et al.* Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. *Nature* **496**, 247–251 (2013).
74. Tanaka, Y., Iwaki, S. & Tsukazaki, T. Crystal Structure of a Plant Multidrug and Toxic Compound Extrusion Family Protein. *Structure* (2017). doi:10.1016/j.str.2017.07.009
75. Tanaka, Y., Iwaki, S. & Tsukazaki, T. Crystal Structure of a Plant Multidrug and Toxic Compound Extrusion Family Protein. *Structure* **25**, 1455-1460.e2 (2017).
76. Miyauchi, H. *et al.* Structural basis for xenobiotic extrusion by eukaryotic MATE transporter. *Nat. Commun.* (2017). doi:10.1038/s41467-017-01541-0
77. Jardetzky, O. Simple allosteric model for membrane pumps [27]. *Nature* (1966). doi:10.1038/211969a0
78. Higgins, C. F. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* (2007). doi:10.1038/nature05630
79. Protein Data Bank. RCSB PDB: Homepage. *RCSB PDB* (2019).
80. He, X. *et al.* Structure of a cation-bound multidrug and toxic compound extrusion transporter. *Nature* **467**, 991–994 (2010).
81. Poole, F. L. *et al.* Defining genes in the genome of the hyperthermophilic archaeon *Pyrococcus furiosus*: Implications for all microbial genomes. *J. Bacteriol.* (2005). doi:10.1128/JB.187.21.7325-7332.2005
82. Fiala, G. & Stetter, K. O. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch. Microbiol.* (1986). doi:10.1007/BF00413027
83. Krahe, M., Antranikian, G. & Märkl, H. Fermentation of extremophilic microorganisms. in *FEMS Microbiology Reviews* (1996). doi:10.1016/0168-6445(96)00018-6

84. Silva, P. J. *et al.* Enzymes of hydrogen metabolism in *Pyrococcus furiosus*. *Eur. J. Biochem.* (2000). doi:10.1046/j.1432-1327.2000.01745.x
85. Sapra, R., Verhagen, M. F. J. M. & Adams, M. W. W. Purification and characterization of a membrane-bound hydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* (2000). doi:10.1128/JB.182.12.3423-3428.2000
86. Bryant, F. O. & Adams, M. W. Characterization of hydrogenase from the hyperthermophilic archaeobacterium, *Pyrococcus furiosus*. *J. Biol. Chem.* (1989).
87. Kietzin, A. & Adams, M. W. W. Tungsten in biological systems. *FEMS Microbiology Reviews* (1996). doi:10.1016/0168-6445(95)00025-9
88. Lundberg, K. S. *et al.* High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* (1991). doi:10.1016/0378-1119(91)90480-Y
89. Waege, I., Schmid, G., Thumann, S., Thomm, M. & Hausner, W. Shuttle vector-based transformation system for *pyrococcus furiosus*. *Appl. Environ. Microbiol.* (2010). doi:10.1128/AEM.01951-09
90. Valentine, D. L. Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* (2007). doi:10.1038/nrmicro1619
91. Koga, Y. Thermal adaptation of the archaeal and bacterial lipid membranes. *Archaea* (2012). doi:10.1155/2012/789652
92. Siliakus, M. F., van der Oost, J. & Kengen, S. W. M. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. *Extremophiles* (2017). doi:10.1007/s00792-017-0939-x
93. Sprott, G. D., Agnew, B. J. & Patel, G. B. Structural features of ether lipids in the archaeobacterial thermophiles *Pyrococcus furiosus* , *Methanopyrus kandleri* , *Methanothermus fervidus* , and *Sulfolobus acidocaldarius* . *Can. J. Microbiol.* (2010). doi:10.1139/m97-066
94. Jain, S., Caforio, A. & Driessen, A. J. M. Biosynthesis of archaeal membrane ether lipids. *Frontiers in Microbiology* (2014). doi:10.3389/fmicb.2014.00641
95. Lobasso, S. *et al.* Coupled TLC and MALDI-TOF/MS Analyses of the Lipid Extract of the Hyperthermophilic Archaeon *Pyrococcus furiosus* . *Archaea* (2012). doi:10.1155/2012/957852
96. Hunte, C. & Michel, H. Crystallisation of membrane proteins mediated by antibody fragments. *Current Opinion in Structural Biology* (2002). doi:10.1016/S0959-440X(02)00354-8
97. Mir, S. H. *et al.* Generation of recombinant antibody fragments for membrane protein

