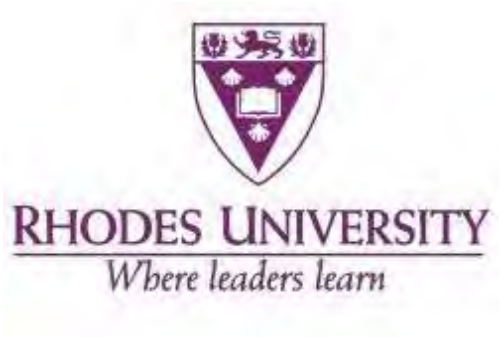


Towards a possible future solution against Multidrug Resistance: An in silico exploration of the Multidrug and Toxic compound Extrusion (MATE) transporter proteins as potential antimicrobial drug targets.

A thesis submitted in partial fulfilment of the requirements for the degree

of

Master of Science in Bioinformatics and Computational Molecular Biology (Coursework and Thesis)

RHODES UNIVERSITY, SOUTH AFRICA
Research Unit in Bioinformatics (RUBi)

DEPARTMENT OF BIOCHEMISTRY AND MICROBIOLOGY
Faculty of science

by

Amira Damji

G12D0762

2023

ABSTRACT

The rise of multidrug resistance (MDR) has become a pressing global issue, hindering the treatment of cancers and infectious diseases, and imposing a burden on healthcare systems and the economy. The Multidrug and Toxic compound Extrusion (MATE) superfamily of membrane efflux transporters is one of the key players contributing to MDR due to their ability to export a wide range of cationic and hydrophilic xenobiotics, including treatment drugs, from cells, diminishing their efficacy. Targeting MATE transporters holds great promise in achieving some cellular control over MDR, but first, a deeper understanding of their structure-function-dynamics link is required. This study aimed to explore the MATE transporters as potential antimicrobial drug targets using a two-fold in silico approach. First, virtual screening of compounds from the South African Natural Compounds Database (SANCDB) was performed to identify prospective lead inhibitory compounds against the MATE transporters using molecular docking, and top hits were selected based on their binding energy and interaction with the active site on the N-lobe of the protein. Second, to investigate the molecular-level dynamics of their extrusion mechanism, the MATE transporter structures were embedded in a POPC membrane bilayer using the CHARMM-GUI online tool and then subjected to MD simulations for 100 ns with the CHARMM 36m force field using GROMACS. The resulting trajectories were evaluated using three standard metrics – RMSD, RMSF, and Rg; significant global structural changes were observed and key functional regions in both membrane- and non-membrane transmembrane (TM) segments were identified, containing more dynamic and flexible residues than other regions. Furthermore, the MATE transporters showed more of a loosely-packed structure, providing flexibility to allow for conformational switching during their substrate-transport cycle, which is typical for proteins whose secondary structures are composed of all α-helices. The scope of this study lied in the preliminary stages of the computer-aided drug design process, and provided insights that can be used to guide the development of strategies aimed at regulating or inhibiting the function of the MATE transporters, offering a possible future solution to the growing challenge of MDR.

DECLARATION

I, Amira Damji, hereby declare that this thesis submitted to Rhodes University is my original work and has never been submitted to any institution for a degree or diploma. All sources, references and literature used during preparation of this work are properly cited and listed in complete reference to the due source.

---------- ----------

Signature Date

DEDICATION

This thesis is dedicated to my parents,

Mahamood and Fatima Damji,

You have always been my constant source of encouragement, love, and strength. Thank you for your unwavering support and belief in my potential – especially through the rough moments, when even I stopped believing in myself. Without you, I would not be where I am today, or have had the opportunities that I have in life. I am grateful to have you as companions through this journey of life.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Prof. Kevin Lobb and, co-supervisor Dr. Vuyani Moses, as well as the Research Unit in Bioinformatics (RUBi) at Rhodes University for the opportunity to work on this study and for their input and guidance during the course of this study.

Prof. Lobb, thank you for your patience with me, your help and support where I needed it, and for the feedback – I am extremely grateful for this.

I would also like to acknowledge the members of the Research Unit in Bioinformatics (RUBi), including some my fellow MSc colleagues, especially Sophakama Zabo, and the PhD students, who took the time to help/advise me when I needed it. Thank you.

I am also grateful to the Dean of Sciences at Rhodes University, Prof. Tony Booth, for his support through the challenging times I faced during the course of my academic career. I truly appreciate it.

In addition, I would like to acknowledge my friend, Dane Brown, for all his help and support and input where I needed it. Thank you for your advice and guidance and for the productive work space.

I would like to extend a huge thank you to my life partner in crime, Ibi. I am grateful beyond words for all your unwavering support and belief in me. This achievement would not have been possible without you! You have been there for my best and worst moments and are a huge part of this achievement. Thank you from the bottom of my heart.

To my brother Aman, I am grateful for all your support, encouragement, and belief in me. It has pushed me through and kept me going throughout various challenges I have experienced during my life and academic journey. I have learned so much from you and feel blessed to have you as my brother. Thank you!

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

HM – Homology modeling (or homology model, context dependent)

PRIMO – PRotein Interactive MOdeling

MDR – Multidrug resistance

MRP – Multidrug-resistant protein

MP – membrane protein

IMP – integral membrane protein

MOP – multidrug/oligosaccharidyl-lipid/polysaccharide

Families of bacterial efflux-pump proteins:

– MATE – Multidrug And Toxic-compound Extrusion

– ABC – ATP-Binding Cassette superfamily

– MFS – Major Facilitator Superfamily (MFS)

– SMR – Small Multidrug Resistance superfamily

– RND – Resistance Nodulation Division family

– PACE – Proteobacterial Antimicrobial Compound Efflux

SCL – Solute carriers

CADD – Computer-aided Drug Design

FBDD – Fragment-based Drug Design

SBDD – Structure-based Drug Design

HTS – High-throughput screening

MTP – membrane transporter protein

MDR – Multidrug resistance

Pf – *Pyrococcus furiosus*

MATE transporter subfamiles:

– DinF – DNA damage-inducible protein F

– eMATE – eykaryotic MATE

– NorM_VC – NorM from Vibrio cholerae

CHARMM – Chemistry at Harvard Macromolecular Mechanic

RCSB – Research Collaboratory for Structural Bioinformatics

MD – Molecular Dynamics

MDS – Molecular dynamics simulation

NRF – Norfloxacin

EP – Efflux pump

EPI – Efflux pump inhibitor

NP – Natural product

SANCDB – South African Natural Compounds Database

POPC – 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

PDB – Protein Data Bank (file format)

IF – Inward-facing

IFC – Inward-facing conformation

OF – Outward-facing

OFC – Outward-facing conformation

TM – Transmembrane helix

TMD – Transmembrane Domain

MW – Molecular weight

RCSB – Research Collaboratory for Structural Bioinformatics (website)

RMSD – Root Mean Square Deviation

RMSF – Root Mean Square Fluctuation

Rg – Radius of gyration

PDBTM – Protein Data Bank of Transmembrane Proteins

Towards a possible future solution against MDR: An in silico exploration of the MATE transporter proteins as potential antimicrobial drug targets.

CHAPTER 1: BACKGROUND AND MOTIVATION

1.1 | MDR and MATE transporters

Multidrug resistance (MDR) impedes drug treatments worldwide, causing a myriad of deaths from various diseases. As the name suggests, MDR refers to the phenomenon of simultaneous resistance to multiple drugs with heterogeneous properties (Larsen *et al.,* 2000; Tanwar *et al.*, 2014; Lopes-Rodrigues *et al.,* 2017) and has become a growing global threat. Agencies, including the Center for Disease Control and Prevention (CDC) and the World Health Organization (WHO), have declared MDR a serious and growing cause of concern. In effect, the WHO has ranked MDR (also referred to as antimicrobial resistance (AMR)) as one of the leading global health issues (Ghosh *et al.,* 2021), with a predicted rise in the number of MDR-related deaths by 10 million by the year 2050, at an estimated economical cost of 100 trillion dollars, if left unchecked (O'Neill, 2014). Furthermore, O'Neill (2014) provided a breakdown of the predicted increase in regional deaths by 2050 for Asia (4,730,000 deaths), Africa (4,150,000 deaths), Europe (390,000 deaths), Latin America (392,000 deaths), North America (317,000 deaths), and Oceania (22,000 deaths), highlighting the urgency of drawing up prospective strategies to dealing with the problem of MDR.

MDR is a natural phenomenon occurring not only in a multitude of (most concerningly pathogenic) bacteria and viruses, but is also a means by which cancer cells evade chemotherapeutic drug treatments – owing to the expression of drug efflux pumps (Wu *et al*., 2011; Kunjan *et al*., 2013). Implicated as a major cause of MDR, membrane efflux transporters are deemed significant drug targets to counteract (the effects of) MDR, thereby allowing for efficacious drug treatments (Tanwar *et al.*, 2014; Lopes-Rodrigues *et al.,* 2017). However, for effective drug targeting, gaining insight into the intricate particulars involved in the functional mechanism of action of these drug targets is crucial (Tanwar *et al.*, 2014). This research explores the family of efflux transporters known as the multidrug and toxic compound extrusion (MATE) proteins – found ubiquitously across most life forms – as potential antimicrobial drug targets in the prevention of MDR. Like other families of efflux proteins, MATE transporters are known to be promiscuous based on their capacity to export structurally and chemically

diverse (drug) substrates of differing targets out of the cell – resulting in the phenomenon of MDR (Kusakizako *et al.*, 2020).

The MATE transporter family confers resistance to a diverse range of therapeutic agents, hindering drug treatments. For example, by causing resistance to hydrophilic fluoroquinolone drugs, they hinder treatments against *Neisseria gonorrhoeae* bacterial infections (McAleese et al., 2005; Kaatz *et al.*, 2005; Kuroda and Tsuchiya, 2009; Du *et al*., 2018), and by causing resistance to erythromycin and chloramphenicol, they hinder treatments against *Streptococcus pneumoniae* infections (Schindler and Kaatz, 2016; Du *et al*., 2018). In addition, they also induce resistance to aminoglycoside antibiotics, as well as anticancer drug treatments, making them promising targets to combat MDR mechanisms that hinder drug treatments on a global scale (McAleese et al., 2005; Kaatz *et al.*, 2005; Kuroda and Tsuchiya, 2009; Claxton et al., 2018).

Being polyspecific – and able to extrude drugs of a diverse nature – MATE transporters in humans and bacteria impede drug treatments such as in the case of cancers, bacterial and viral infections, and even diabetes, highlighting their importance as drug targets to prevent MDR (Omote *et al.,* 2006; Kuroda and Tsuchiya, 2009). Additionally, resolving the molecular-level mechanism underlying the ability of MATE transporters to identify and expel polyspecific drugs will provide a wider knowledge base on MDR in the broader sense (Higgins, 2007).

Spanning the cell membrane, MATE transporters are known to switch between an inward-facing conformation (IFC) – forming a Λ-shaped architecture towards the cytoplasm – and an outward-facing conformation (OFC) – forming a V-shaped architecture towards the extracellular environment, driving (cellular) substrate extrusion (Symersky *et al.*, 2015; Jin *et al.*, 2016; Claxton *et al.*, 2018). Based on this knowledge, they are known to fall under the "alternating-access" model known as the "rocker-switch" mechanism – shared with other families of efflux proteins (Law *et al.*, 2008; Du *et al.*, 2015; Drew and Boudker, 2016; Drew *et al.*, 2021), which will be discussed in detail the next chapter. However, their exact mechanism of action remains elusive (Jin *et al.,* 2016; Eisinger *et al.,* 2018; Kusakizako *et al.*, 2020).

Advances in computational methods have taken research capabilities to a higher level, offering an economical, time-efficient means to study drug targets. This study employs the two commonly used in-silico techniques in the drug discovery and design/development process – molecular docking followed by molecular dynamics simulations – described in the approach section below that, when combined, meet the aim of exploring the MATE transporters as antimicrobial drug targets.

1.2 | Problem statement

Increasing at an alarming pace, outrunning the discovery and design/development of new drugs, MDR remains a hindrance to drug therapies, taking a toll on the healthcare system and the economy worldwide. This form of cellular resistance has been traced back to various drug efflux proteins, such as the family of MATE transporter proteins, known to exhibit a conserved structure through evolution in all kingdoms of life. These proteins are one of the most important transport proteins involved in exporting toxic compounds and drugs from the cells via a process known as 'extrusion'. The molecular-level mechanism by which extrusion occurs – conferring MDR – is yet to be understood. As such, a better understanding of the way in which they function would open the possibility of targeting them to prevent their efflux action, offering a means to combat MDR and allowing cells to reach sufficient levels of drug dosage, which would otherwise be compromised.

1.3 | Aim

This study aims to explore the transmembrane MATE transporter proteins as potential drug targets using a bioinformatics approach, offering a possible future solution to the problem of MDR. The aim will be addressed via a two-fold in-silico approach, employing the two commonly used CADD techniques – molecular docking and molecular dynamics – to achieve the objectives below.

1.4 | Objectives

1. To identify potentially optimal inhibitory compounds from the established South African natural compounds database (SANCDB) against both conformations of the MATE transporter using molecular docking, following which, the results will be filtered (based on binding at the correct site as well as lowest binding energies) to select the ideal top docked IFC and OFC complexes.

2. To generate membrane systems of the *apo*-proteins and docked protein-ligand complexes (selected above) - which serve to mimic the native (membrane) environments of the transmembrane MATE transporters for the following simulations - by embedding them into homogenous POPC membrane bilayers. In addition, the available OFC MATE complex structure with a derivative of the known drug, norfloxacin, will be used as a control, and this complex also embedded into a POPC lipid membrane bilayer for the following objective.

3. To investigate the dynamic behaviour of the protein within the above-generated membrane systems over time via molecular dynamics simulations (MDS), to gain insight into the dynamic nature of the extrusion mechanism implicated in MDR.

1.5 | Approach of the study

Given the time- and cost-effectiveness of using in silico methods, computer-aided drug design (CADD) techniques have gained popularity in the drug discovery and design/development industry and have greatly progressed over the past few decades. This study employs a two-fold bioinformatics approach to study the extrusion mechanism of the MATE transporter proteins. Firstly, molecular docking will be used to identify potential inhibitors against the MATE proteins from the SANC database, followed by a molecular dynamics (MD) approach to shed light on the dynamic behaviour of the MATE transporters, modeled in their native membrane environment – in both their *apo*- and substrate-bound forms. Together, combining these techniques in succession would allow for the identification of possible inhibitory lead drug compounds against MATE transporters, as well as provide a better understanding of their functional mechanism of drug efflux – both of which feed into the potential of eventually developing drugs that target these proteins and thus address the challenge of MDR. The flowchart in Figure 1 below outlines the overall methodology of this study.

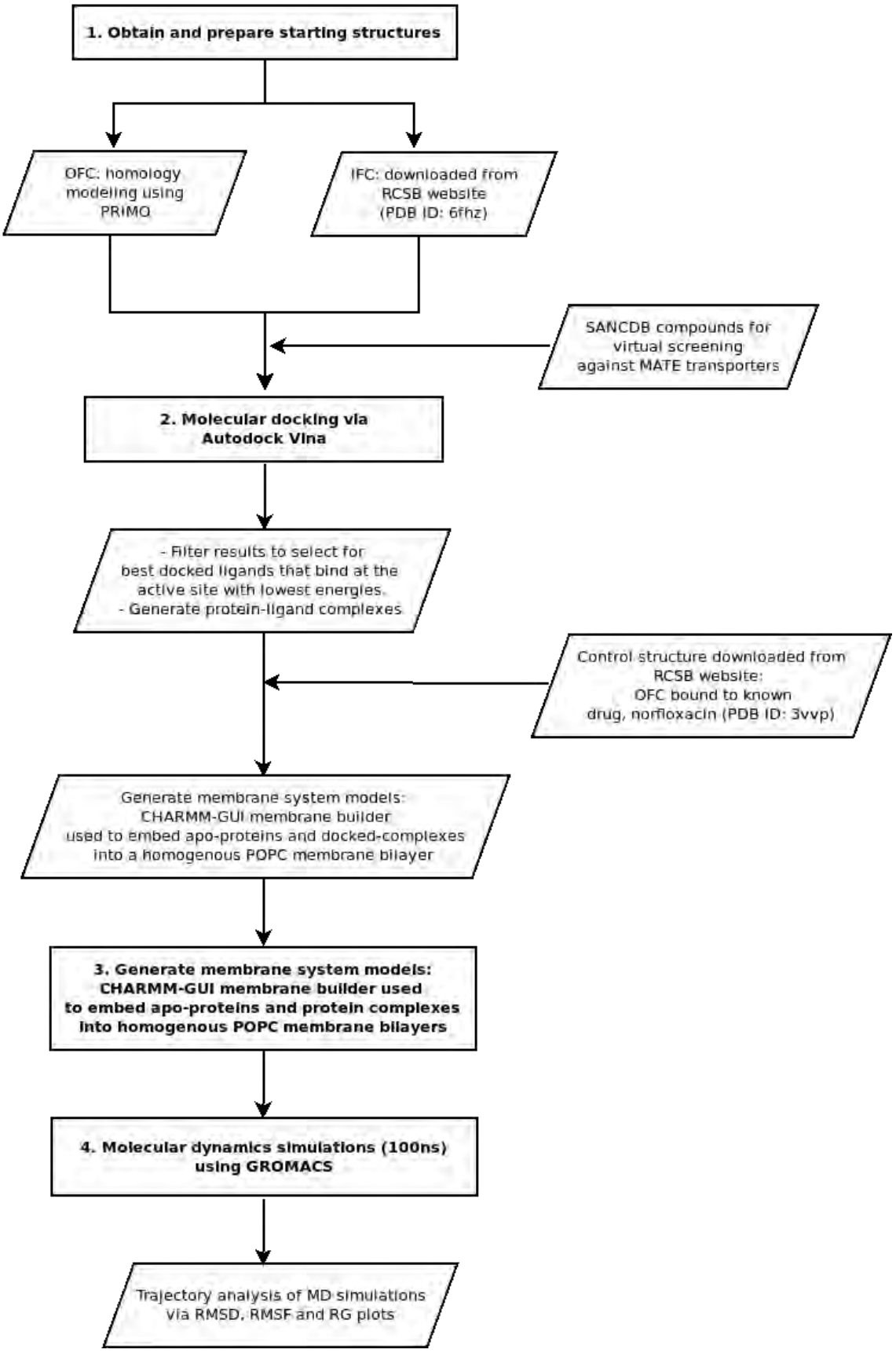


Figure 1.1 | Summary flowchart of study approach: This study will use a two-fold approch, combining molecular docking and MD simulations to explore and gain a better understanding of structure-function-dynamics link of the MATE transporters.