- crystallization. in *Methods in Enzymology* (2015). doi:10.1016/bs.mie.2014.12.029
98. Pardon, E. *et al.* A general protocol for the generation of Nanobodies for structural biology. *Nat. Protoc.* (2014). doi:10.1038/nprot.2014.039
99. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3 resolution. *Nature* (1985). doi:10.1038/318618a0
100. Henderson, R. & Unwin, P. N. T. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* (1975). doi:10.1038/257028a0
101. Pebay-Peyroula, E., Rummel, G., Rosenbusch, J. P. & Landau, E. M. X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science* (80-.). (1997). doi:10.1126/science.277.5332.1676
102. Landau, E. M. & Rosenbusch, J. P. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proc. Natl. Acad. Sci.* (2002). doi:10.1073/pnas.93.25.14532
103. Caffrey, M. Crystallizing membrane proteins for structure-function studies using lipidic mesophases. *NATO Sci. Peace Secur. Ser. A Chem. Biol.* (2013). doi:10.1007/978-94-007-6232-9-4
104. Kolbe, M., Besir, H., Essen, L. O. & Oesterhelt, D. Structure of the light-driven chloride pump halorhodopsin at 1.8 Å Resolution. *Science* (80-.). (2000). doi:10.1126/science.288.5470.1390
105. Cherezov, V. *et al.* In Meso Structure of the Cobalamin Transporter, BtuB, at 1.95 Å Resolution. *J. Mol. Biol.* (2006). doi:10.1016/j.jmb.2006.09.022
106. Kato, H. E. *et al.* Crystal structure of the channelrhodopsin light-gated cation channel. *Nature* (2012). doi:10.1038/nature10870
107. Rasmussen, S. G. F. *et al.* Crystal structure of the β 2 adrenergic receptor-Gs protein complex. *Nature* (2011). doi:10.1038/nature10361
108. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H. W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* (2008). doi:10.1038/nature07063
109. Huang, J., Chen, S., Zhang, J. J. & Huang, X. Y. Crystal structure of oligomeric β 1-adrenergic G protein-coupled receptors in ligand-free basal state. *Nat. Struct. Mol. Biol.* (2013). doi:10.1038/nsmb.2504
110. Doré, A. S. *et al.* Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature* (2014). doi:10.1038/nature13396

111. <https://blanco.biomol.uci.edu/mpstruc>.
112. Wüthrich, K. Protein structure determination in solution by NMR spectroscopy. *Journal of Biological Chemistry* (1990).
113. Dubochet, J., Frank, J. & Henderson, R. THE NOBEL PRIZE IN CHEMISTRY 2017 They captured life in atomic detail. *NOBEL* (2017).
114. Ghosh, E., Kumari, P., Jaiman, D. & Shukla, A. K. Methodological advances: The unsung heroes of the GPCR structural revolution. *Nature Reviews Molecular Cell Biology* (2015). doi:10.1038/nrm3933
115. Higgins, C. F. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* **446**, 749–757 (2007).
116. Lu, M. Structures of multidrug and toxic compound extrusion transporters and their mechanistic implications. *Channels* (2016). doi:10.1080/19336950.2015.1106654
117. *Extremophiles Handbook*. (2011). doi:10.1007/978-4-431-53898-1
118. Matyash, V., Liebisch, G., Kurzchalia, T. V., Shevchenko, A. & Schwudke, D. Lipid extraction by methyl- tert -butyl ether for high-throughput lipidomics . *J. Lipid Res.* (2008). doi:10.1194/jlr.d700041-jlr200
119. Hartler, J. *et al.* Lipid data analyzer: Unattended identification and quantitation of lipids in LC-MS data. *Bioinformatics* (2011). doi:10.1093/bioinformatics/btq699
120. Li, D. & Caffrey, M. Lipid cubic phase as a membrane mimetic for integral membrane protein enzymes. *Proc. Natl. Acad. Sci.* (2011). doi:10.1073/pnas.1101815108
121. Kabsch, W. *XDS. Acta Crystallogr. Sect. D Biol. Crystallogr.* (2010). doi:10.1107/S0907444909047337
122. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2004). doi:10.1107/S0907444904019158
123. Adams, P. D. *et al.* PHENIX: Building new software for automated crystallographic structure determination. in *Acta Crystallographica Section D: Biological Crystallography* (2002). doi:10.1107/S0907444902016657
124. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2010). doi:10.1107/S0907444909052925
125. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. in *Acta Crystallographica Section D: Biological Crystallography* (2006). doi:10.1107/S0907444906045975
126. Read, R. J. Pushing the boundaries of molecular replacement with maximum likelihood.

- Acta Crystallogr. - Sect. D Biol. Crystallogr.* (2001). doi:10.1107/S0907444901012471
127. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* (2007). doi:10.1107/s0021889807021206
128. Liebschner, D. *et al.* Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. Sect. D Struct. Biol.* (2017). doi:10.1107/s2059798316018210
129. Terwilliger, T. C. *et al.* Iterative-build OMIT maps: Map improvement by iterative model building and refinement without model bias. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2008). doi:10.1107/S0907444908004319
130. Afonine, P. V. *et al.* FEM: Feature-enhanced map. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2015). doi:10.1107/S1399004714028132
131. Chen, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2010). doi:10.1107/S0907444909042073
132. Zhang, Y. & Skolnick, J. TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* (2005). doi:10.1093/nar/gki524
133. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* (2008). doi:10.1002/jcc.20945
134. Alexov, E. G. & Gunner, M. R. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophys. J.* (1997). doi:10.1016/S0006-3495(97)78851-9
135. Georgescu, R. E., Alexov, E. G. & Gunner, M. R. Combining conformational flexibility and continuum electrostatics for calculating pK_as in proteins. *Biophys. J.* (2002). doi:10.1016/S0006-3495(02)73940-4
136. Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* (2012). doi:10.1021/ct300400x
137. Klauda, J. B. *et al.* Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* (2010). doi:10.1021/jp101759q
138. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* (1996). doi:10.1016/0263-7855(96)00018-5
139. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* (2015). doi:10.1016/j.softx.2015.06.001
140. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for

- Ewald sums in large systems. *J. Chem. Phys.* (1993). doi:10.1063/1.464397
141. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for molecular simulations. *J. Comput. Chem.* (1997). doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H
142. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* (1984). doi:10.1063/1.448118
143. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* (1985). doi:10.1103/PhysRevA.31.1695
144. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* (1981). doi:10.1063/1.328693
145. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* (2014). doi:10.1016/j.cpc.2013.09.018
146. <https://www.nanotempertech.com>.
147. Gault, J. *et al.* High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat. Methods* (2016). doi:10.1038/nmeth.3771
148. Ma, P. *et al.* The cubicon method for concentrating membrane proteins in the cubic mesophase. *Nat. Protoc.* (2017). doi:10.1038/nprot.2017.057
149. Wiegand, I., Hilpert, K. & Hancock, R. E. W. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* (2008). doi:10.1038/nprot.2007.521
150. Daniel, A. S., Fuller-Pace, F. V., Legge, D. M. & Murray, N. E. Distribution and diversity of hsd genes in *Escherichia coli* and other enteric bacteria. *J. Bacteriol.* (1988). doi:10.1128/jb.170.4.1775-1782.1988
151. Weinberg, M. V., Schut, G. J., Brehm, S., Datta, S. & Adams, M. W. W. Cold shock of a hyperthermophilic archaeon: *Pyrococcus furiosus* exhibits multiple responses to a suboptimal growth temperature with a key role for membrane-bound glycoproteins. *J. Bacteriol.* (2005). doi:10.1128/JB.187.1.336-348.2005
152. Surade, S. Comparative analysis and 'expression space' coverage of the production of prokaryotic membrane proteins for structural genomics. *Protein Sci.* (2006). doi:10.1110/ps.062312706
153. Nie, L. *et al.* Identification of the high-affinity substrate-binding site of the multidrug and toxic compound extrusion (MATE) family transporter from *Pseudomonas stutzeri*.