1.6 | Scope of study

Contemporary drug discovery and development involves three main steps – identifying disease-related biological targets, identifying, and optimizing new 'lead' compounds, and developing a novel drug (Liu et al., 2018). The scope of this study lies in the preliminary stages of the drug discovery and design process.

Firstly, molecular docking will be performed to identify potential inhibitor compounds from the SANC-database against MATE transporters via virtual screening. The potential inhibitory compounds identified will serve as lead compounds for future drug design/development stages.

MD simulations are frequently employed successfully in contemporary drug discovery and design as they provide insight into the dynamic structural nature of biomacromolecules as well the energetic nature of protein-ligand interactions by simulating their behaviour within their natural environments (Salo-Ahen *et al.,* 2021). This allows for a better overall understanding of the link between the structure and function of the target macromolecule at a molecular level, as well as the underlying protein-ligand interactions, providing information valuable to the drug discovery and design process (Liu *et al.,* 2018). As such, the molecular docking method will be complemented by molecular dynamics (MD) simulations of the docked complexes. This is done with the MATE transporters in a POPC lipid membrane bilayer (modelling its native environment) in order to elucidate the possible mechanism of the protein and the way in which it interacts with potential inhibitory drug compounds. It will also be explored how these findings correlate to the literature. The combination of these two methods will feed into the future possibility of targeting MATE transporters to prevent their drug extrusion function, in turn offering a potential solution to preventing MDR.

1.7 | Thesis structure

The remainder of this thesis follows the below structure:

Chapter 2: Literature review

Chapter 3: This section briefly introduces molecular docking and its role in drug discovery and design/development, followed by the docking methodology used in this study. It then provides the results and selection of docked complexes on which MD simulations were performed. Finally, a chapter summary is provided.

Chapter 4: This section provides a brief introduction to MD simulations, which is followed by the methodology used, the results and discussion, and finally, a chapter summary.

Chapter 5: This section discusses how the two-fold docking and MD approaches used meet the aims of the study. In addition, it highlights the limitations of this study as well as importance of this study in the process of (future) drug development and design.

CHAPTER 2: LITERATURE REVIEW

2.1 | Multidrug resistance and membrane transporters

Multidrug resistance (MDR) – the simultaneous resistance to multiple drugs with (chemically and structurally) diverse properties and varying targets (Szakács et al., 2006) – has become a growing threat worldwide, responsible for millions of deaths annually. MDR leads to death by impeding effective treatments of cancer, bacterial infections such as tuberculosis (TB), and viral infections including HIV and malaria – whereby treatment drugs fail to reach their biological targets in sufficient concentrations to cause the desired effect(s) (Blanco et al., 2016). As a direct consequence, treatment of infections is now more challenging than ever owing to the rise in multidrug-resistant bacteria (and other pathogens). In 2019, infections caused by multidrug bacteria lead to the deaths of over 1.27 million people, and likely millions more. (Antimicrobial Resistance Collaborators, 2022). As such, agencies like the CDC and WHO have declared MDR as a serious cause of concern (Tacconelli *et al.,* 2018).

Resistance to multiple drugs (MDR) in bacteria can occur through two main ways. The first is by the accumulation of multiple genes – on resistance (R) plasmids or transposons – each of which code for resistance to a particular drug, and secondly via the secretion of drugs by multidrug efflux pumps; both methods allow for the expulsion of various types of drugs (Nikaido, 2009). Antibiotic resistance via efflux pump action was first discovered as a means of resistance in bacteria found to exhibit resistance to tetracycline (McMurray *et al*., 1980), although later efflux pumps were found to be mediators of resistance (against multiple drugs) that are present in all living organisms (Morita et al., 1998; Brown et al., 1999; Nikaido, 2009).

The capacity of membrane transporters to export a wide spectrum of structurally and chemically heterogeneous toxic compounds has been implicated as a major underlying cause of MDR. Based on their promiscuity toward substrate specificity, they have earned the title of multidrug transporters or MDR efflux pumps (EPs) (Webber et al., 2003; Piddock, 2006; Claxton et al., 2018). EPs are widely prevalent in all living organisms and are said to be highly conserved and evolutionarily primitive elements (Saier et al., 1998; Martinez et al., 2009a; 2009b; Nikaido, 2009; Martinez, 2009a; Alcalde-Rico et al., 2016; Sharma et al., 2019). They are known to facilitate the efflux of foreign chemical substances – known as xenobiotics – and metabolites out of the cell via active transport, a vital process through which homeostasis is maintained (Dijun et al., 2015).

Homeostasis, or the maintenance of an optimal constant internal environment, is vital for normal functioning and, thus, the survival of organisms. The survival of the cell/organism is largely reliant upon the crucial process of the active extrusion of metabolites and xenobiotics from the cell, a role played by membrane transporters. However, prokaryotes and eukaryotes have also evolved mechanisms that harness these specialised transport systems, evading drug therapies and diminishing the effects of antibacterial and even anticancer drug therapies by drastically reducing their cellular concentrations (Higgins, 2007; Du et al., 2015). As such, membrane transporters are deemed a leading cause of MDR (Gottesman et al., 2002; Dijun et al., 2015; Alibert et al., 2017; Shinya et al., 2022); they are chief influencers of drug pharmacokinetics, hindering drug therapies (Higgins, 2007; Du et al., 2015; Claxton, 2021).

In 2017, and for the first time, the WHO published a list of 12 MDR pathogens, that still today, are a danger to the lives of millions, requiring the rapid design and development of newer antibiotics to fight off MDR-related infections (WHO, 2017). In many of these organisms, efflux pump action has been pinpointed as a major cause of the development of antibiotic resistance (Alav et al., 2018). EPs are membrane-bound proteins found in all organisms. They play an important role in the resistance of organisms to antibiotics and other toxic substances by pumping out these compounds from the cell (Saier et al., 2001; Martinez et al., 2009; Nikaido, 2009).

Several families of membrane EPs known as drug transporters are known to function as mediators of cellular drug transport out of the cells (Higgins, 2007; Du et al., 2015) that, through varying mechanisms, play a role in MDR. Drug transporters in humans are key players in the pharmacokinetics of drugs in the body. This means that they regulate the ADME/tox (absorption, distribution, metabolism, elimination/toxicity) properties of drugs in the body. Drug transporters are divided into two major groups; the solute carriers (SCL) and the ATP-binding cassette (ABC) group. As the name suggests, ABC transporters are driven by ATP, whereas SLC transporters depend on an electrochemical gradient to drive substrate transport. Typically, SCL transporters are involved in the uptake of drugs into the cell. However, MATE transporters are the exception and, contrastingly, play a role in exporting organic cationic substances from the cell (Liu and Pan, 2019). In bacteria, seven families of drug transporters have been identified; classified based on differences in their primary sequences, structural folding, and mechanisms of cellular efflux. These families include the ATP-binding cassette (ABC) family, small multidrug resistance (SMR) family, resistance-nodulation-division (RND) family, major facilitator superfamily (MFS), the more recently understudied multidrug and toxic compound extrusion (MATE) family, the proteobacterial antimicrobial compound efflux (PACE) family, and lastly, the AbgT proteins (Morita et al.,1998; Brown et al., 1999; Putman et al., 2000; Saier et al., 2001; Hvorup et al., 2003; Piddock, 2006; Delmar and Yu, 2016; Hassan et al., 2018; Claxton, 2021). Figure 2.1 below shows a

schematic diagram of five of these families that are found spanning the lipid membrane in bacteria, highlighting their different energy-coupling mechanisms driving substrate efflux.

Based on their energy source, EPs can be classified as primary transporters – that require ATP for substrate export such as the ABC transporters – or secondary (active) transporters – that in contrast, require a proton gradient to drive substrate efflux such as the MATE, SMR, MFS, and RND transporters (Kumar and Schweizer, 2005; Tanaka *et al.,* 2013; Spengler *et al.,* 2017; Alav *et al.,* 2018; Lamut *et al.,* 2019), shown in Figure 2.1 below. The characteristics of MATE transporters as secondary active transporters are discussed in detail in Subsection 2.4.

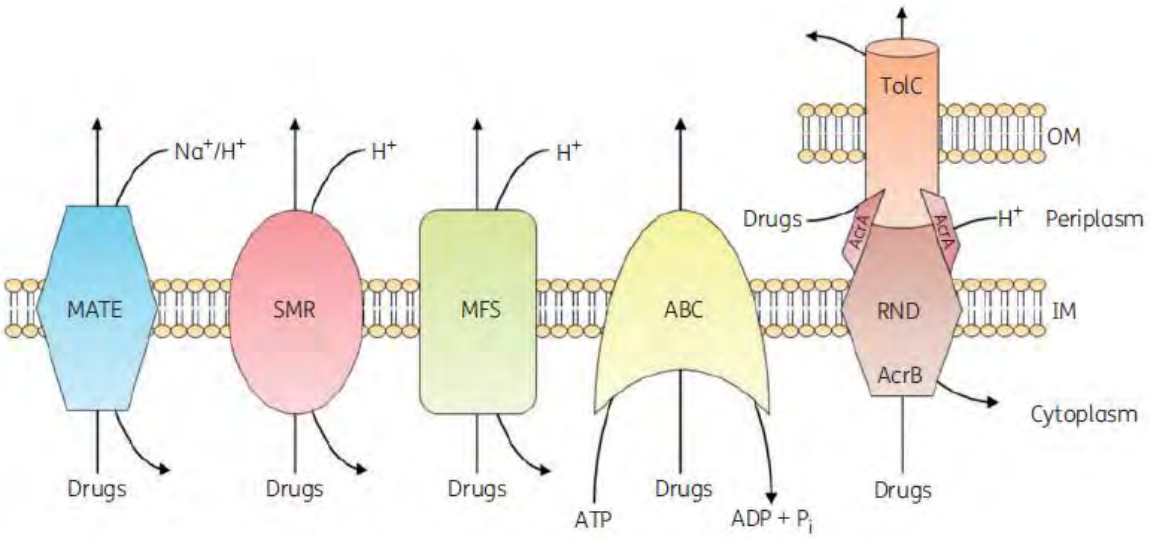


Figure 2.1 | Five families of membrane EPs found in bacteria that play a role in drug efflux (adapted from Alav et al., 2018): the energy-coupling mechanisms fuelling the active efflux of drug substrates for each of these EPs is shown. In contrast to the other families of EPs, MATE transporters utilize an H^+/Na^+ inward-directed gradient, expelling drugs from the cell to the periplasm.

While these transporters exhibit divergence in their efflux mechanisms, there is an overlap in their substrate specificity, suggesting that they work as a unified network in the export of potentially toxic substances (Claxton, 2021). As such, determining the underlying structural and functional frameworks of multidrug transporters (or MDR EPs) is key to a better understanding of MDR and to figuring out possible future strategies to combat MDR. These aspects of MATE transporters will be discussed in the subsections to follow.

2.1.1 | Drug efflux pumps and efflux pump inhibitors

EPs are key facilitators of drug efflux, giving rise to multidrug-resistant pathogens and hindering drug therapies by significantly reducing cellular drug concentrations. Antimicrobials of all kinds are expelled from cells by the action of drug EPs, leading to the development of MDR. (Poole, 2007; Sun et al., 2014). As such, preventing their efflux function would, in turn, prevent the occurrence of MDR, making it a worthwhile strategy to potentially evade the problem of resistance and allow for effective drug treatments (Poole et al., 2006; Kapp et al,. 2018; Sharma et al., 2019). Membrane proteins make up about two-thirds of the determined cellular drug targets, known to bind to approximately 50% of recognized small-molecule drugs (Lappano and Maggiolini, 2011; Tautermann, 2014).

The following are five main ways by which drug efflux can be prevented: 1) via the downregulation of EP gene expression; 2) structural reconstruction of existing drugs to make them unidentifiable as EP substrates; 3) obstructing the assembly of functional units of EPs; 4) targeting the energetic driving force underlying the efflux mechanism; 5) preventing substrates from binding to EPs by obstructing the binding site (Bhardwaj and Mohanty, 2012). The latter of these – protein inhibition via binding of an inhibitor at the active site – is of interest to the current research. Compounds able to inhibit the function of efflux pumps are accordingly termed efflux pump inhibitors (EPIs) and have the potential to serve as pharmaceutical agents through inhibition of the target EP (Sharma et al., 2019). In theory, EPIs would block the drug efflux action of their drug EP targets, which would allow for the maintenance of sufficient cellular concentrations of antimicrobial treatment drugs, thereby increasing the susceptibility of the multidrug-resistant pathogen to the antimicrobials (Soto, 2013; Pule et al., 2016; Reza et al., 2019). EPIs can thus be a useful treatment option for MDR-related infections since they function well in conjunction with antimicrobial medications (Shinya et al., 2022).

In conclusion, EPs – such as MATE transporters – have become popular as drug targets, as they are vital to the survival of bacterial pathogens through their efflux function. As such, the discovery/design of efflux pump inhibitors (EPIs) is a feasible potential solution for combating the occurrence of MDR (Poole et al., 2006; Kapp et al., 2018; Sharma et al., 2019). Therefore, investigating their structure-function relationships is vital to understanding their underlying mechanism(s) of action in order to utilize them as antimicrobial drug targets.

2.2 | Membranes and Membrane Proteins

Cell membranes – primarily made up of phospholipids and membrane proteins – are responsible for creating and actively sustaining a significantly different internal (or cellular) environment and also maintain electrical and solute gradients necessary for cellular function (Yeh et al., 2020). Membrane proteins help to maintain this optimal internal environment (homeostasis), playing a role in electrical polarization, signalling, remodelling of the membrane, and aiding in transport across the membrane, as well as other critical processes (von Heijne, 2007). Membrane proteins make up 30-40% of proteins encoded in most genomes (Krogh et al., 2001; Birch *et al.,* 2020) and are split into two groups – peripheral and integral proteins.

Peripheral proteins temporarily attach or tether to the membrane using electrostatic or other interactions with a concurrent soluble portion within the outer aqueous surroundings (Whited and Johs, 2015). On the other hand, integral membrane proteins – such as MATE transporters – are permanently embedded into the membrane matrix (partially or fully). They span either one or both leaflets of the membrane bilayer, either via a transmembrane domain (TMD) topology or are tethered using a membrane anchor (Whited and Johs, 2015). Furthermore, integral protein folding involves the formation of a transmembrane topology belonging to one of two classes – α-helical bundles or β-barrels. Those with α-helical bundle arrangements have been found to exhibit a greater degree of flexibility, allowing for conformational changes, usually seen in receptors, channels, and transport proteins (Kobilka, 2007; Shimada et al., 2019).

Transport proteins such as MATE transporters are integral membrane proteins (IMPs) that, in the human body, regulate the entry of xenobiotics (drugs) and endogenous chemicals throughout numerous sites (Giacomini *et al.,* 2010; Giacomini and Sugiyama, 2017). In contrast, drug-metabolizing proteins (or enzymes) are primarily found localized in the liver and intestines, transporter proteins are found in all bodily tissues, where they are involved in the vital roles of drug absorption, distribution, tissue-specific drug targeting, and finally, elimination. As such, MATE transportes greatly influence drug pharmacokinetics and pharmacodynamics through their role in the cellular export of drugs and other endogenous compounds, influencing drug-drug interactions in humans (Hillgren *et al.,* 2013). This role of MATE transporters will be discussed in Section 2.5.

Several medications interact with membrane proteins; in fact, membrane proteins make up 60% of the pharmacological targets of marketed drugs (Bakheet and Doig, 2007; Birch *et al.,* 2020), highlighting their significance extending past their physiological cellular functions (Yeh et al., 2020). As such, deciphering their 3D structure-function relationship is highly important to scientists as well as beneficial to the healthcare/pharmaceutical industry. Structural biology provides information on the

conformational and molecular dynamics of these proteins as well as drug binding mechanisms, enhancing our understanding of membrane proteins as therapeutic targets (Birch *et al.,* 2020), which is useful for structure-aided drug discovery (Kermani, 2021).

The properties of membrane proteins make them significant pharmacological targets, and, in order to develop drugs and achieve some degree of cellular control, it is necessary to gain insight into their structure and function (Yeh et al., 2020). Although, given their complex physical characteristics, membrane proteins are challenging to study. Hence, several gaps still need to be filled to gain a strong understanding of their structure and mechanism underlying their function(s). Hitherto, X-ray crystallography is the most efficient tool for the determination of 3D structures of membrane proteins, but high-resolution of 3D structures is an arduous process reliant upon the effective extraction of proteins, solubilization, stabilization, and finally, producing protein crystals through diffraction. In the case of membrane proteins especially, each step of this process of structure resolution can be rather cumbersome (Kermani, 2021). In addition, membrane proteins interact with lipids of the membrane bilayer, which is crucial to their function, but this adds another layer of complexity to the process of extraction by detergents, which can damage membrane protein complexes or affect both structural and functional aspects of the extracted protein (Guo, 2020). Therefore, studying membrane proteins is challenging and necessitates a membrane environment that supports and sustains the protein in its folded, functional state (Yeh et al., 2020).