- J. Biol. Chem.* (2016). doi:10.1074/jbc.M116.728618
154. Srinivasan, L., Baars, T. L., Fendler, K. & Michel, H. Functional characterization of solute carrier (SLC) 26/sulfate permease (SulP) proteins in membrane mimetic systems. *Biochim. Biophys. Acta - Biomembr.* (2016). doi:10.1016/j.bbamem.2016.01.006
155. Kaur, J., Olkhova, E., Malviya, V. N., Grell, E. & Michel, H. A L-lysine transporter of high stereoselectivity of the amino acid-polyamine-organocation (apc) superfamily production, functional characterization, and structure modeling. *J. Biol. Chem.* (2014). doi:10.1074/jbc.M113.510743
156. Jin, X. *et al.* Insights into conformational regulation of PfMATE transporter from *Pyrococcus furiosus* induced by alternating protonation state of Asp41 residue: A molecular dynamics simulation study. *Biochim. Biophys. Acta - Gen. Subj.* (2016). doi:10.1016/j.bbagen.2016.02.007
157. Lamb, A. L., Kappock, T. J. & Silvaggi, N. R. You are lost without a map: Navigating the sea of protein structures. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2015). doi:10.1016/j.bbapap.2014.12.021
158. von Heijne, G. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* (1992). doi:10.1016/0022-2836(92)90934-C
159. Heck, A. J. R. Native mass spectrometry: A bridge between interactomics and structural biology. *Nat. Methods* (2008). doi:10.1038/nmeth.1265
160. Leney, A. C. & Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* (2017). doi:10.1007/s13361-016-1545-3
161. Gupta, K. *et al.* Identifying key membrane protein lipid interactions using mass spectrometry. *Nat. Protoc.* (2018). doi:10.1038/nprot.2018.014
162. Bolla, J. R. *et al.* Direct observation of the influence of cardiolipin and antibiotics on lipid II binding to MurJ. *Nat. Chem.* (2018). doi:10.1038/nchem.2919
163. Yen, M. R., Chen, J. S., Marquez, J. L., Sun, E. I. & Saier, M. H. Multidrug resistance: phylogenetic characterization of superfamilies of secondary carriers that include drug exporters. *Methods Mol. Biol.* (2010). doi:10.1007/978-1-60761-700-6_3
164. Kuk, A. C. Y., Mashalidis, E. H. & Lee, S. Y. Crystal structure of the MOP flippase MurJ in an inward-facing conformation. *Nat. Struct. Mol. Biol.* (2017). doi:10.1038/nsmb.3346
165. Kwon, S. K., Kim, S. K., Lee, D. H. & Kim, J. F. Comparative genomics and experimental evolution of *Escherichia coli* BL21(DE3) strains reveal the landscape of toxicity escape from membrane protein overproduction. *Sci. Rep.* (2015).

- doi:10.1038/srep16076
166. Kühlbrandt, W. The resolution revolution. *Science* (2014). doi:10.1126/science.1251652
167. Cheng, Y. Single-particle cryo-EM-How did it get here and where will it go. *Science* (2018). doi:10.1126/science.aat4346
168. Khoshouei, M., Radjainia, M., Baumeister, W. & Danev, R. Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. *Nat. Commun.* (2017). doi:10.1038/ncomms16099
169. Lavery, D. *et al.* Cryo-EM structure of the human $\alpha 1\beta 3\gamma 2$ GABAA receptor in a lipid bilayer. *Nature* (2019). doi:10.1038/s41586-018-0833-4
170. Martens, C. *et al.* Direct protein-lipid interactions shape the conformational landscape of secondary transporters. *Nat. Commun.* (2018). doi:10.1038/s41467-018-06704-1
171. Wagner, A., Duman, R. & Mykhaylyk, V. The Long-Wavelength MX Beamline I23 at Diamond Light Source. *Acta Crystallogr. Sect. A Found. Adv.* (2015). doi:10.1107/s2053273314093942