While techniques, including crystallography and electron microscopy, have benefited structural analyses, they still need to be improved in several aspects. These include the characterization of essential functional interactions with membrane lipids, small molecules, and membrane modulators, and in addition and more importantly, their conformational diversities and dynamics (Yeh et al., 2020). The alternative method is nuclear magnetic resonance (NMR) spectroscopy, which avoids the necessity for protein crystallization, preserves molecular flexibility and dynamics throughout the investigation, and keeps the protein in its functional state. Moreover, NMR provides structural data from entire cells and describes conformational fluctuations of ligands and lipids linked to membrane proteins in sub-atomic detail and is thus a chief technique for studying both protein structure and dynamics (Yeh *et al.,* 2020).

However, technological advances and the development of computational techniques, including homology modelling and MD simulations, have been monumental in exploring structural, functional, and dynamic aspects of proteins in a time- and cost-effective manner. These techniques were employed in the current study and will be discussed in their respective chapters.

2.3 | Protein dynamics: sequence-structure-function relationship

Proteins are exceptionally dynamic structures, and their intrinsic motions allow them to undergo the conformational changes they need to interact with other molecules and/or adjust to changes in the surrounding environment, which is essential to their physiological function (Teilum et al., 2009). Conformational changes in proteins involve the cleavage of some non-covalent interactions/bonds and the creation of others. This capacity to alter their conformation or accommodate changes is crucial for many biological and physiological processes, such as signal transmission, antigen recognition, protein transport, and enzyme catalysis (Teilum et al., 2009).

The adaptive nature – or flexibility – of a protein brings forth either subtle conformational changes that occur via the movement of amino acid side chains during substrate binding or more radical conformational changes when specific ligand(s) induces folding in some proteins (Teilum et al., 2009). Rackovsky and Scheraga (2020) recently demonstrated that the link between structure and dynamics is determined by the global structural flexibility of a protein, as opposed to more local side-chain properties. Furthermore, a converse link between protein stability and physiological function has been established in both enzymes and protein hormones, highlighting the importance of protein flexibility for their function (Schoichet et al., 1995; Jomain et al., 2006).

Describing the dynamic motions taking place within a protein requires the use of both experimental – most notably spectroscopic – as well as computational methods - most notably molecular dynamics (MD) simulations. The latter of these is a powerful tool allowing for the characterization of rapid dynamic motions on the (sub-)nanosecond timescale (Dodson and Verma, 2006; Dodson et al., 2009; Teilum et al., 2009). MD simulations are thus a promising method for the characterization of protein functional dynamics. Details of this technique are described in Chapter 4.

2.4 | Secondary active transporters

All living organisms contain widely dispersed secondary active transporters, acting as molecular transport machines of cells (Forrest et al., 2011). This subsection investigates the common structural and functional features of secondary transporters, such as MATE transporters as the basis upon which the MATE family of secondary active transporters will be discussed in the subsections to follow.

While their transport methods differ, a distinguishing shared characteristic of secondary active transporters is their need for an energy source provided by an electrochemical potential in the form of an ion gradient to propel the movement of a substrate (or ligand) (Zhang et al., 2015). Additionally, binding of a substrate to the protein on one side of the membrane induces a conformational change in

these transporters, allowing for the release of the substrate on the converse side of the lipid membrane. Their substrates are numerous and can include neurotransmitters, nutrients, as well as a range of drugs (Boudker and Verdon, 2010).

Although their primary sequences and functions differ, and they have thus been classified as separate families (Hediger et al., 2004 ; Saier et al., 2006), many secondary transporters evolved from a common ancestral gene and are structurally related (Lolkema and Slotboom, 2003; Bret et al., 2005; Lolkema and Slotboom, 2008), indicating the possibility that some membrane proteins later evolved various structural folds that account for the wide range of distinct functions they perform within the cell (Oberai et al., 2006; Theobald and Miller, 2010).

All secondary transporters have been found to exhibit structural repeats, which is often manifested in form of a specific 3D fold comprising a number of adjoining transmembrane segments (TMs), typically repeated twice or thrice. These structural repeats may have emerged as a result of gene duplication and thereafter by fusion and sequence divergence (Steiner-Mordoch et al., 2008). As a result, these structural repeats can go undetected, demanding the resolution of 3D crystal structures (Boudker and Verdon, 2010). In addition, structural repeats exhibiting parallel alignments are linked via a pseudosymetrical axis that is perpendicular to the lipid membrane plane (Boudker and Verdon, 2010).

The symmetry of the structural repeats is known to be a key component governing the molecular-level mechanism of secondary transporters (Boudker and Verdon, 2010). It is thought that the pseudo symmetrical TM bundles generate rocking motions that rotate around the bound substrate and are responsible for the alternating access of the substrate-binding site to either side of the membrane (Guan and Kaback, 2006; Law et al., 2008). This mechanism, known as the ‘rocker switch alternating access’ model, seems to fit the proposed mechanism of the substrate efflux function of MATE transporters and will be discussed in detail in the following subsections.

Studies on the transport mechanisms of secondary transporters indicate an integration of two distinct conformational states involving the isomerization between inward- and outward- facing conformational states. Additionally, the presence of further intermediate conformational states are likely to occur (in both substrate-bound and *apo* forms) between the two main conformations during the transport cycle (Forrest et al., 2008; Shimamura et al., 2010; Boudker and Verdon, 2010). Thermodynamically, the difference in the electrochemical potentials of ions/substrates on both sides of the membrane governs the transport cycle and dictates how many intermediates are formed and how stable they are (in terms of their free energies) (Forrest et al., 2008; Shimamura et al., 2010; Boudker and Verdon, 2010).

The effects of conformational transitions between the inward and outward states on the ligand-binding site have also been investigated, yielding contradictory findings based on their structures. In the case where the substrate-binding site is fully formed within the protein core, transmembrane movements of the core have no influence on the substrate selectivity and binding as the network of hydrogen bonds and ion interactions at the core remain unaffected (Padan et al., 2009 ; Appel et al., 2009). On the other hand, secondary transporters with their substrate-binding sites formed within interfaces between structural components are likely to move as a result of the conformational changes that occur during the transport cycle; The transport cycle is complete when the transporter reverts to the outward-facing conformation (Boudker and Verdon, 2010).

The ion-coupling feature of secondary transporters has also been studied (Boudker et al., 2007; Groeneveld and Slotboom, 2010). In theory, isomerisation between the inward- and outward- facing conformations should only occur provided both the substrate- and ion-binding sites are either occupied or unoccupied, respectively. This allows for the ion-coupled transport of substrates as well as the prevention of wasteful leakage. As such, the presence of a stringent thermodynamic coupling between the binding of a ligand and the generation of a transport-effective conformational state is expected (Boudker and Verdon, 2010). Moreover, a difference in the ion preference as well as the transport stoichiometries has been found amongst evolutionarily related transporter families (Krishnamurthy et al., 2009), said to be associated with the type of ion gradients present as well as the measure of the electrochemical gradients of the substrates, respectively (Boudker and Verdon, 2010).

Given the functions of secondary membrane transporters – such as the efflux of (treatment) drugs – they are deemed an essential class of proteins involved in physiology and disease and are often popular drug targets. However, compared with soluble proteins, their placements within the lipid bilayers of cells as well as their hydrophobic nature, makes them rather complex to study (Guan, 2022). As drug targets, researchers are often interested in inhibiting their functions contributing to disease. A full transport cycle includes the binding of a substrate followed by several conformational changes, any stage of which can be targeted via inhibition of the protein to prevent the substrate transport process (Boudker and Verdon, 2010). As such the substrate-binding site itself becomes the target of inhibitors (or competitive blockers) typically made up of chemical elements of a similar nature to that of the actual substrate along with added large aromatic groups (Boudker and Verdon, 2010). Several studies have found that the size threshold of the inhibitor/blocker divides the transported substrates and the non-transportable inhibitors (Hirayama et al., 2001; Singh et al., 2008; Loo et al., 2008; Kalra, 2014). For example, several antidepressants act as competitive inhibitors of neurotransmitter sodium symporters (NSS) by binding to the outward-facing transporter conformations using an Na^+ and Cl^- ion gradient (Beuming et al., 2008; Andersen et al., 2009; Tavoulari et al., 2009), highlighting the likelihood of overlap in the binding

site of the inhibitor and substrate (Boudker and Verdon, 2010). A classic example of this overlap in binding sites is observed in the competitive binding of cocaine versus dopamine-to-dopamine transporters (Beuming et al., 2008). Structural studies on competitive inhibition and size of inhibitors/blockers have been investigated in the case of the mammalian glutamate transporters whereby the competitive inhibitor, TBOA, which is similar in coordination to the substrate, is found to bind within the substrate-binding sites. However, the large size of TBOA prevents any substrate from approaching the transporter (Pavić et al., 2019).

Finally, the resolution of 3D structures of active secondary transporters captured in their various conformations during their function in the transport cycle, has found that the substrate- and/or inhibitor-binding sites are at or close to the interfaces of structural domains of the transporter. Typically the transport cycle involves the conformational exchange of the pseudosymmetrical structural repeats within these domains responsible for inducing the transition between the two major inward- and outward- facing conformations (Boudker and Verdon, 2010). The following subsections will discuss the structural and functional features of the MATE transporter family in light of the above discussion and that of the previous sections.

2.5 | MATE family of EPs

MATE transporters are mediators of the active transport of a range of solutes across cellular membranes and have a conserved structure across all life forms, including bacteria, archaea, fungi, plants, and animals, where they play numerous diverse significant roles (Claxton et al., 2021).

MATE transporters of the plant kingdom are well established and known to be one of the largest transporter families, whereby each individual species contains several MATE proteins encoded in their genome, each of which is involved in the transport of specific substrate(s) (Tanaka et al., 2021). The genome of *Arabidopsis thaliana*, for example, encodes 58 MATE proteins, whereas that of *Glycine max* encodes 117 MATE proteins (Hvorup et al., 2003; Liu et al., 2016), which is indicative of the diverse evolution in functions of plant MATEs (Kuroda and Tsuchiya, 2009; Takanashi et al., 2014; Upadhyay et al., 2019). Similar to MATEs of other kingdoms, plant MATEs transport a wide range of substrates such as organic acids, plant hormones, and secondary metabolites, and in addition have been found to several other diverse roles in physiological processes across plant development (Diener et al., 2001; Takanashi et al., 2014). For example, plant MATE transporters are involved in a myriad of roles centred around detoxification, specifically aluminium (Al3+) (via the secretion of secondary metabolites) and defence against plant pathogens and predators (Morita et al., 2009; Takanashi et al., 2014; Miyauchi et

al., 2017). Other physiological roles played by plant MATEs include efflux of xenobiotics (as in other organisms), gathering of secondary metabolites (such as alkaloids and flavonoids), translocation of iron (Fe), and hormone signalling, to name a few – all of which are vital to plant development (Takanashi et al., 2014). Plant MATEs have received a lot of attention and have been studied in depth in comparison to other MATE proteins. Studies conducted on over 40 plant MATEs not only show that they are important agents in plant physiology, but they also exhibit restricted substrate specificities (Takanashi et al., 2014). In fact, MATE transporters of the plant kingdom have evolved to such a degree that each MATE protein primarily exports a specific substrate(s) (Kuroda and Tsuchiya, 2009; Takanashi et al., 2014; Upadhyay *et al.,* 2019). MATE proteins are mainly located in both the plant vacuoles as well as in the plasma membrane, where they use an H+ ion gradient for their substrate export (Yazaki, 2005). Plant MATE proteins constitute families that can grow from dozens to hundreds of members (Liu et al., 2016; Zhu et al., 2016; Li et al., 2017). The occurrence of considerable numbers of plant MATE proteins is thought to possibly be linked to their ability to export a range of secondary metabolites (Omote et al., 2006). Plant MATEs are vital transport proteins contributing greatly to plant survival and development, and are thus of importance to the agricultural industry.

MATE transporters in both bacteria and humans extrude multiple drugs from the cytoplasm out into the extracellular environment and are fewer in number compared to plants. MATE transporters in bacteria have thus been implicated as a major cause of pathogenic resistance (i.e., MDR) (Morita et al., 1998; Morita et al., 2000; Braibant et al., 2002; Su et al., 2005). For example, the nosocomial infection caused by the multidrug (or multi-antibiotic) resistant pathogen *Staphylococcus aureus* (MRSA), which has become a notorious clinical pathogen, causing disease and death in humans through multiple drug resistance (Tanaka et al., 2013). Bacterial MATEs are known to cause resistance to cationic drugs such as ethidium bromide (EtBr), tetraphenylphosphonium, berberine, acriflavine, and norfloxacin, exhibiting a preference for the efflux of organic cationic substrates (Tanaka et al., 2013). It should be noted, however, that the molecular-level mechanism of organic cation efflux is not well understood (Otsuka et al., 2005; Omote et al., 2006).

Following genomic analysis, mammalian MATE proteins have been grouped under the eukaryotic subfamily of MATEs, and each species was found to have some MATE orthologs (Otsuka et al., 2005). Humans have 2 MATE orthologs, MATE1 (gene symbol, SLC47A1) and MATE2 (gene symbol, SLC47A2), whose genes are found in close proximity to one another on chromosome 17, as well as asplice variant, MATE2K. (Otsuka et al., 2005; Hillgren *et al.,* 2013). While MATE1 is found to be expressed throughout the body, it is largely expressed in the kidney and liver, concentrated at the bile canaliculi and brush border membranes, respectively; MATE2, on the other hand, is particularly expressed in the kidneys, concentrated at the brush border membranes (Otsuka et al., 2005). Through

cis-inhibition studies, it is known that transport substates of MATE1 proteins are organic cations of vastly different chemical structures (Otsuka et al., 2005). MATE2 proteins are also known to transport diverse organic cations such as the type 2 anti-diabetic drug, metformin, the histamine cimetidine, tetraethylammonium (TEA), 1-methyl-4-phenylpyridinium (MPP), and N-methylnicotinamide (NMN), doing so against an H+ ion gradient (Masuda et al., 2006). Biochemical and pharmacological profiles of MATE orthologs in both humans and mice, along with their observed centralised subcellular locations have revealed properties of these renal and hepatic transporters that make them suitable for the transport of organic cations, essentially accountable for the final step of excretion (Omote *et al.,* 2006). In other words, MATE transporters in humans and mice are responsible for the end excretion of possible harmful organic cationic compounds – including cationic treatment drugs – using a coupled H^+ electrochemical (Otsuka et al., 2006; Masuda et al., 2006; Hiasa et al., 2006). Furthermore, MATE transporters are known to function alongside organic cation transporters (OCT1 and OCT2), whereby a substrate overlap is observed (Hillgren *et al.,* 2013), adding to the complexity when it comes to studying these transporters.

In tumour cells, MATE transporters are known to confer cellular resistance against anticancer drugs, simultaneously exporting structurally and chemically diverse drugs out of the cell, preventing them from reaching the required cellular levels for effective treatments (Tanaka et al., 2013). MATE transporters of bacteria and cancer cells directly lead to MDR, demanding the urgency for discovery and design of drug compounds with the capacity to prevent, via inhibition for example, the efflux action of MATE transporters. Thus far, however, screening for lead drug compounds have come up short (Tanaka et al., 2013). In addition, the lack of understanding of the underlying mechanism of their efflux action also requires urgent attention, given the importance of MATEs as drug targets and agents of MDR.

The MATE family transporters are known to provide (cellular) resistance against a diverse range of therapeutic agents, including aminoglycoside antibiotics, fluoroquinolones, and anticancer treatments, making them promising drug targets to combat MDR mechanisms that hinder treatments globally (Higgins, 2007; Claxton et al., 2018). They stand out from other known multidrug transporters for their capacity to access the stored energy of either the Na^+ or H^+ (or both) electrochemical gradient to drive their transport cycle, as in the case of the human and most bacterial MATE proteins, respectively (Lu et al., 2013a), discussed further in the subsections below.

2.5.1 | Classification, structure, and topology of the MATE family of EPs

Initially, MATE transporters were classified within the MFS superfamily. However, phylogenetic studies and subsequent structural characterization revealed distinctions in both sequence and in topology amongst the two subfamilies (Morita et al.,1998; Brown et al., 1999; Kusakizako et al., 2020), earning

them a separate name. Furthermore, the MATE subfamily has been classified under the multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) flippase superfamily (Hvorup et al., 2003).

Most MATE transporters usually comprise of about 400-500 amino acids, that make up their 12 TMDs – in contrast to other multidrug transporter families such as the MSF transporters with 6 TMDs – and are found to exhibit a broad substrate specificity on account of their core domain, which consists of non-conserved amino acid sequences (Omote et al., 2006; Santos et al., 2017). MATE transporters are known for their role in the efflux of cationic drugs, including norfloxacin and ethidium (Omote et al., 2006). Moreover, on account of their primary sequence similarity, MATE transporters have been divided into three subfamilies, which include the NorM, DinF (DNA damage-inducible protein F), and the Eukaryotic (eMATE) branches (Brown et al., 1999), shown in the phylogenetic tree in Figure 2.2.