8. APPENDIX



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 25, 2018 – 11:06 AM BST

PDB ID : 6GWH
Title : Outward-facing conformation of a MOP transporter.
Deposited on : 2018-06-25
Resolution : 2.80 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

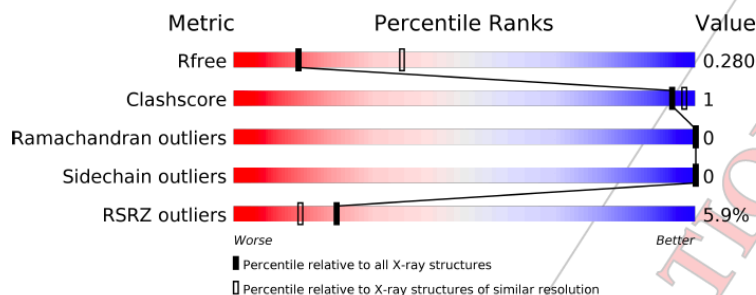
MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : rb-20031172
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031172

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	461	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 3311 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

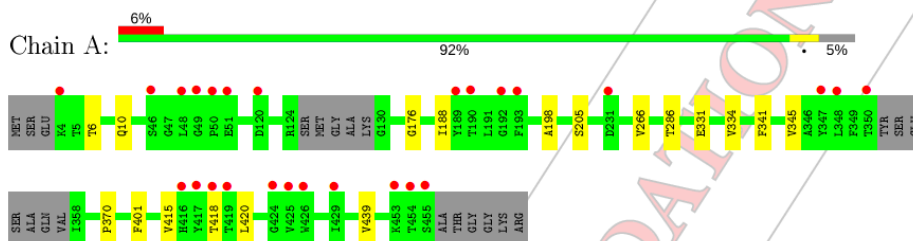
- Molecule 1 is a protein called MOP transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	440	3311	2176	539	576	20	0	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: MOP transporter



4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	95.83Å 95.83Å 369.11Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.97 - 2.80 19.97 - 2.80	Depositor EDS
% Data completeness (in resolution range)	97.7 (19.97-2.80) 97.7 (19.97-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.05 (at 2.79Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.258 , 0.280 0.258 , 0.280	Depositor DCC
R_{free} test set	809 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	80.7	Xtriage
Anisotropy	0.258	Xtriage
Bulk solvent $k_{sol}(e/\text{Å}^3)$, $B_{sol}(\text{Å}^2)$	0.32 , 51.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3311	wwPDB-VP
Average B, all atoms (Å ²)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.24	0/3369	0.38	0/4568

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3311	0	3497	9	0
All	All	3311	0	3497	9	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (9) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:331:GLU:HA	1:A:334:VAL:HG22	1.89	0.53
1:A:176:GLY:O	1:A:205:SER:OG	2.28	0.50
1:A:286:THR:HG22	1:A:370:PRO:HG3	1.97	0.46
1:A:188:ILE:HD11	1:A:198:ALA:HB2	1.97	0.46
1:A:418:THR:HG23	1:A:420:LEU:HB2	1.98	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:266:VAL:HG13	1:A:415:VAL:HG21	1.98	0.45
1:A:401:PHE:CE2	1:A:439:VAL:HG11	2.54	0.42
1:A:6:THR:O	1:A:10:GLN:HG2	2.20	0.41
1:A:341:PHE:O	1:A:345:VAL:HG23	2.22	0.40

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	434/461 (94%)	427 (98%)	7 (2%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	347/362 (96%)	347 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	407	GLN
1	A	416	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

CONFIDENTIAL VALIDATION REPORT

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	440/461 (95%)	0.50	26 (5%) 22 14	54, 73, 103, 130	6 (1%)

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	50	PRO	6.6
1	A	417	TYR	4.6
1	A	46	SER	4.6
1	A	231	ASP	4.5
1	A	347	TYR	4.2
1	A	192	GLY	4.1
1	A	4	LYS	3.8
1	A	419	THR	3.5
1	A	350	THR	3.5
1	A	418	THR	3.3
1	A	193	PHE	3.2
1	A	51	GLU	2.9
1	A	455	SER	2.9
1	A	49	GLY	2.8
1	A	189	TYR	2.7
1	A	48	LEU	2.6
1	A	453	LYS	2.5
1	A	454	THR	2.5
1	A	429	ILE	2.3
1	A	348	LEU	2.3
1	A	416	HIS	2.2
1	A	425	VAL	2.1
1	A	426	TRP	2.1
1	A	424	GLY	2.1
1	A	190	THR	2.1
1	A	120	ASP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.