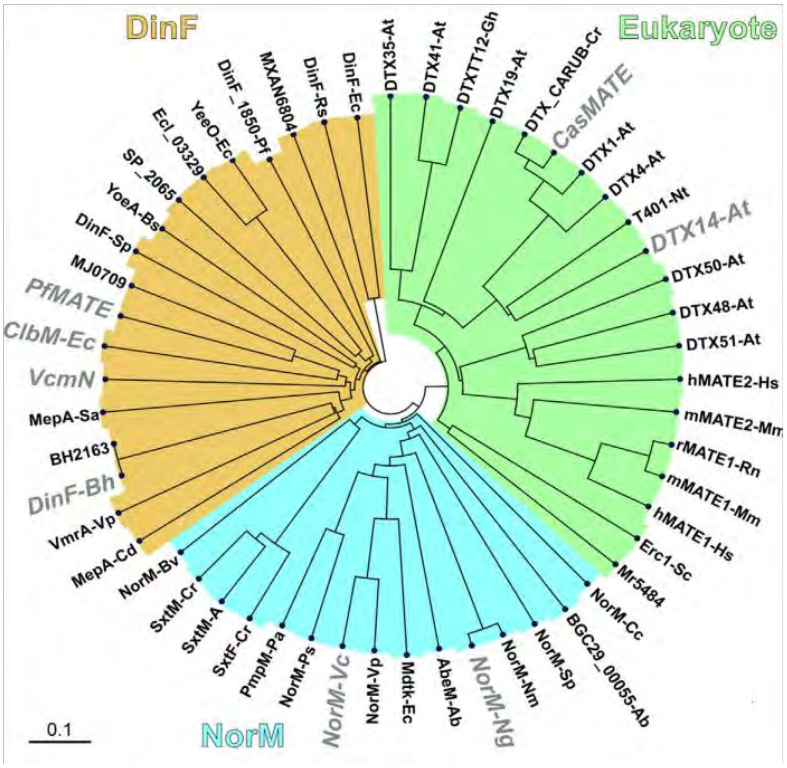


Figure 2.2 | Phylogenetic tree of MATE transporters highlighting the three subfamilies (adapted from Claxton *et al.*, 2021): The evolutionary relationship between the 53 classified members of the DinF, NorM, and eMATE subfamilies is shown, with the transporters that have available structures in bold grey.

Available 3D crystal structures illuminate that all MATE transporters share a conserved fold consisting of 12 transmembrane (TM) helices organized into two bundles of 6 α-helices each, exhibiting an intramolecular two-fold symmetry about an axis perpendicular to the membrane plane (Lu, 2016; Kusakizako et al., 2020). As discussed previously, secondary transporters with α-helical bundles are flexible and able to undergo conformational changes. The two bundles, referred to as the N- and C-lobes – consisting of TM1-TM6 and TM7-TM12, respectively – form a V-shaped cleft open to the extracellular side in the case of the OFC, and an Λ-shape in the case of the IFC (Lu et al., 2013a; Tanaka et al., 2013; Ranchenko et al., 2015; Moussa et al., 2016; Miyauchi et al.,, 2017; Zakrzewska, 2019; Claxton et al., 2021), which is a classic architecture that arises between the two anatomically identical outward-facing bundles in rocker-switch proteins (Yan, 2013; Yan, 2015), expanded on in the following subsections. In contrast to other secondary-active transporters, the N- and C-lobes of MATE proteins can be described as pseudosymmetrical with parallel TM topologies, as in the case of the RND family of MDR transporters (Eicher et al., 2014). A large intracellular loop between TM6 (found within the N-lobe) and TM7 (found within the C-lobe) is responsible for linking these bundles together (Zhao et al., 2021). In addition, a central cavity is formed between the N- and C- lobes, which provides surfaces for ion and substrate binding, and can be split into the N- and C-lobe cavities (Tanaka et al., 2021).

The nature of this cavity has been found to vary from subfamily to subfamily, influencing their various preferred substrates. For example, the central cavity found in the NorM branch of MATEs is said to be more negatively charged than those of the DinF branch, transporting more positively charged substrates than the latter group (Tanaka et al., 2017). Members of the DinF branch include VcmN, ClbM and PfMATE. The amino acid sequences of crystal structure representatives from the Eukaryotic group, CasMATE and AtDTX14, reveal fair similarities to the NorM MATEs, both possessing negatively charged cavities (Miyauchi et al., 2017; Tanaka et al., 2017; Tanaka et al., 2021). It is thus no surprise that MATE transporters can export a range of substrates from the cell.

Sequence alignments of some MATE transporters from each of the three subfamilies, found several conserved residues in various TMDs within both the N- and C-lobes, depicted in Figure 2.3 below, grouped by MATE subfamily.

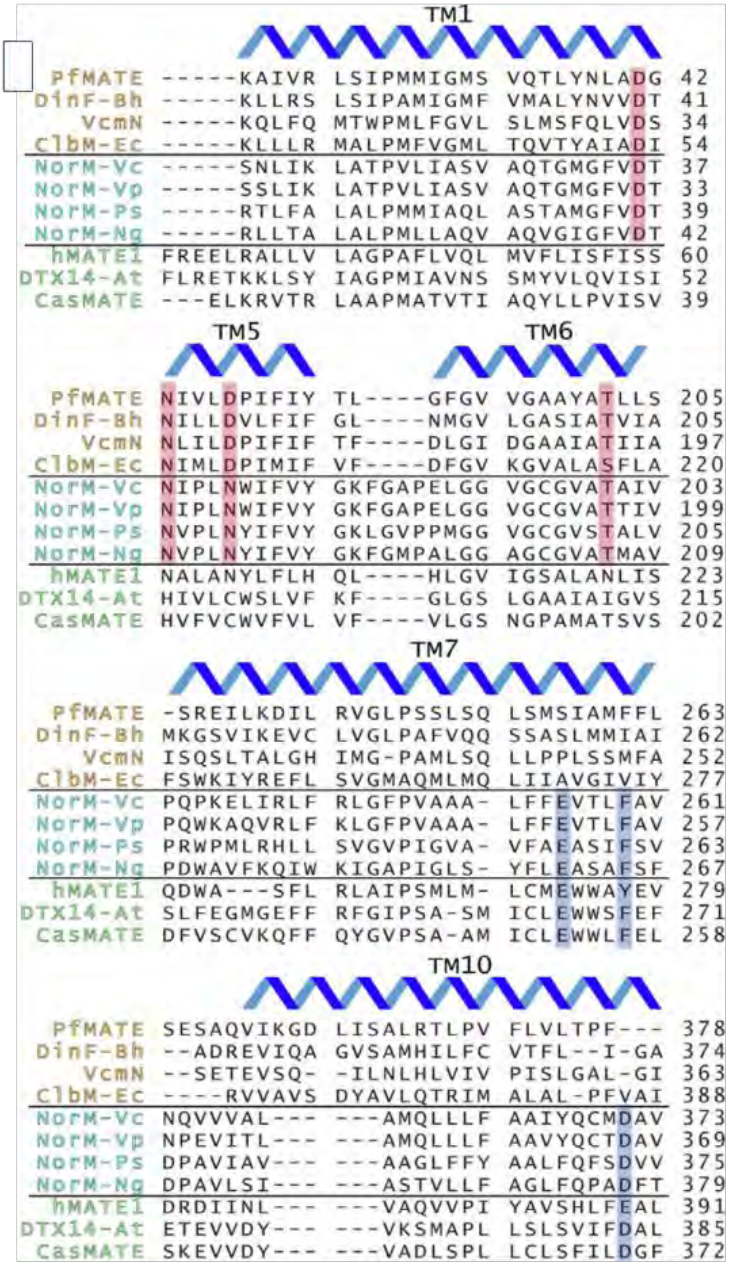


Figure 2.3 | Sequence analysis diagram of 11 representatives of MATE transporters, grouped by subfamilies (adapted from Claxton *et al.*, 2021): conserved residues found to be vital for ion- and/or substrate-binding located in various TMs within the N- and C-lobes are highlighted in red and blue, respectively.

Sites for cation or water binding are present at specific locations within the N- and C-lobes of MATE transporters, containing conserved amino acids, and are deemed vital components of their function (He et al., 2010; Lu et al., 2013b; Krah et al., 2017; Krah et al., 2020; Castellano et al., 2021). These conserved residues have been found to be involved in the binding of the ion and/or substrate to the

transporter, allowing for substrate efflux (Claxton et al., 2021). These aspects of MATE transporters will be discussed in the subsections to follow.

By resolving several predicted conformational states from members of many transporter families, including the MATE subfamily, and subsequently determining their relative ion- and substrate-binding sites, biophysical analyses of the available crystal structures stand by the fundamentals of a shared alternating access mechanism (Drew and Boudker, 2016; Dastvan et al., 2016; Kazmier et al., 2014; Reading et al., 2017; Verhalen et al., 2017; Jia et al., 2020). These studies will be discussed in the following subsections, exploring the ion-specificity and ion-binding site, the substrate-binding site, as well as other structural and mechanistic characteristics of MATE transporters, to gain a better understanding of the efflux action of MATE transporters that is responsible for MDR.

2.5.2 | Antiporter function and ion-binding site and specificity of MATE EPs

Antiporters refer to secondary active membrane transporters that transport their substrate and coupling ion in opposite directions of the membrane (Mitchel, 1967). MATE transporters are secondary active transporter mediators, that use an electrochemical gradient to fuel substrate efflux against the concentration gradient (Mitchell, 1957; Jardetzky, 1966; Drew and Boudker, 2016). In this way, they serve as antiporters, whereby a coupling of the substrate efflux to inwardly directed Na+ or H+ gradients is observed (Kuroda and Tsuchiya, 2009). In other words, they fuel the drug-efflux process by linking it to the influx of monovalent cations along a pre-existing electrochemical gradient (Ficici et al., 2018). According to the antiport model, the conformation that is occupied, whether the IFC or OFC, is influenced by the respective binding of substrate and ions (Claxton et al., 2018).

Various studies have focussed on exploring the cation specificity of MATE transporters, but the results remain inconclusive; even at the amino acid level of any available homolog, streamlining the agents that influence cation specificity has been difficult (Ficici et al., 2018). Some MATE antiporters have been shown to function using an Na+ gradient, such as NorM from Vibrio parahaemolyticus (Morita et al., 2000), NorM-NG (Long et al., 2008) and ClbM (Moussa, 2016), whilst others were found to use an H+ gradient, such as PfMATE (Tanaka et al., 2013), DinF-BH (Radchenko et al., 2015), and NorM-PS (Nie et al., 2016), and interestingly, NorM-VC MATEs were found to be able to use both Na+ and H+ gradients (Jin et al., 2014). In addition, substrate efflux of some MATEs was shown to be driven by other monovalent cations including K+ and Rb+, exhibiting a lack of cation specificity (Huda, et al., 2001; Xu et al., 2003; Hashimoto et al., 2013; Moussa et al., 2017). The cation specificity, however, is frequently examined using cell-based resistance experiments, and thus it should be noted that it is not always evident that the observed effects are due solely or primarily to MATE activity (Ficici et al.,

2018). This issue has also not been definitively resolved by the available prokaryotic MATE structures, given that the resolution of the crystallographic data is often inadequate to determine the existence of an ion-binding site or its precise geometry (Ficici et al., 2018). The structure of PfMATE, with a resolution of around 2.5Å, however, remains an exception. PfMATE has been said to utilise an H^+ gradient, containing a binding site in the N-lobe with the capacity to detect the ion (Tanaka et al., 2013). What is perplexing however, is that the configuration of the side chain of this potential site is remarkably comparable to that of other Na^+ binding sites (Schlegel et al., 2012; Schulz et al., 2013; Guskov et al., 2016). Whilst these sites may often also bind H^+ and even be H^+ selective, at physiological pH and Na^+ concentrations, it is anticipated that substrate efflux is coupled to an Na^+ gradient (Leone et al., 2015). The literature on cation preference of MATE transporters, thus remains inconclusive.

In addition to the apparent lack of ion specificities amongst and within MATE subfamilies, for some MATE proteins, the location of the ion- and substrate-binding sites within the N- and/or C-lobes remains unclear. It has however been established that the majority of MATE transporters have a conserved glutamic/aspartic acid (Otsuka et al., 2005) located near the center of the protein, that is exposed to water molecules, and which may be involved in ion coordination (Vanni et al., 2012).

As antiporters, regardless of cation specificity, however, how exactly do MATE transporters carry out their substrate efflux function? The mechanism behind the substrate-efflux of MATE transporters is yet another long-standing, unsolved question about MATE transporters. The following subsections discuss the literature exploring this question.

2.5.3 | Substrate binding site

As mentioned above, the structure of MATE proteins includes a central cavity which is formed between the N- and C- lobes, providing surfaces for ion and substrate binding within the N- and C-lobe cavities (Tanaka et al., 2021). The ion-binding site as well as the ion specificity, as discussed in the previous subsection, differs amongst MATE transporters. This subsection addresses the literature about the substrate-binding site of MATE transporters.

Structural comparisons and functional analyses have revealed that the substrate-binding site, just like the ion-binding site, differs amongst members of the MATE subfamilies (Zhao et al, 2021). That is, for some MATE transporter subfamilies including NorM-VC (from the NorM subfamily) and AtDTX-14 (from the eMATE subfamily), the substrate-binding site has been found within the C-lobe cavity, formed by negatively charged residues (Kusakizako et al., 2021); whilst for the DinF MATE subfamily, which includes DinF-BH and PfMATE, substrates have been found to bind within the N-lobe cavity (Lu et al.,

2013a; Tanaka et al., 2013). Contrastingly, NorM-NG complex structures found substrate-binding to occur on the extracellular loop, with the substrate-binding site in the central cavity (Lu et al., 2013a), although this proposition remains contested (Kusakizako et al., 2021).

In addition to the dissimilar locations of substrate-binding sites amongst MATE subfamilies, considerable dissimilarities have also been found with regards to the composition and configuration of amino acid residues at the substrate-binding site (Zhao et al., 2021). The subsections below provide a deeper look into the conformational mechanics of MATE transporters, where more sub familial differences are discussed with regard to the structure-function relationship of MATE transporters and their substrate extrusion cycle.

2.5.4 Conformations and possible mechanistics of MATE transporters

MATE transporters exist in two main conformations, an inward-facing (IF) and an outward-facing (OF) conformation (or IFC and OFC, respectively), and the switching from one conformation to another involves various transitions that generate intermediate conformations. This subsection discusses these conformations and what is known about the molecular-level changes that take place during the conformational switching that drive substrate efflux.

Several representatives from each MATE subfamily have been resolved experimentally, capturing the MATE transporters in various conformations. Hitherto, a total of 29 crystal structures from seven different organisms, with representatives from all three MATE subfamilies have been resolved and characterized: NorM-VC (He et al., 2010), NorM-NG (Lu et al., 2013a; Radchenko et al., 2015), PfMATE (Tanaka et al., 2013; Zakrzewska, 2019), ClbM (Mousa et al., 2016), CasMATE (Tanaka et al., 2017), AtDTX14 (Miyauchi et al., 2017), and DinF-BH (Lu et al., 2013b). Most of these crystal structures were captured in the outward-facing conformation (OFC), until recently, when the first crystal structure in the inward-facing conformation (IFC) was captured by Zakrzewska *et al.* (2019) from the archaeon, *Pyrococcus furiosus* (PfMATE) (PDB ID: 6FHZ), exhibiting an Λ-shaped architecture open to the intracellular side. Figure 2.4 below shows a ribbon representation of PfMATE in both distinct OF (left) and IF (right) conformations embedded within a membrane bilayer, with the N- and C-lobes depicted in pink and blue, respectively. It can be seen that in the OFC, the upper V tips of both lobes are found on the extracellular side, with the base on the cytoplasmic side, and vice versa for the IFC.

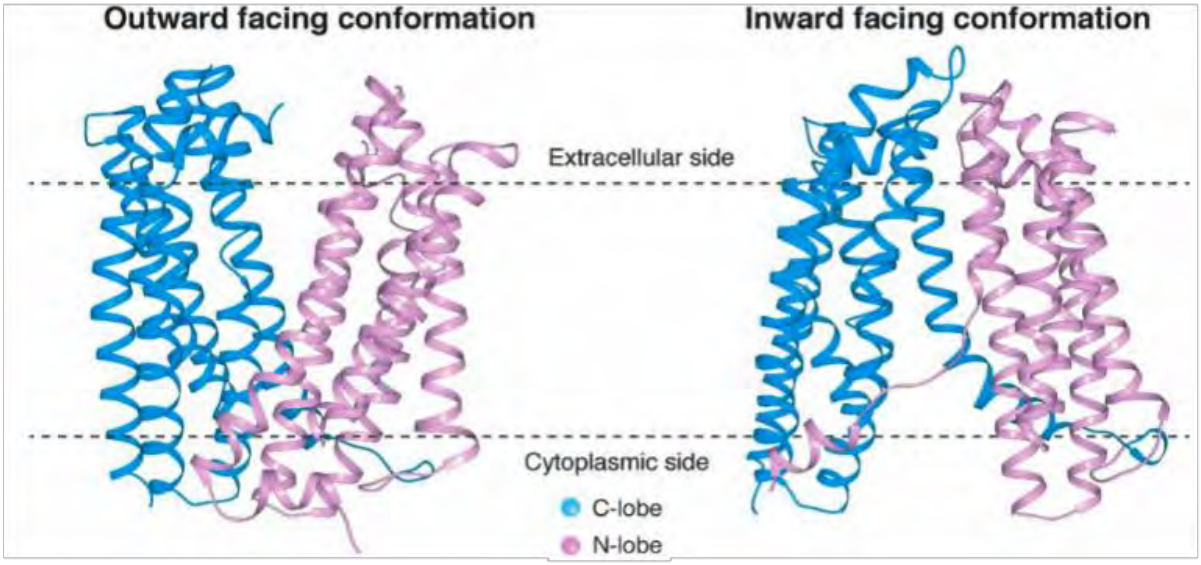


Figure 2.4 | Illustration of PfMATE in their distinct OFC (left) and IFC (right) (adapted from Zakrzewska *et al.*, 2019): MATE transporters span membrane bilayer, exhibiting a V-shaped architecture, with a central cavity open to the extracellular side in the OFC and an Λ-shaped architecture whose central cavity is open to the cytoplasmic side in the IFC. The N- and C-lobes are represented in pink and blue ribbons, respectively.

Table 2.1 lists the 29 crystal structures along with various information including their subfamily, crystal structure resolution, known transported substrates, etc, with the structures used in this study underlined. Of these structures, it should be noted that irrespective of bound ion, substrate, or inhibitors, all of them except for one (PDB ID: 6fhz) have been captured in the OFC, indicating that the OFC is a lower energy state, and is thus the favoured intermediate during crystal lattice formation (Claxton et al., 2021). In addition to the IFC and the OFC, the latest crystallography research complemented by biophysical analyses, report other intermediate conformations of these transporters within the membrane. Six MATE proteins (NorM-VC, NorM-NG, PfMATE (excluding 6FHZ), ClbM, CasMATE, and AtDTX14) in this table have been captured in the OFC, exhibiting a typical V-shape, with the N- and C-lobes facing the intracellular side, protruding away from each other towards opposite sides of the membrane, and a big central cavity open to the extracellular side, and sideways to the lipid bilayer. The structures (PDB ID: 4LZ6 and 4LZ9) from DinF-BH, however seem likely to be a possible intermediate conformation in which the hairpin of TM7-TM8 is observed to be detached from the C-lobe, and instead clustered with the N-lobe (Lu et al., 2013).

Table 2.1 | Structurally characterized MATE transporters and known transported antimicrobial substrates (adapted from Claxton et al., 2021): Shown are a list of 29 available MATE 3D crystal structures from 7 different organisms, representative of the three subfamilies along with the drug substrates/ligand(s) they are known to transport, their classifications, and structure resolution. Five MATE proteins – NorM-VC, NorM-Ng, DinF-Bh, VcmN, and ClbM-Ec – are bacterial, one MATE protein, PfMATE, is archaeal, and the remaining two – DTX-14-At and CasMATE – belong to the plant kingdom. The underlined structures were used in this study (PDB IDs: 6FHZ as the IFC structure and 3VVP as a complex with bound known drug).

Protein	Subfamily	Organism	Bound/Transported Substrates	PDB ID	Resolution (Å)	Ligand
NorM-Vc	NorM	*Vibrio cholerae*	Et, CFX, DXR, DNR, RBXL, R6G, NFX Refs. 48,51,75,83	3MKT 3MKU	3.65 4.20	 Rb^+
NorM-Ng	NorM	*Neisseria gonorrhoeae*	Et, TPP, R6G, CFX, NFX, AFh, NME, BE, BC Refs. 49,84,85	4HUK 4HUL 4HUM 4HUN 5C6P	3.60 3.80 3.50 3.60 3.00	TPP, monobody Cs^+, monobody ethidium, monobody R6G, monobody verapamil, monobody
PfMATE	DinF	*Pyrococcus furiosus*	Et, NFX, R6G Refs. 53,56	3VVN 3VVO 3VVP 3WBN 3VVR 3VVS 3W4T 6FHZ 6GWH 6HFB 4MLB	2.40 2.50 2.90 2.45 3.00 2.60 2.10 2.80 2.80 3.50 2.35	 Br-norfloxacin Cyclic peptide MaL6 Cyclic peptide MaD5 Cyclic peptide MaD3S Cs^+
DinF-Bh	DinF	*Bacillus halodurans*	CFX, TPP, R6G, Dequalinium, Et, NFX, TPA Refs. 50,54	4LZ6 4LZ9 5C6N 5C6O	3.20 3.70 3.00 3.00	 R6G verapamil
VcmN	DinF	*Vibrio cholerae*	CFX, Et, OFX, Hoechst 33342, Kan, NFX, Streptomycin Refs. 55,86	6IDP 6IDR 6IDS	2.20 2.50 2.80	
ClbM-Ec	DinF	*Escherichia coli*	Et, R6G, Precolibactin Refs. 66,67	4Z3N 4Z3P	2.70 3.30	 Rb^+
DTX14-At	Eukaryotic	*Arabidopsis thaliana*	NFX Ref. 79	5Y50	2.60	
CasMATE	Eukaryotic	*Camelina sativa*		5XJJ	2.90	

The poor resolution of crystal structures (above 3Å) of MATE transporters (as shown in Table 2.1) is partly responsible for the obscurities and discrepancies in the literature regarding their ion-coupled mechanism of efflux (Ficici et al., 2018). However, structures of PfMATEs as well as VcmN have been captured at a higher degree of resolution, making them ideal structures for studying the features of MATE transporters that result in MDR.

2.5.5 | Sub familial and intra familial contrasts and similarities amongst MATE transporters

Studies of MATE transporters from the three subfamilies – DinF, NorM, and eMATE – have revealed unexpected contrasts. Some of these – the ion specificities and binding sites, substrate binding sites, and nature of the central cavity, all of which impact their function – have been highlighted in the subsections above. This section summarises these differences and addresses important additional contrasts amongst MATE transporters of different subfamilies, but also within a subfamily.

Following multiple sequence alignments, bacterial MATE transporters are grouped into both the DinF and NorM subfamilies (Omote *et al.,* 2006; Tanaka *et al.,* 2017). Till now, 5 prokaryotic MATE crystal structures have been resolved (see Table 2.1), in which contrasts have been found with respect to the ion driving their function. These include one from the NorM-Vc subfamily (He *et al.,* 2010) driven simultaneously by inwardly-directed Na^+ and H^+ ion gradients (Jin *et al.,* 2014), two that are coupled to an inwardly-directed Na^+ ion gradient, NorM-Ng (Lu *et al.,* 2013b) and ClbM-Ec (Mousa *et al.,* 2016), and two that are coupled to an inwardly-directed H^+ ion gradient, PfMATE (Tanaka *et al.,* 2013) and DinF-Bh (Lu *et al.,* 2013a; Radchenko *et al.,* 2015).

MATE transporters in the OFC from members of different subfamilies have been resolved (with or without ions/substrates) in two different structural forms – a ‘bent’ and ‘straight’ form. It has been suggested that these structural variations of OF conformations are representative of distinct phases in the transport cycle of these transporters (Claxton, 2021).

From the DinF subfamily, two *apo* outward-open structures of PfMATE (archaeal) have been captured in these bent (PDB ID: 3VVO) and straight (PDB ID: 3VVN) forms – at 2.5Å and 2.4Å resolutions, respectively (Tanaka *et al.,* 2013). The bent and straight conformations obtained at pH 6.0 and pH 7.0-8.0, respectively, describe the structure of the TM1 helix within the N-lobe (where the substrate binding site is located). In the bent form, TM1 exhibits a kink at Pro 26 and Gly 30, whereby it bends towards the side of the TM2 helix, whereas on the other hand, it remains as a single, straight helix in the straight form (Tanaka *et al.,* 2013). In addition, Tanaka *et al.* (2013) found that in the bent form, the regions of

TM5 and TM6 helices exposed to the extracellular side of the membrane as well as the loop that links the two helices, have a slight outward shift, resulting in a wider open-outward conformation. As such, in the bent conformation, the N-lobe cavity primary comprising TM1 and TM2, collapses, resulting in the release of the bound substrate into the extracellular space. In summary, the findings of this study, through both structural investigation and functional analysis, proposed that, following the structural transition from the inward-open to outward-open conformation, deprotonation occurs promptly and in this state, the substrate alone is bound to the transporter. An Asp 41 is then protonated, resulting in the sheltering of its side chain into the surrounding hydrophobic environment; this results in the bending of TM1 at Pro 26, collapsing the the N-lobe cavity, which in turn releases the substrate into the extracellular surroundings. As a final step, the protonated transporter, now with its empty substrate-binding site, undergoes a structural transition back to the inward-open conformation (Tanaka *et al.,* 2013).

In the eukaryotic MATEs, CasMATE and AtDTX14, as well as human MATE proteins, the substrate binding site found within the C-lobe cavity contains a conserved Glu within TM7, which has been identified as a proton acceptor, and is also involved in an important interaction with an Asp within TM10 (Kusakizako et al., 2019). The interaction between these amino acids is responsible for inducing a structural change from a straight-TM7 to a bent-TM7 helix, which results in the release of the substrate (found at the C-lobe in the outward-open form) from the substrate-binding site. These findings show a contrast in comparison to members of the in bacterial DinF group and that of PfMATE, where the substrate-binding site is found within the N-lobe, and significant conformational transitions occur in TM1 as opposed to TM7, resulting in the structural change from a straight-TM1 to a bent TM1 helix, followed by the release of the substrate (Kusakizako *et al.,* 2019).

The (eukaryotic) MATE transporters of the tobacco species, *Nicotiana tabacum*, has two MATE transporters, NtMATE1 and NtMATE2, whose amino acid sequences are 95% similar (Shitan and Yazaki, 2007; Shoji *et al.,* 2009), but only have a sequence identity of 35% to the above mentioned plant MATE transporters, CasMATE and AtDTX14 (Tanaka *et al.,* 2021). A more recent study reported the structure of NtMATE2 in the bent and straight forms of their OFC (Tanaka *et al.,* 2021). However, the conformational changes were found to be specific to NtMATE2, in that an unexpected inclining movement of TM7 observed involved the conserved Glu in TM7, but an Asn403 was found in TM10, instead of the Asp in TM10 found in previous studies to be conserved in the other eukaryotic MATEs.
The above studies suggest varying conformational mechanisms even within the same subfamily of MATE transporters and amongst MATE subfamilies. For some MATE transporters (PfMATE from the DinF subfamily and bacterial DinF), the substrate-binding site is located within the N-lobe, and TM1 rearrangements (and an Asp 41 residue) are important in the MATE transporter cycle. Conversely, for some eMATE transporters, the substrate-binding site is found within the C-lobe cavity, and structural

rearrangements involving a conserved Glu in TM7 (CasMATE, AtDTX14, human MATEs and NtMATE2) and a conserved Asp in TM10 (CasMATE, AtDTX14 and human MATEs) or Asn in TM10 (NtMATE2) are important during the transport cycle of these MATE transporters. These sub familial and intra familial contrasts contribute to the perplexities about the molecular-level mechanisms of MATE transporters.

Although multiple sequence alignments revealed no consensus sequence amongst MATE transporters subfamilies (with roughly 40% sequence similarity), various TM domains as well as inter-TM linkers were found to be more conserved than other regions (Chen *et al.,* 2015). These regions are located within the perimeters of TM1 and TM7 along with the extracellular loops linking TM1 with TM2 and TM7 with TM8, the cytoplasmic loops that connect TM2 with TM3 and TM8 with TM9, and finally in the loops that link TM4 with TM5 and TM10 with TM11 (Otsuka et al., 2005). It is likely that these conserved domains contain sites that are crucial to the function of MATE transporters (Chen *et al.,* 2015).

Several authors have highlighted the need for better-resolved structures and methods to identify subtle but vital structural and dynamic details of the underlying molecular-level mechanism of MATE transporters (Tanaka *et al.,* 2013; Tanaka *et al.,* 2017; Ficici *et al.,* 2018; Kusakizako *et al.,* 2019; Krah *et al.,* 2020; Claxton *et al.,* 2021; Tanaka *et al.,* 2021). While monumental progress has been made in the understanding of these membrane efflux pumps, more investigations are required to fill in the gaps and gain a better understanding of their mechanism and dynamics in the scheme of cellular control. This will shed more light on the contribution of MATE transporters and other drug efflux pumps to the widespread problem of MDR.

2.5.6 | Proposed mechanistic model of MATE EPs

Mechanistically, MATE transporters have been proposed to fall under an alternating-access model with rocker-switch type movements and function as secondary active antiporters. The resolved 3D crystal structures of MATE transporters in their two distinct conformations (IFC and OFC) in their *apo* and complex forms, along with those of intermediary conformations, as well as studies on ion-specificity provide support for this model.

Figure 2.5 provides a schematic summary of the discussed structure-function dynamics of MATE transporters as antiporters under the alternating access rocker-switch model.

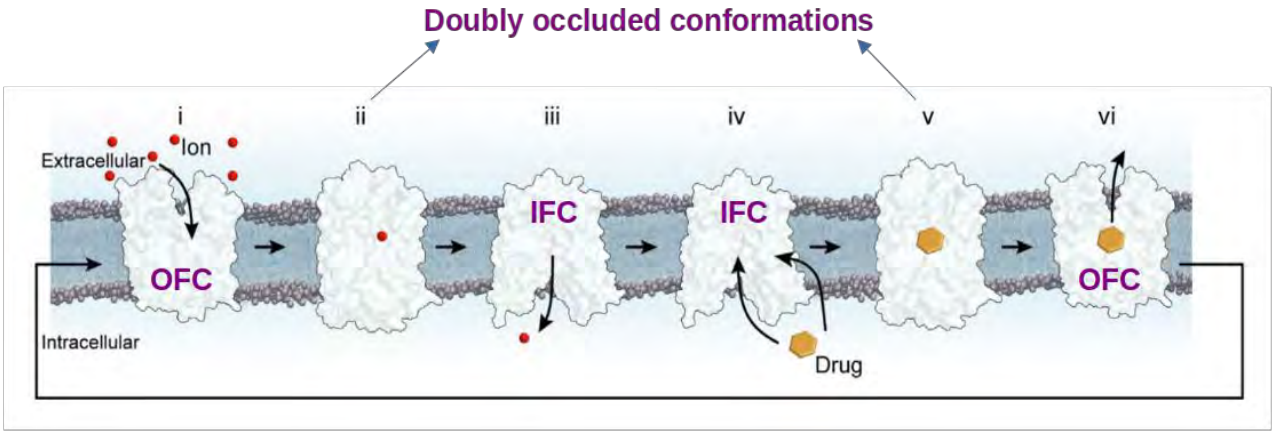


Figure 2.5 | Proposed mechanism for the substrate-efflux transport cycle in MATE transporters (adapted from Claxton *et al.*, 2021): In the OFC (i), the ion and substrate binding sites are unoccupied; an inwardly-directed (H^+/Na^+) ion gradient results in the binding of an ion from the extracellular space, inducing the collapse of a TM helix forming an intermediate doubly occluded conformation (ii); rocker-switch movements of the two TM bundles (N- and C- lobes) result in the conversion to the IFC (iii) and the release of the ion into the intracellular space; the drug molecule binds to the exposed substrate-binding site (iv), resulting in the collapse of a TM helix and the (re)formation of an intermediate doubly occluded conformation (v); rocker-switch movements cause structural rearrangements forming the OFC (v), releasing the drug to the extracellular space and concluding the transport cycle.

While the above mechanism – a rocker-switch alternating access antiport model – is now largely agreed upon, as discussed in the above sections, the molecular-level understanding of the mechanism remains questionable. It is known that conformational changes are brought about by the rearrangements of some of the TM helices, which involve specific amino acid residues, which in turn are responsible for the expulsion of bound substrates. However, at the current level or resolution, and with only one available structure in the IFC, subtle but important structural changes underlying the molecular mode of action of members of MATE transporters have yet to be elucidated in more detail. As such, the need for alternative investigatory methods, and better quality of resolved structures is highlighted to target MATE transporters as antimicrobial targets in the future, offering a possible solution to MDR.

CHAPTER 3: MOLECULAR DOCKING

3.1 | Introduction

Molecular docking was applied in this study to identify potentially optimal inhibitory compounds from the SANCDB via virtual screening against (both IF and OF conformations of) the MATE transporter protein. This chapter provides a brief introduction to molecular docking, followed by the methodology used and the results generated.

3.2 | Molecular Docking and drug discovery

Drug discovery aims to identify chemical compounds with the capacity to produce an intended biological outcome (Liao *et al*., 2013). The (experimental) practice of drug discovery and development is costly both in terms of time and money investments (Brown and Superti-Furga, 2003; Workman, 2003). However, the development of computer-aided drug design (CADD) methods has been monumental in the pharmaceutical industry to address both these issues (Gomeni *et al.,* 2001; Veselovsky and Ivanov, 2003). CADD allows for narrowing down scientists' attention to the compounds that are more likely to produce the desired biological effect, thereby downsizing synthetic and biological experimental resources (Zhang, 2011). These compounds exhibiting the sought-after effects are referred to as 'lead' compounds (Koehn and Carter, 2005; Khazir *et al.,* 2013), whose properties are assessed and later optimized for 'drug-likeness' (discussed below).

CADD comprises of two chief approaches – Structure-based drug design (SBDD) and ligand-based drug design (LBDD). This study utilized a SBDD approach, which makes use of the 3D structural data of the macromolecular target (or receptor) – obtained experimentally or via in silico homology modelling – providing information on various structural attributes of the receptor. This allows for the identification and/or design of potential optimal ligands against the target receptor (Mandal *et al.,* 2009; Blaney*,* 2012).

In response to the necessity to solve problems of structural molecular biology and structure-based drug discovery, and fueled by an increase in computational power and the ready access to compound databases, molecular docking has evolved into a fundamental SBDD (CADD) technique, widely used by the pharmaceutical industry (Berman *et al.*, 2000; Irwin and Shoichet, 2005; Hu *et al.*, 2005; Morris and Lim-Wilby, 2008; Ferreira *et al*., 2015).

Molecular docking is an appealing tool that enables the study of the bio-molecular interactions between a drug molecule (ligand) and a protein (receptor), amongst other types of interactions and applications (Rohs *et al.*, 2005; Guedes *et al.*, 2014). The current study employed this type of docking, known as ligand-protein docking. Ligand-protein docking serves as a means of emulating the atomic-level interactions occurring between the ligand and the protein, shedding light not only on the behavior of the ligand at the active site of the protein but also on important biochemical processes (McConkey *et al.*, 2000).

There are several applications of molecular docking in the drug discovery process, including structure-activity investigations, virtual screening, and optimization of lead compounds, amongst several others (Morris and Lim-Wilby, 2008). Virtual screening allows in silico sampling of compounds to observe their behavior and whether they produce the desired biological effect; thereby, the list of compounds thereafter that require (physical) experimental screening is drastically reduced (Klebe, 2006; Sousa *et al.,* 2010). In turn, the time and costs of required experimental resources are minimised. This study applied molecular docking to identify possible lead inhibitory compounds against MATE transporters via virtual screening against the SANCDB natural compound library.