CONFIDENTIAL VALIDATION REPORT



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2018 – 02:50 PM GMT

PDB ID : 6FHZ
Title : Inward-facing conformation of a MOP transporter.
Deposited on : 2018-01-16
Resolution : 2.80 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

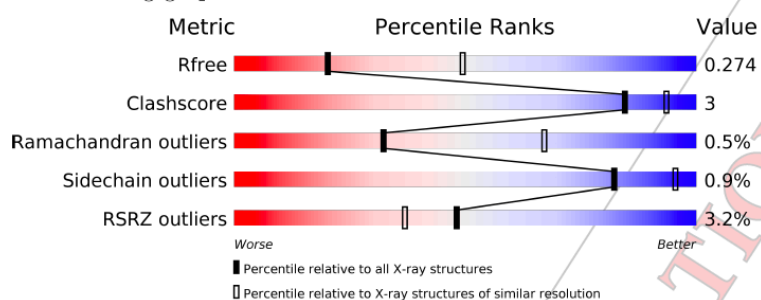
MolProbity : 4.02b-467
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030736
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030736

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	440	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 3264 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

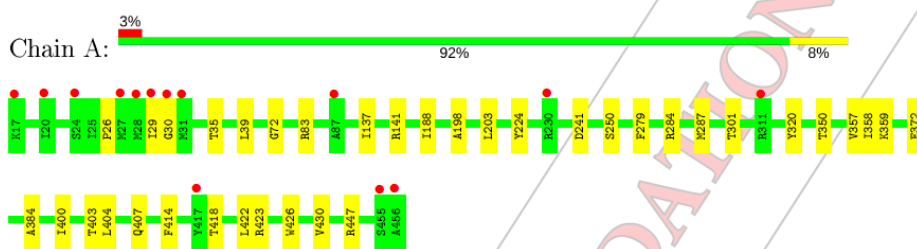
- Molecule 1 is a protein called Putative MOP flippase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	440	3264	2139	530	577	18	0	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Putative MOP flippase



CONFIDENTIAL VALIDATION REPORT

4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	107.12Å 134.31Å 69.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.80 - 2.80 19.80 - 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.80-2.80) 99.8 (19.80-2.80)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.05 (at 2.79Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.229 , 0.274 0.227 , 0.274	Depositor DCC
R_{free} test set	612 reflections (4.83%)	DCC
Wilson B-factor (Å ²)	72.9	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent $k_{sol}(e/\text{Å}^3)$, $B_{sol}(\text{Å}^2)$	0.34 , 47.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3264	wwPDB-VP
Average B, all atoms (Å ²)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.24% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.24	0/3323	0.39	1/4515 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	26	PRO	N-CA-CB	6.04	110.55	103.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3264	0	3371	18	0
All	All	3264	0	3371	18	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (18) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:414:PHE:O	1:A:418:THR:OG1	2.28	0.50
1:A:188:ILE:HD11	1:A:198:ALA:HB2	1.96	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:72:GLY:HA3	1:A:250:SER:HB3	1.95	0.47
1:A:423:ARG:HA	1:A:426:TRP:CD1	2.49	0.46
1:A:372:PHE:HB3	1:A:430:VAL:HA	1.96	0.46
1:A:39:LEU:HD13	1:A:284:ARG:HB3	1.99	0.44
1:A:357:VAL:HG13	1:A:358:ILE:HG23	1.99	0.44
1:A:35:THR:HG21	1:A:284:ARG:NH1	2.34	0.43
1:A:400:ILE:HG23	1:A:404:LEU:HD12	2.01	0.43
1:A:279:PHE:HD1	1:A:426:TRP:HZ3	1.67	0.42
1:A:350:THR:HG21	1:A:359:LYS:HA	2.02	0.42
1:A:83:ARG:HH21	1:A:241:ASP:HB3	1.86	0.41
1:A:447:ARG:HA	1:A:447:ARG:HD2	1.81	0.41
1:A:422:LEU:HA	1:A:422:LEU:HD23	1.96	0.41
1:A:301:THR:HG22	1:A:384:ALA:O	2.21	0.41
1:A:403:THR:O	1:A:407:GLN:HB2	2.21	0.41
1:A:137:ILE:O	1:A:141:ARG:HG3	2.21	0.41
1:A:284:ARG:NH2	1:A:287:MET:HG3	2.37	0.40