Using software, ligand-protein docking aims to provide insight into and make both structural and energetic predictions by identifying various binding modes and their corresponding binding affinities, respectively, about the atomic level interactions taking place (Morris and Lim-Wilby, 2008). This is achieved by positioning the ligand in close proximity to the binding site of the target receptor, allowing the formation of a complex via non-covalent bonding between the two structures (Rohs *et al.*, 2005; Guedes *et al.*, 2014), with the objective of obtaining a stable ligand-protein complex of optimal conformation with the lowest binding free energy (Agarwal *et al.,* 2015; Dar and Mir, 2017).

3.3 | Theory of Molecular Docking

Molecular docking entails two fundamental steps. The first step involves making predictions of the ligand ‘pose’, which refers to the conformation, position, and spatial orientation of the ligand with respect to the binding site of the target protein, following which, the second step evaluates the binding affinities for each predicted pose (Meng *et al.,* 2011). Prior knowledge about the location of the docking site correlates to a considerably higher docking efficiency (Meng *et al.,* 2011). This first step entails ‘sampling’ various poses of the ligand and the receptor binding site, which are subsequently ranked via a ‘scoring function’. The scoring function is designed to accurately and time-effectively determine feasible poses from those that would be impractical, doing so by making estimations of the binding

affinities between the ligand and protein, based on either force-fields, prior knowledge, and/or empirical premise, and simplifications (Kitchen *et al.,* 2004; Meng *et al.,* 2011). The scoring function(s) varies across different docking software. For example, this study used Autodock Vina for docking, which employs a hybrid scoring function based on empirical data and prior knowledge (Trott and Olson, 2010).

The sampling algorithm considers translational, rotational, and conformational degrees of freedom for both the ligand and protein; this would result in numerous possibilities of binding modes, becoming too computationally taxing. As such, various sampling algorithms are available, each with its own advantages and limitations (Meng *et al.,* 2011). Essentially, it is expected that the binding mode generated by the sampling algorithms is equivalent to those achieved experimentally and that the scoring function would rank it at the top of the various outputs (Meng *et al.,* 2011). There are several different docking software packages available, each employing various scoring functions and sampling algorithms.

3.4 | Natural products and drug discovery and design/development

Having evolved a range of diverse chemical and biological features, plants, animals, and marine life are a rich source of natural products (NPs), typically exhibiting stereochemical complexity, containing a range of functional groups, and being able to form highly specific interactions with biological targets. As such, NPs are attractive structural templates for drug discovery and design (Khazir *et al.,* 2013). Ancient records dating back nearly 5000 years ago document the successful use of NPs in treating a multiplicity of diseases (Dias *et al*., 2012; Cragg and Newman, 2013). NPs are thus of significant importance to the field drug discovery and are especially promising in treating cancers and infectious disease (Atanasov *et al.,* 2015; Harvey *et al.*, 2015). This study screened NPs from SANCDB - a South African natural compound database (Hatherley *et. al.,* 2015) – for possible inhibitory action against MATE transporters (in both their *apo*- IFC and OFC).

3.5 | Lipinski's rule of five, 'drug-likeness', and NPs

Lipinski's 'rule-of-five' (Lipinski *et al.,* 2001), commonly referred to as 'Lipinski's rule of drug-likeliness' denotes specifications relating to the important properties – mainly solubility and permeability – of prospective drug molecules known to influence their (oral) bioavailability (Gu *et al.,* 2013; Valli *et al.,* 2013; Ntie-Kang *et al.,* 2014). These properties termed ADME/Tox (absorption, distribution, metabolic, elimination, and toxicity), are imperative measures used to assess the 'drug-likeliness' of discovered (lead) compounds (Zhang and Wilkinson, 2007). Drug compounds with

desirable ADME/Tox properties are referred to as drug-like. Albeit serving as substantial guidelines in the case of bioavailability of orally administered small molecules during drug design, these rules have some limitations, for example, they fail to address NPs and semisynthetic NP drugs, consisting of over one-third of small-molecule drugs on the drug market (Zhang, 2011).

Whilst over 60% of SANCDB compounds are not in violation of any of Lipinski's rules and about 80% violate no more than one of these rules (Hatherley *et. al.,* 2015), they are attractive NPs worth examining as potential lead drug compounds. As such, prior to the utilization of the SANCDB NPs for the purposes of this study, they were not tested via Lipinski's rule of drug-likeness, as the compounds in the SANCDB have already been tested against this criteria.

3.6 | Methodology

Docking was performed using AutoDock Vina. The figure below, Figure 3.1 (adapted from Morris and Lim-Wilby, 2008), shows a summary flowchart of the docking process, which was utilized in this study.

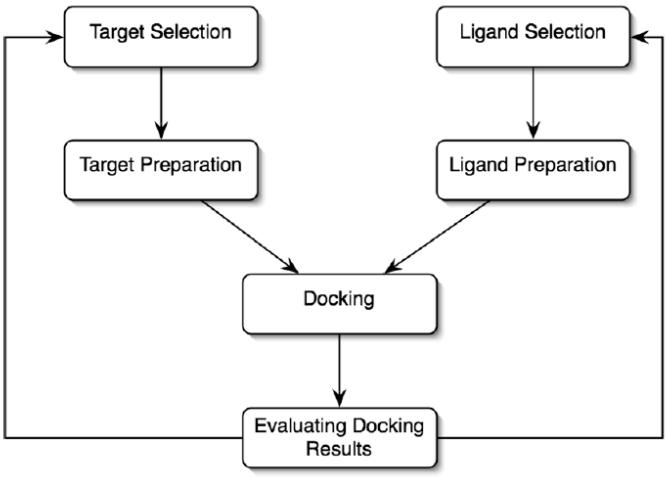


Figure 3.1 | Basic flowchart of the molecular docking process (Morris and Lim-Wilby, 2008): 3D structures of both the ligands and target protein are required for docking. These ligands and target protein are prepared prior to the docking process, following which they undergo docking, generating hit compounds. Following the basic methodology shown in the figure above, the process used in this study is detailed below.

3.6.1 | Target selection

The MATE transporter protein, in both the IFC and the OCF were the target receptors of this study. The only available crystal structure of the IFC (PDB ID: 6FHZ) from the archaea *Pyrococcus furiosus* was downloaded from the RCSB website, as the target IFC MATE receptor.

For the OCF, homology modelling was employed using PRIMO (PRotein Interactive MOdeling, Hatherley et al., 2016) pipeline, and the top ranked model was selected. The model fit the criteria for as a good model based on inspection of the Ramachandaran plot (see Figure 3.2), which showed a higher percentage of residues (98.1 %) fell in the most favoured regions, whilst the remaining residues (1.9 %) fell in the additional allowed regions. This model thus served as the target OFC MATE receptor.

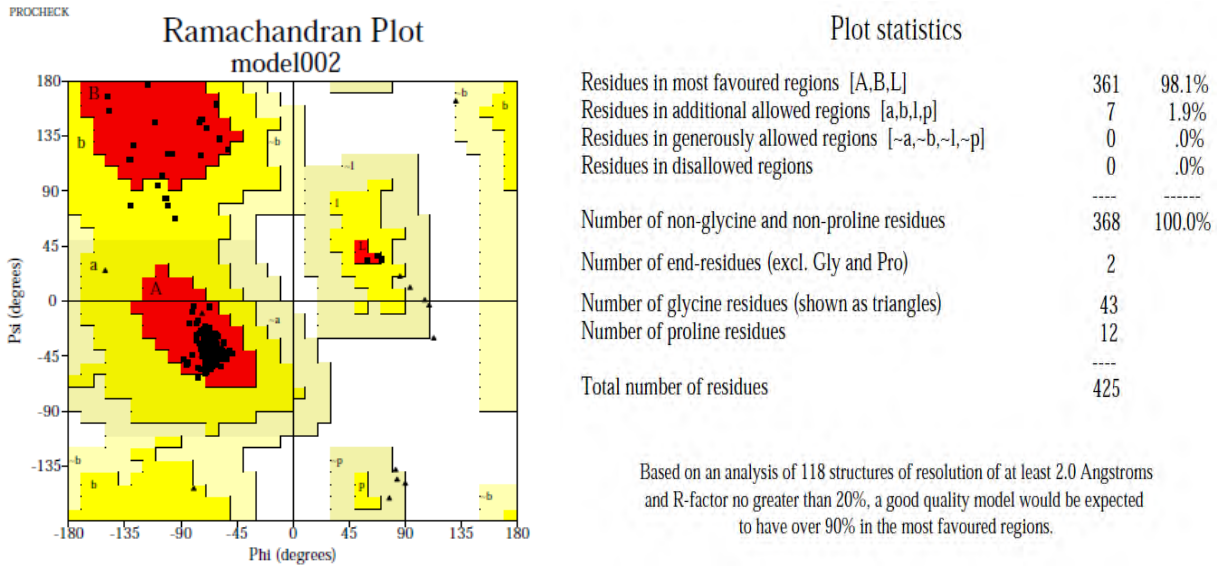


Figure 3.2 | Ramachandaran plot of the selected homology model for the target MATE OFC protein: The plot indicates the model selected is a successful model given that 98.1% of the residues fell within the most favoured regions, whilst the remaining 1.9% fell in the additional allowed residues.

A complex crystal structure of the OFC MATE protein, also from the archaea *Pyrococcus furiosus,* with a derivative of a known drug (norfloxacin) bound to it was available on the RCSB website (PDB ID: 3VVP, see Figure 3.3). It served two purposes – firstly providing information on the identified target binding site, and secondly acting as a control OFC complex structure during the MD simulations performed in the next chapter.

3.6.2 | Target and ligand preparation

Once the target IFC and OFC starting structures were selected, they were prepared for docking, along with the SANCDB ligands, using the available AutoDock Tools' python scripts. While MATE transporters are membrane-bound proteins, the area demarcated by the grid box for docking consisted of only the regions exposed to either side of the membrane (depending on whether the IFC or OFC was the target protein). The docking method used was a blind docking approach whereby despite knowing the binding site on the target protein, all exposed regions of the protein were demarcated by the grid box. This was done to increase the accuracy of selecting optimum ligands that freely bound to the protein at the correct site during the post-docking evaluation process.

The Orientations of Proteins in Membranes (OPM) database (Lomize *et al.,* 2012) was used to visualize the way in which the MATE transporter is embedded within the membrane bilayer, allowing for an accurate demarcation of the area within the grid box. The figure below shows the selected IFC structure embedded in a lipid bilayer, depicting how it would be found in its native (biological) environment. Although not depicted in Figure 3.3, the central cavity of the remains open to the intracellular/cytoplasmic side for substrate binding (Zakrewska *et al.,* 2019), and was included within the grid box.

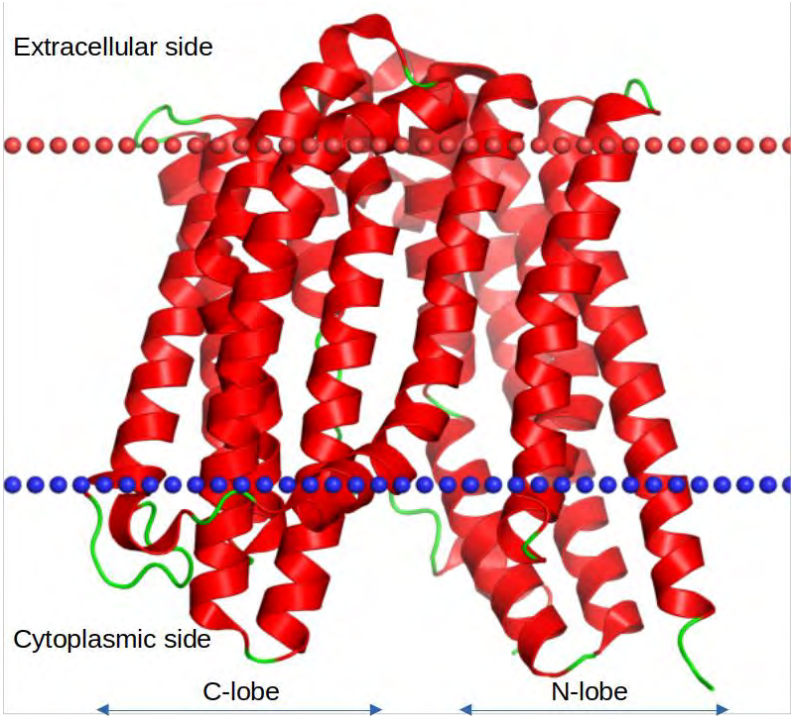


Figure 3.3 | Visualization of the IFC of the MATE transporter (PDB ID: 6FHZ) embedded within a membrane bilayer generated via the OPM server: the N- and C- lobes meet within the membrane to form an Λ-shape with a central cavity open to the cytoplasm for substrate binding.

The active site area was easy to define given the available crystal structure complex (PDB ID: 3VVP), offering the benefit of increasing the docking efficiency. The active site is located on the N-lobe of PfMATE transporter proteins, as shown in Figure 3.4.

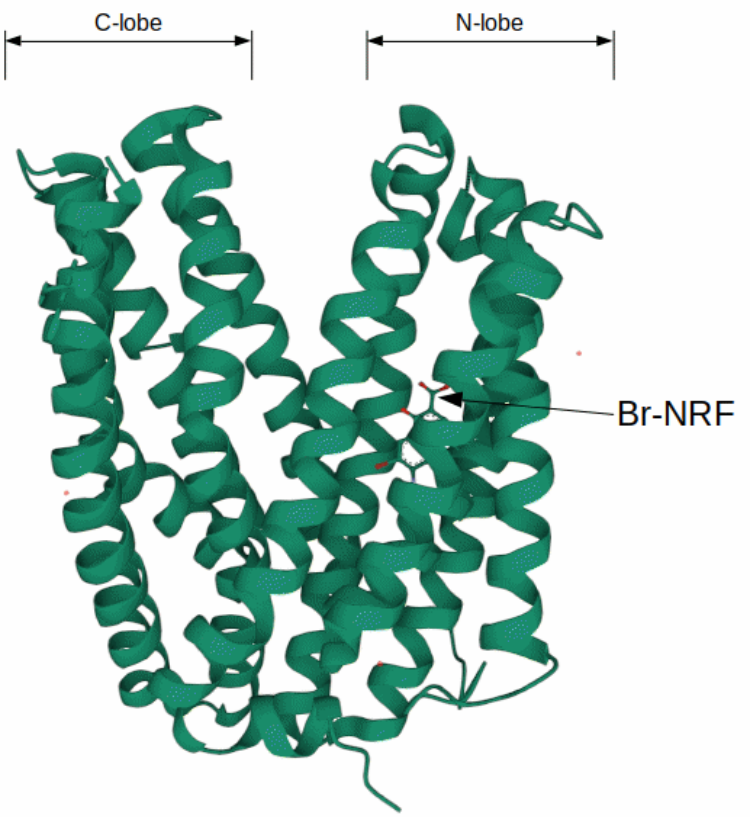


Figure 3.4 | MATE OFC in complex with Br-NRF (PDB ID: 3VVP): A derivative of the drug norfloxacin is bound to the active site of the MATE transporter (ribbon representation) found at the N-lobe of the protein.

The availability of the 3D complex structure (3VVP) provided information on the location of the docking site within the N-lobe of the MATE transporter, offering the benefit of increasing the docking efficiency.

3.6.3 | Docking, hit validation, and selection of docked complexes

With the ligands and target proteins prepared, docking was performed using Autodock Vina, generating several poses for each ligand, ranked according to their binding affinity energies. Figure 3.5 portrays the individual steps taking place during the docking process.

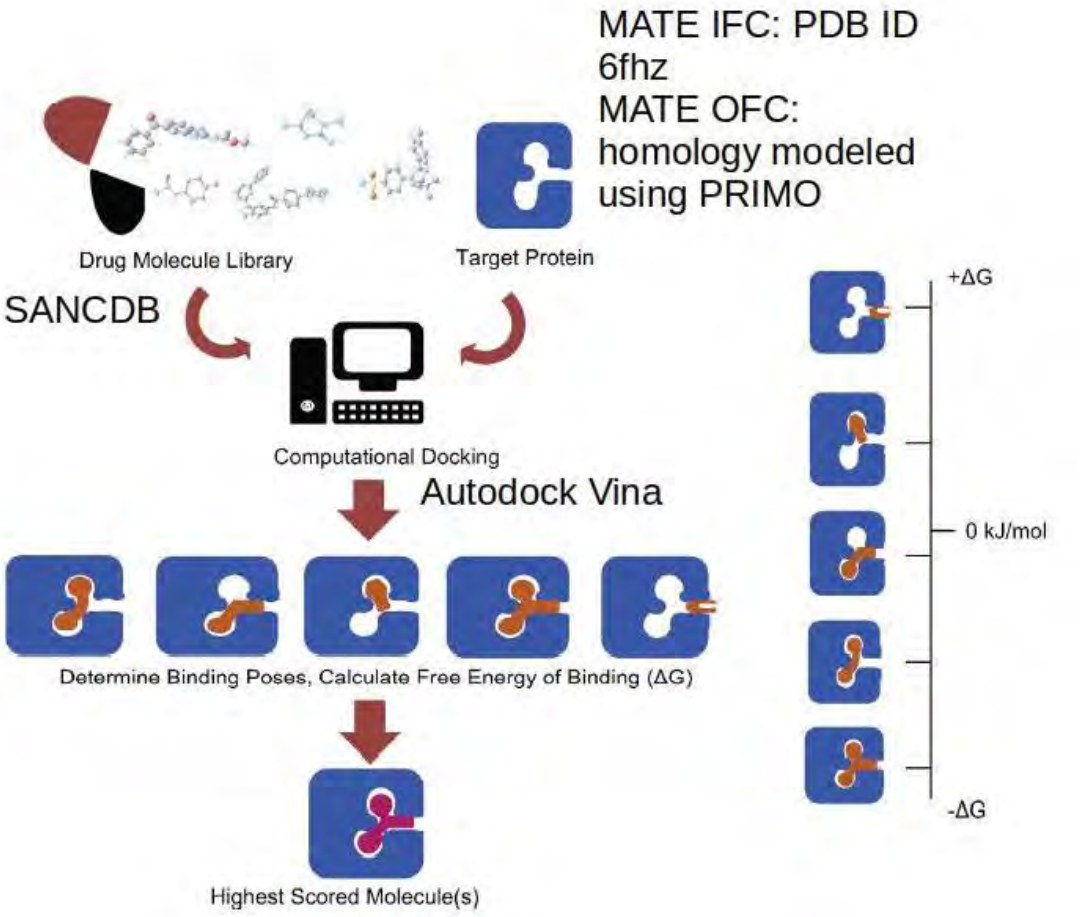


Figure 3.5 | Steps performed during the docking process (adapted from Issa et al., 2019): ligands were obtained from the SANCDB and prepared for docking against the obtained and prepared 3D structures of the target IFC- and OFC-MATE target proteins. Docking was carried out using Autodock Vina which involved determination of feasible poses and calculation of their binding free energies. The final determined poses were then ranked according to their corresponding binding free energies.

Next, a python script was written to rank the top (lowest energy) poses in order. The poses were then validated by visual inspection against the available complex structure (Figure 3.4) to select the compounds of lowest energies that were also bound to the correct binding site on the N-lobe of the MATE transporter.

To ensure validation accuracy and selection of potentially optimal inhibitory ligands, site-specific docking was, however, not used; instead, the grid box demarcated a slightly larger docking site area around the active site, falling within the regions of the protein that would be exposed out of the membrane. As such, the ligands were allowed to freely bind to their preferred region, allowing for more accurate selections (of potential optimal inhibitors). This was apparent upon a visual analysis of the generated poses against the target IFC- and OFC-MATE transporters, whereby those ligands that preferably bound to regions other than the known binding site were weeded out.

The ligands were then filtered via a written python script to select the best two ligands with lowest energies against both the target IFC- and OFC-MATE proteins. Complexes of the MATE transporters with these bound ligands were generated for the second objective via MD simulations in the following chapter.

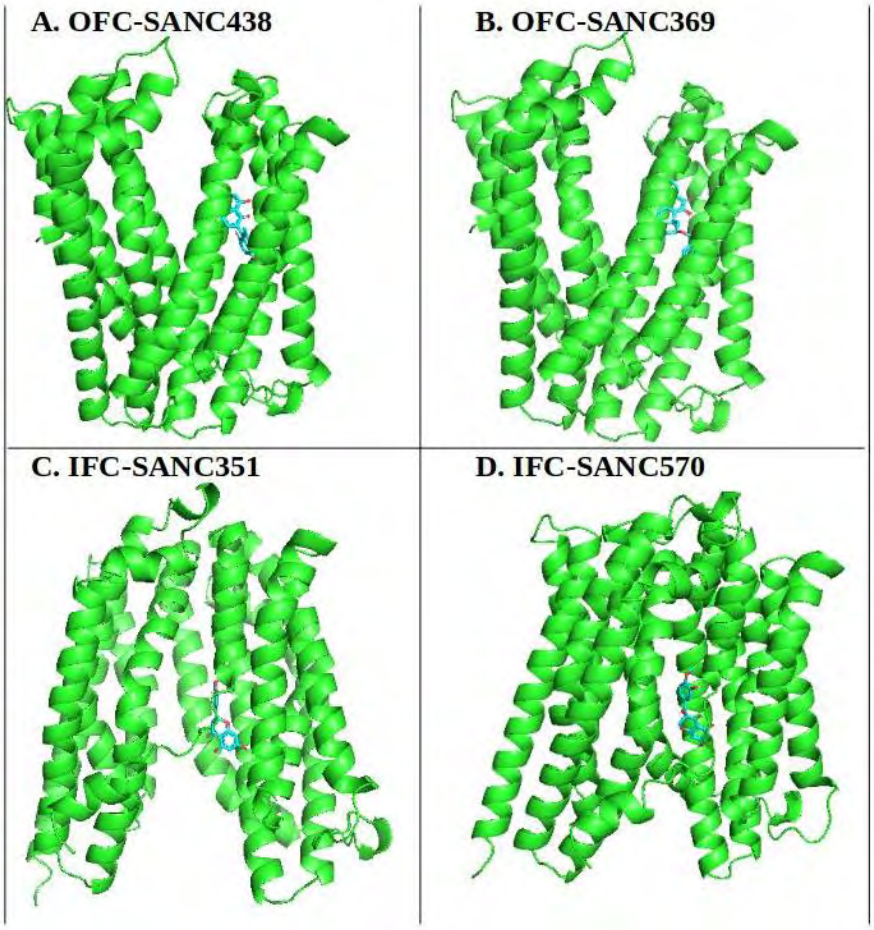


Figure 3.6 | Selected docked IFC and OFC complexes for MD simulations: A: OFC-SANC438, B: OFC-SANC369, C: IFC-SANC351 and D: IFC-SANC570, shown in ribbon form, with bound SANCDB ligands at the N-lobe.

Figure 3.6 shows the ligand molecules of the above complexes selected for MD simulations with their molecular weights, including that of the control downloaded complex (PDB ID: 3VVP).

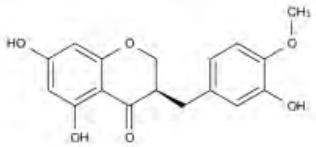
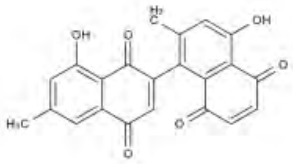
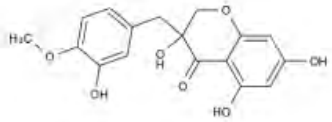
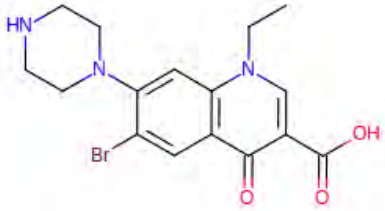
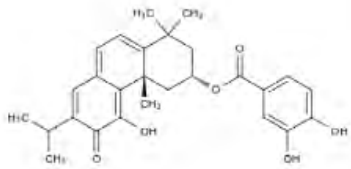
Ligand structures from selected docked complexes	
SANC0351 (from docked MATE IFC complex), [5,7-Dihydroxy-3-(3'-hydroxy-4'-methoxybenzyl)-4-chromanone, MW = 316.3 g/mol] 	SANC00438 (from docked MATE OFC complex), [Neodiospyrin, MW = 374.34 g/mol] 
SANC00570 (from docked MATE IFC complex), [3,5,7-Trihydroxy-3-(3'-hydroxy-4'methoxybenzyl)-4-chromanone, MW = 332.3 g/mol] 	NRF (from MATE OFC complex, PDB ID: 3VVP), [6-bromo-1-ethyl-4-oxo-7-(piperazin-1-yl)-1,4-dihydroquinoline-3-carboxylic acid, MW = 380.24 g/mol] 
SANC00369 (from docked MATE OFC complex), [Parviflorone F, MW = 450.52 g/mol] 	

Figure 3.7 | Molecular structures of the SANCDB ligands from the selected docked complexes and the ligand of the control OFC (NRF) with their molecular weights

MD simulations were performed on the selected docked complexes shown in Figure 3.7, as well as on the control OFC complex structure, which will be covered in the next chapter.

CHAPTER 4: MOLECULAR DYNAMICS SIMULATIONS

4.1 | Introduction

This study aimed to explore the family of MDR membrane proteins, MATE transporters, as potential antimicrobial drug targets. As mentioned previously, proteins do not remain static but are in constant motion under physiological conditions to perform their functions. Investigating a protein for drug-targetability, thus, necessitates a deep understanding of its dynamic nature within its natural environmental conditions, where MD simulations become a valuable technique.

Membrane proteins, comprising drug transporters, make up a vast and significant class of proteins to which the MATE transporters belong. Therefore, it is crucial to gain knowledge of their structure-function-dynamics relationships at a molecular level, in order to achieve some degree of cellular control. However, due to the difficulty of conducting experimental studies with proteins within membrane environments, these proteins remain relatively poorly understood (Byrne and Iwata, 2002; Bowie, 2004). As such, reaching past the static image of single protein structures, however, has proven to be quite difficult, although numerous techniques, including NMR relaxation, fluorescence spectroscopy, or time-resolved X-ray crystallography – which offer insight into the intrinsic conformational flexibility – have been developed. Despite this plethora of available experimental techniques, their main limitation is that they fail to determine spatio-temporal details at the nano- to microsecond timescales – in turn failing to provide information on the possible conformational space a protein (or other macromolecule) can occupy *in vivo*. Notably, specifics about the pathways amongst different known conformations – vital to protein function – frequently remain ambiguous.

Computer simulations have been used to complement experimental methods, illuminating the atomic-level function and dynamics of biological molecules, especially in the case of membrane proteins. Of these in silico techniques, MD simulations are a popular technique routinely used in the drug design and development industry. As such, this study employed MD simulations, which allowed for the simulation of the transmembrane MATE proteins within a membrane environment, providing insight into their probable behaviour *in vivo*. This section discusses the theory behind MD simulations, followed by the methodology used in this study, and then presents and discusses the results generated. To finish off this chapter, the limitations of this study are highlighted along with the corresponding future work.

4.2 | Brief theory of MD simulations

MD simulations are an in silico technique, popularly and effectively used for CADD. They bridge the structure-function-dynamics gap in the study of macromolecules, such as proteins that are highly dynamic in nature. MD simulations solve Newton's laws of motion for systems comprising interacting molecules, providing insight into the intrinsic physical motions of atoms and molecules – that are key to their mechanism of action – generating a corresponding trajectory over the simulation period (Liu *et al.,* 2018). In the current study, this system consists of MATE transporters (in both their *apo* and complex or ligand-bound forms) spanning a lipid bilayer environment, surrounded by water molecules and neutralizing ions – discussed in the Methodology subsection 4.3.2. Furthermore, MD simulations also utilize a molecular mechanics (MM) approach to calculate the potential energies of structural atoms in relation to their geometries, from which several 'force-fields' are generated (Liu *et al.,* 2018). Subsection 4.3.3 on energy minimization (under the methodology section) discusses the concept of force-fields in more detail. In contrast to experimental methods, which offer a more static view, MD simulations provide a dynamic depiction of proteins or peptides including folding, misfolding, and aggregation, as well as structural changes resulting from environmental variables (pH, temperature, and residue mutation) (Santini and Derreumaux, 2004; Campos *et al.,* 2010; Chen *et al.,* 2014). As such, the applications of MD extend far and wide, such as in the study of pathophysiology in disease as well drug discovery and design (Salo-Ahen *et al.,* 2020). The following methodology section describes the process of performing MD simulations in detail.

4.3 | Methodology

4.3.1 | Starting structures for MD simulations

In the previous section, molecular docking studies of MATE transporters in both their IFC and OFC was performed, where virtual screening of SANCDB compounds against these MATE proteins was carried out, following which several docked poses for each ligand were generated. In this section, MD simulations were performed on the selected docked IFC and OFC MATE complexes (following docking analyses), both the MATE apo-IFC and apo-OFC proteins, as well as on an available MATE OFC transporter in complex with a derivative of the known drug, norfloxacin, a fluoroquinolone antibiotic. These (seven) structures are listed below:

MATE IFC structures:

a) MATE apo-IFC

b) MATE IFC-SANC00351 docked complex

c) MATE IFC-SANC00570 docked complex

MATE OFC structures:

d) MATE apo-OFC (homology model)

e) MATE OFC-SANC00369 docked complex

f) MATE OFC-SANC00438 docked complex

g) MATE OFC-NRF complex (PDB ID: 3VVP) as a control

4.3.2 | Generation of a membrane environment for simulations

The main objective for the use of MD simulations in this study was to simulate the possible behaviour of MATE transporters under physiological conditions (or *in vivo*). Given that they span the entire bilayer, this largely includes their association with the lipid membrane bilayer, where they perform their function of cellular export of xenobiotics, including drug molecules. Their structural and dynamic nature allows them to act as secondary antiporters; they continually undergo cyclic conformational transitions and various dynamic motions driven by an inwardly directed ion gradient, thereby transporting their substrates across the lipid bilayer. Simulating larger and/or more complex systems demands more time and computational power. Even though simulation software can now handle simulating large and/or complex biological systems, there is always a trade-off between the system size and/or complexity and constraints of time and computational power. As such, despite performing virtual screening of all the SANCDB compounds against the MATE structures, MD simulations were only performed on a few selected docked complexes.

MATE transporters span the lipid membrane bilayer *in vivo*, where they function as transporters, moving (drug) substrates out of the cell, undergoing dynamic motions and conformational changes driven by an inwardly directed ion gradient. The first step thus involved generating a membrane environment with the MATE protein spanning a lipid membrane in order to simulate (or mimic) their conditions *in vivo*. This is an important step because their association with the membrane bilayer is known to influence their behaviour and function.

This study used the web-based CHARMM-GUI membrane builder (Wu *et al.,* 2014) to embed the above-listed MATE structures into a lipid bilayer. Amongst the parametrized lipids, POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) lipids were selected to generate a homogenous-lipid bilayer.

While biological (cellular and organelle) membranes are composed of a diversity of lipid types in varying ratios, POPC lipids are known to be the most plenteous lipid type making up lipid bilayers of the different cellular components; they are composed of mixed saturated/unsaturated-chain lipid chains (Petrache, 2012). In addition, POPC lipids have been well studied with available experimental data and have been widely used in performing MD simulations of membrane transporters.

CHARMM-GUI membrane builder is largely automated, simplifying the process of generating a membrane system in a step-by-step manner. Jo *et al.* (2009) provides details on this process, which is shown in Figure 4.1. Briefly, the steps involved are listed below:

1) reading in the starting structure pdb files (structures listed above).
2) changing the protein orientation – the option to use the OPM (Orientations of Proteins in Membranes) database (Lomize *et al.,* 2006) was selected for protein orientation.
3) determination of system size – where parameters – including the lipid type (POPC was selected), the number of lipids in the upper and lower bilayer leaflets, shape of the generated system, and finally the thickness of water surrounding the system – are specified.
4) generation of components including the lipid bilayer (homogenous POPC), surrounding water (TIP3P water model), and neutralizing ions.
5) assembly of the entire system
6) equilibration of the system

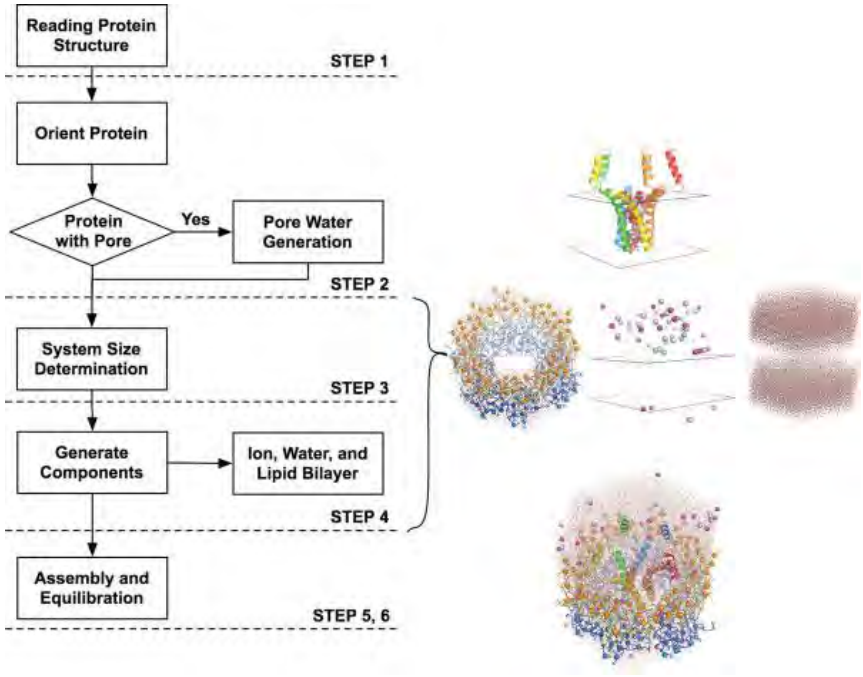


Figure 4.1 | Summary flowchart of the steps followed by the CHARMM-GUI membrane builder (adapted from Jo *et al.*, 2009): Using the web-based online tool, a homogenous POPC lipid membrane bilayer system was generated for the MATE transporter structures prior to MD simulations, according to the steps shown in the flowchart.

A representation of the entire system with all the components generated by the CHARMM-GUI membrane builder is shown in Figure 4.2 and was generated for all of the seven MATE structures listed above. As annotated in the figure, the systems generated for each of the structures listed in Section 4.3.1, comprise the MATE transporter (ribbon representation) spanning the homogeneous POPC lipid membrane bilayer, surrounded by TIP3P water molecules (red) and some neutralising ions (blue spheres). While such a system *in vivo* would have much more complexity, simulating too many components (thereby increasing the system constraints) would be more time-consuming and demand more computational power. As such, there is always a trade-off between system complexity and computational power and time.