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	438/440 (100%)	423 (97%)	13 (3%)	2 (0%)	32 67

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	29	ILE
1	A	30	GLY

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	332/345 (96%)	329 (99%)	3 (1%)	82 95

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	203	LEU
1	A	224	TYR
1	A	320	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	99	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

CONFIDENTIAL VALIDATION REPORT

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	440/440 (100%)	0.12	14 (3%) 48 37	47, 61, 93, 132	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	456	ALA	5.0
1	A	31	MET	4.6
1	A	20	ILE	4.3
1	A	28	MET	3.4
1	A	24	SER	3.3
1	A	27	MET	3.2
1	A	29	ILE	3.0
1	A	30	GLY	2.9
1	A	230	ARG	2.7
1	A	455	SER	2.6
1	A	311	ARG	2.5
1	A	87	ALA	2.2
1	A	417	TYR	2.1
1	A	17	LYS	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.

CONFIDENTIAL VALIDATION REPORT

ACKNOWLEDGMENTS

Firstly, I would like to express my gratitude to my supervisor Prof. Dr. Dr. h.c. Hartmut Michel for assigning me with this daring project and providing the opportunity to learn about membrane proteins and structural biology. I am grateful for his support during the time of my PhD studies, for giving me trust and independence in order to become a scientist I am now. I truly appreciate his proof-reading of this thesis and my manuscript.

Besides my advisor, I would like to thank my thesis committee member and my university supervisor Prof. Dr. Klaas Martinus Pos for insightful discussions.

My sincere acknowledgements are extended to Conny Münke who helped me immensely at the beginning of my studies, introduced me into the membrane protein world and answered millions of my questions. I am grateful for her patience, kindness and for sharing her tremendous knowledge and experience with me.

In particular, I am also grateful to Barbara Rathmann. We worked together on crystallization of my target protein and I very much appreciate her enthusiasm, willingness to set up multiple plates and her amazing skills to harvest LCP crystals. I am deeply indebted to her for contributing to the success of my project.

Many thanks also go to Dr. Hao Xie for all the help in the lab, critical and reflective thoughts and his constant support.

I would like to express my gratitude to all collaborators who contributed to different parts of this work for the fruitful cooperation:

- PD Dr. Winfried Hausner for making my stay at the University of Regensburg possible, for access to the laboratory and fermentation facility. Renate Richau for excellent

technical help and teaching me how to deal with *Pyrococcus furiosus* cells. Dipl.-Ing. Thomas Hader and Konrad Eichinger for their assistance with bioreactors and cultivation of the archaeal cells. Without their precious support it would not be possible to conduct this research.

- Dr. Ahmad Reza Mehdipour for performing MD simulations and valuable discussions.
- PD Dr. Harald Köfeler, Dipl.-Ing. Dr. Martin Trötz Müller and Christine Pein for performing MS-based lipidomics.
- Dr. Jani Reddy Bolla for performing native MS analysis and his prompt feedback.
- Dr. Ramona Duman for measuring the crystals at the high wavelength beamline I23.

Many thanks to all co-authors of our publication, also to other past and present group members that I have had the pleasure to work with or alongside.

I also thank my beloved friends for sharing not only cheerful, but also tough moments, for the stimulating discussions, long days and nights working together, for our memorable dinners, trips and all the fun we have had in the last years. I am grateful to all of you for making my PhD life more enjoyable.

Last but definitely not least, I wish to express my gratitude to my wonderful family: my dearest mom (I am so grateful for everything you have done for me), my sister Nataalka, and my brother Kamilek. Without their strength and belief in me, it would be harder to accomplish my PhD studies. I am beyond grateful for their endless support along the rocky way, for all their love and encouragement. To you I dedicate this thesis.

CURRICULUM VITAE

SANDRA ZAKRZEWSKA



Address	Altenhöferallee 70/1.03.14 60438 Frankfurt am Main
Date of birth	April 23 rd 1989
Place of birth	Eschweiler
Nationality	Polish
E-mail	sanzak89@gmail.com

EDUCATION

From 05/2014

Ph.D. candidate

Max Planck Institute of Biophysics, Frankfurt am Main, Germany
Department of Molecular Membrane Biology

Supervisor: Prof. Dr. Dr. h.c. Hartmut Michel

Dissertation: Structural and functional characterization of a MATE family multidrug resistance transporter from *Pyrococcus furiosus*

2/2012 - 9/2013

Master of Science Degree in Molecular Biotechnology

Gdansk University of Technology, Gdansk, Poland
Department of Molecular Biotechnology and Microbiology

Supervisors: Prof. Dr. Józef Kur and Dr. Marcin Olszewski

Thesis: Heterologous production and purification of a DNA polymerase from the thermophilic bacterium *Thermovibrio ammonificans*

10/2008 - 2/2012

Engineer's Degree in Molecular Biotechnology

Gdansk University of Technology, Gdansk, Poland

Department of Molecular Biotechnology and Microbiology

Supervisors: Prof. Dr. Józef Kur and Dr. Marcin Olszewski

Thesis: Bioinformatic studies on the genes encoding family B-DNA polymerases from *Euryarchaeota*

PROFESSIONAL EXPERIENCE

From 05/2014

Max Planck Institute of Biophysics, Frankfurt am Main

Designed and performed research on a structural and functional characterization of the MATE (multidrug and toxic compound extrusion) family transporter

Collaborated with the lab teams to develop experimental and computational approach on protein and lipid interplay

Organized and coordinated the 1st joint PhD Student Retreat of the Max Planck Institute of Biophysics and International Max Planck Research School (IMPRS), Hamburg, Germany, June 2018

Participated in the workshops:

In Situ Serial Crystallography Workshop, Paul Scherrer Institute, Villigen, Switzerland

6th International School on Biological Crystallization, Granada, Spain

Project Management in Biotech Industries, Frankfurt am Main, Germany

Advanced training seminar in accordance with §15 GenTSV for project managers and assigned persons for biological security, Max Planck Institute of Biochemistry, Martinsried, Germany

Attended the conferences and symposiums:

Gordon Research Conference on Ligand recognition and Molecular Gating, Ventura, USA

CRC 807 International Symposium on New Horizons in Membrane Transport and Communication, Frankfurt am Main, Germany

EUFEPS Annual Meeting 2019, Goethe University, Frankfurt am Main, Germany

Barcelona BioMed Conference on Transporters and other Molecular Machines, Institute for Research in Biomedicine, Barcelona, Spain

10/2008 – 9/2013. **Gdansk University of Technology, Gdansk**

Heterologously produced and purified a DNA polymerase from the thermophilic bacterium

Conducted bioinformatic studies on the genes encoding DNA polymerases

Internship at the RWTH Aachen, ISA Institute of Environmental Engineering (Prof. Dr. rer. nat. habil. Horst Friedrich Schröder), Aachen, Germany

Attended the training:

Internal Auditor of Management System in Laboratory PN-EN ISO/IEC 17025:2005, Gdansk, Poland

Attended the conferences:

14th National Academic Seminar of Biotechnology Students and

4th International Students' Conference of Biotechnology, Gdansk, Poland

AWARDS

2018 **Best Poster Award - 1st place in the poster session**

Poster: Molecular structures of two conformational states of an MOP superfamily transporter

Gordon Research Conference on Ligand recognition and Molecular gating, USA

2012 **2nd place in the poster session**

Poster: Novel thermostable family B DNA polymerases from *Euryarchaeota*

7th International Conference of Young Naturalists “From Biotechnology to Environment Protection”, Poland

LANGUAGES

Polish (native), English (fluent), German (intermediate)