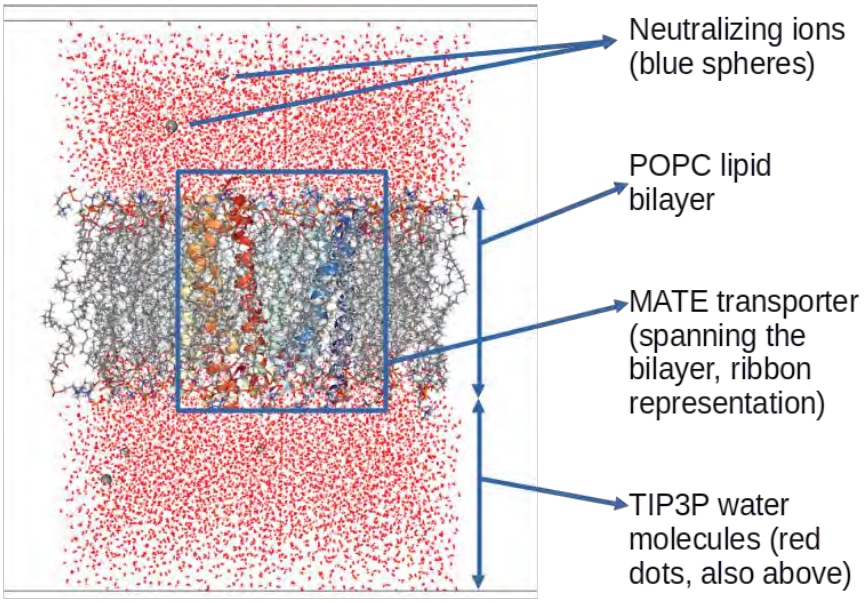


Figure 4.2 | Representation of the entire system generated via the web-based CHARMM-GUI membrane builder tool: each system comprises MATE transporters either in their distinct IFC or OFC, and in their *apo*- and complex-forms, embedded and oriented into a homogenous POPC lipid bilayer; in addition, neutralizing ions were added to the system; finally, the system was solvated with TIP3P water molecules.

Following the above steps, the required input files (structure (.gro) and topology (.top)) for MD simulations in GROMACS were successfully generated and downloaded.

4.3.3 | Energy minimization

Energy minimization is an important, routine practice performed prior to running MD simulations, and aims to determine a set of coordinates representative of the minimal energy conformation for the given structure (Gautam, 2020). The stability of a protein (or in this case a system) is correlated to its energetic state and can be described as a function of its atomic coordinates. Several elements are entailed in this energy function, including: a) the bond and angle energy of covalent bonds; b) energy from dihedral angles; c) van der Waals energy (to prevent atomic steric clashes); d) electrostatic energy due to forces between charged atoms (Gautam, 2020). The entire set of these energy quantifications that have been parameterized is referred to by the term 'force-field', of which several options are available; these include CHARMM, AMBER, AMBERJOPLS, and GROMOS forcefields (Gautam, 2020). Energy minimizations are thus necessary to ensure the system components are in a relaxed state, and at their local energy minima (Jaidhan *et al.,* 2014). This is done by calculating the potential energy for a single conformation as a sum of the individual aforementioned bond energies, which then undergoes energy minimization, frequently performed using gradient optimization, via different available methods

(Jaidhan *et al.,* 2014). As mentioned above, calculations of force-fields are typically derived from MM methods; classical MM methods consider the particle-particle interactions that can replicate structural and conformational changes (Liu *et al.,* 2018). Furthermore, for membrane systems like in the current study, energy minimization is also useful in resolving possible bad lipid-lipid contacts (Jo *et al.,* 2009), as well as for correcting other structural abnormalities (Gautam, 2020).

In the current study, all system minimizations were run for 5000 steps using GROMACS, which utilizes the steepest descent algorithm (for gradient optimization) and set to converge when the maximum force was < 1000 kJ/mol. The selected force-field was the CHARMM 36m force-field, whose lipid and protein parameters have been experimentally validated with the TIP3P water model (also used in this study) and has been found to be a highly suitable choice of force-field for simulations involving membrane-bound proteins (Reid *et al.,* 2022). A minimization curve was then generated in GROMACS and inspected using the software, Grace (Turner, 2005); the minimizations were determined as successful when the generated graphs showed a steady decline in potential energy and reached convergence.

4.3.4 | System equilibration runs

As with energy minimization, it is also standard practice to equilibrate the system prior to running MD simulations. Once the system is in a relaxed state, it is brought to desired temperature and pressure conditions to mimic the *in vivo* thermodynamic state, which is known as equilibration. In the current study, the system was equilibrated following the six-phase consecutive equilibrations recommended upon generation of the system using the CHARMM-GUI membrane builder; these are described in detail by Jo *et al.* (2007), and shown in Table 4.1.

Table 4.1 | Description of the six-phase equilibration steps performed for each system (adapted from Jo *et al.*, 2007): the system was equilibrated using NVT ensembles (steps 1 and 2, 50 ps), followed by equilibration via an NPAT ensemble in the remaining steps (325 ps).

Step	Ensemble[1]	Timesteps	Equilibration Time	Force Constants for Harmonic Restraint[2]				
				Protein Backbone[3]	Protein Sidechain[3]	Water[4]	Lipid[5]	Ion[3]
1	NVT	1 fs	25 ps	10.0	5.0	2.5	2.5	10.0
2	NVT	1 fs	25 ps	5.0	2.5	2.5	2.5	0.0
3	NPAT	1 fs	25 ps	2.5	1.0	1.0	1.0	0.0
4	NPAT	2 fs	100 ps	1.0	0.5	0.5	0.5	0.0
5	NPAT	2 fs	100 ps	0.5	0.1	0.1	0.1	0.0
6	NPAT	2 fs	100 ps	0.1	0.0	0.0	0.0	0.0

[1]NVT stands for constant volume and temperature, and NPAT for constant pressure, area, and temperature.
[2]Force constants are in kcal/(mol·Å^2).
[3]Positional harmonic restraints.
[4]Harmonic restraints to keep water molecules away from the membrane hydrophobic region.
[5]Harmonic restraints to keep the lipid tail in -5 Å$<Z<5$ Å, and lipid head groups close to the membrane surface ($Z=\pm17$ Å for DMPC and $Z=\pm19$ Å for DPPC and POPC).
doi:10.1371/journal.pone.0000880.t001

The first two steps involved the equilibration of the system (protein backbone and sidechains, water, membrane lipids and neutralizing ions) using a temperature ensemble (NVT – constant volume and temperature), each running for 25ps. For typical MD simulations the pressure is equilibrated using an NPT (constant number of particles, constant pressure and constant temperature) ensemble. However, in the case of membrane proteins, like MATE transporters, the anisotropic nature of membranes – specifically, the thermodynamic variations between the lateral and normal orientations of membranes – necessitates the use of an NPAT (constant pressure, area, and temperature) ensemble. As shown in table 4.1, the last four steps, running for a total of 325ps, utilize an NPAT ensemble for pressure equilibration, at 303.15K (for POPC lipids).

4.3.5 | MD production run and trajectory generation

Once energy minimization and temperature and pressure equilibrations have been performed, the production run can be performed. The simulations ran for a total of 100 ns using GROMACS, generating trajectories for the given time period. Given the demand for high computational power, the production run (as well as the above equilibrations), were performed on the CHPC (Center for High Performance Computing) supercomputer.

4.4 | Results and discussion

The trajectories generated from the 100ns MD simulations were analysed using three standard metrics – the root mean square deviation (RMSD), root mean square fluctuation (RMSF), and radius of gyration (Rg), each of which provide different information relating to the dynamic nature and structural stability of the protein structure.

The RMSD provides a measure of the overall structural stability of the protein, and can indicate whether the protein is undergoing significant conformational changes over time. On the other hand, the RMSF provides a measure of the flexibility of individual regions of the protein, and can be used to identify specific residues or regions that are particularly flexible or rigid. Lastly, the Rg can be used to monitor changes in the size and shape of a protein over time, and to identify regions of the protein that are undergoing significant expansion or contraction.

4.4.1 | RMSD (Root Mean Square Deviation)

The RMSD is a metric of structural comparison; whilst initially used to compare structures obtained via x-ray crystallography, advances in technology, however, have led to its use as a conventional tool employed in the analysis of biomolecular simulations. The RMSD measures the global energetic stability, computed as the average distance between atoms of two superimposed structures – the initial (energy minimized and equilibrated pre-simulation protein/ligand) structure and that of the structure at various time-points throughout the simulation. Simply put, the RMSD measures the difference in the atomic positions between two static structures with respect to time, which is related to its stability, and which gives an indication of possible conformational changes (Kufareva et al., 2011).

The RMSD is known to be influenced by various factors including the size (molecular weight) of the macromolecule as well as the structural resolution (Carugo, 2003; Sargsyan *et al.,* 2017). In addition, smaller proteins (< 200 residues) require the use of *weighted* RMSD analysis (Sargsyan *et al.,* 2017); in the current study, however, non-weighted RMSD analyses were used as MATE proteins are large (typically > 400 residues). Generally, the protein size is known to affect the RMSD; proteins of larger lengths tend to exhibit greater RMSD values.

The MATE protein OFC structures used in this study consisted of the homology modelled MATE *apo* OFC which comprises 425 amino acid residues, two docked complexes of the MATE OFC HM in complex with ligands from the SANCDB – MATE OFC-SANC369 and OFC-SANC438 –, as well as a downloaded structure of the MATE OFC in complex with a derivative of the drug, norfloxacin (PDB ID: 3VVP, resolution: 2.91 Å), which comprises 461 amino acid residues. For the IFC, the only available

IFC structure (PDB ID: 6FHZ, resolution: 2.80 Å) was used, which comprises 456 amino acids. As such, MATE transporters are considered an average sized protein.

In the current study, the RMSD plots of the backbone atoms of the MATE protein (the *apo* and complex structures of both the IFC and OFCs), as well as the ligands (for the docked complexes and the available MATE OFC in complex with a derivative of the known drug norfloxacin (NRF) were generated using the software Grace.

4.4.1.1 | RMSD of MATE OFC protein backbone

MD simulations were run for 100 ns on four MATE OFC structures. These included the *apo* OFC (HM), the two selected docked complexes OFC-SANC369 and OFC-SANC438, and an available complex (PDB ID: 3VVP) of the MATE OFC in complex with a derivative of the known drug, norfloxaxin, OFC-NRF. The resulting trajectories were analysed via RMSD plots, firstly with respect to the protein backbone atoms (Figure 4.3), and secondly for the ligands of the OFC complexes (Figure 4.4).

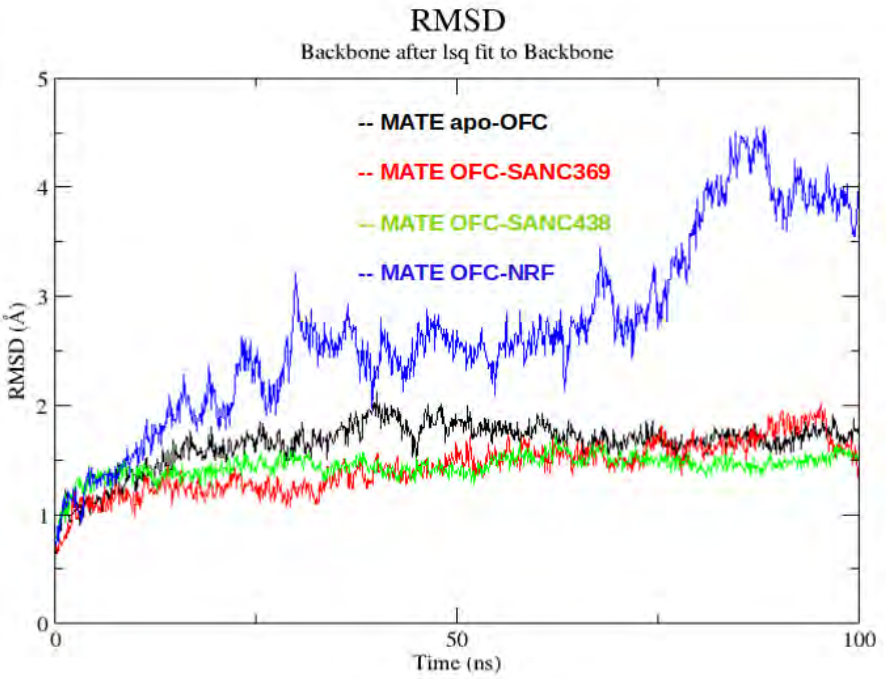


Figure 4.3 | RMSD plot of backbone atoms of MATE OFC structures during the 100 ns MD simulations: The plot shows the RMSD (Å) values of four MATE OFC structures including the MATE OFC apo-protein (black), the MATE OFC-SANC369 docked complex (red), the MATE OFC-SANC438 docked complex (green), and lastly, that of the available MATE OFC-NRF complex.

The general trend for all structures was an increase in RMSD over the simulation time period, although the extent varied amongst the structures. Immediately noticeable from the RMSD plot in Figure 4.3 is that the OFC-NRF complex (blue) showed higher RMSD values and fluctuations throughout the 100 ns of the simulation in comparison to the remaining three MATE OFC structures; at 0.1 ns, the RMSD value is 0.7 Å ± 0.01, which increases to a maximum of 4.55 Å ± 0.01 at 88.2 ns, and eventually begins to maintain RMSD values within a smaller range, reaching 4.0 Å ± 0.01 by the end of the 100 ns simulation. An overall increase of 3.3 Å ± 0.01 was observed from the start to the end of the simulation. The trend in the RMSD of the OFC-NRF complex suggested that the protein complex underwent greater global structural changes and was highly dynamic and less stable in comparison to the other MATE OFC structures. A small contribution to this could be that in addition, larger proteins are known to exhibit higher RMSDs in comparison to smaller ones, and the OFC-NRF MATE complex was somewhat larger (453 residues) than the HM OFC structure (425 residues)

Next, for the majority of the simulation time, the RMSD values of the *apo* OFC (in black) remained higher than that of the two docked complexes, but significantly lower than that of OFC-NRF, with an RMSD value of 0.88 Å ± 0.01 at 0.1 ns, and plateauing at 1.80 Å ± 0.01 at 100 ns. This trend in RMSD suggests that the *apo* OFC remained rather stable and underwent few conformational changes and motions through out the simulation time. The overall change in RMSD from start to finish was 0.92 Å ± 0.01, significantly less than the OFC-NRF complex.

For the docked complex, OFC-SANC369 (red), the RMSD values increased moderately overtime; at 0.1 ns, the RMSD was 0.76 Å ± 0.01, peaking to 1.66 Å ± 0.01 at 98 ns, and eventually decreasing and plateauing, with an RMSD of 1.42 Å ± 0.01 by the end of the 100 ns simulation time. An overall increase of 0.66 Å ± 0.01 was observed, which was the lowest increase in comparison to all other structures. As such, the complex OFC-SANC369 seems to have undergone minimal global structural changes and fewer motions throughout the simulation time, suggesting it remained relatively energetically stable.

The docked complex, OFC-SANC438 (green) remained relatively stable throughout the simulation time. An increase from 0.76 Å ± 0.01 at ns to 1.56 Å ± 0.01 at 100ns was observed, with an overall increase of 0.8 Å ± 0.01. The RMSD trend suggests that this complex seemed rather stable energetically throughout the simulation period on account of possible fewer global structural changes.

It should be noted that in the current study, the OFC of the MATE transporter (apo-OFC) was obtained via homology modelling – from which the two docked complexes (OFC-SANC369 and OFC-SANC438) were generated following molecular docking (chapter 3). The RMSD trend of these three structures, as such, seem to correlate, whereas that of the downloaded complex (OFC-NRF) stood out. The latter complex structure (PDB ID: 3VVP) was crystallized at a resolution of 2.9 Å, which primarily

exhibited higher RMSD values and fluctuations throughout the simulation time (and a resulting greater RMSD difference between the start and end). As such, the marked deviation in the trend could be as a result of the resolution at which the downloaded structure was captured, or because of the difference in the process by which the structures were obtained (x-ray crystallization verses HM). It is also possible that the conditions under which the x-ray crystal structure was obtained (in the absence of membrane) meant that there was some deviation in conformation from when present in the membrane, and that this high RMSD reflects this difference. With regards to global protein stability, the results suggest that the OFC SANC-438 complex is the most stable of all four OFC structures, undergoing fewer conformational changes, followed by the apo-OFC and the SANC369 complex, and lastly the OFC-NRF complex.

4.4.1.2 | RMSD of MATE OFC ligands

The RMSD of the ligands of the OFC complexes (MATE OFC-SANC369, MATE OFC-SANC438, MATE OFC-NRF) was also determined (Figure 4.4). NRF (Br-norfloxacin, MW = 319.3 g/mol, from the MATE OFC-NRF complex (green) exhibited lower RMSD values initially, which gradually peaked to 6.03 Å ± 0.01 by the 27.9 ns. This range of RMSD values was maintained for the rest of the simulation, reaching 5.81 Å ± 0.01 by the end of the simulation. In comparison to the other ligands, the RMSD values of Norfloxacin were generally higher for the latter part of the simulation.

In the case of the docked complex, OFC-SANC369 (black), the ligand (Parviflorone F, MW = 450.52 g/mol) RMSD peaked to 4.53 Å ± 0.01 by 27.7 ns, and this range (± 0.2) was generally maintained for most of the simulation, eventually reaching an RMSD of 4.64 Å ± 0.01, by the end of the simulation. Overall, this ligand maintained RMSD values lower than that of Br-norfloxacin, but higher than that of the other docked complex, OFC SANC-438).

The RMSD values of the ligand (Neodiospyrin, MW = 374.34 g/mol) from the docked complex, OFC SANC-438 (red), were lower than those of the other two ligands throughout the simulation. Until about 90.1 ns, the RMSD ranged between 1 – 2.3 Å ± 0.01, after which it increased and maintained a slightly higher range from 2.5 – 2.9 Å ± 0.01, eventually reaching 2.75 Å ± 0.01 by the end of the simulation time.

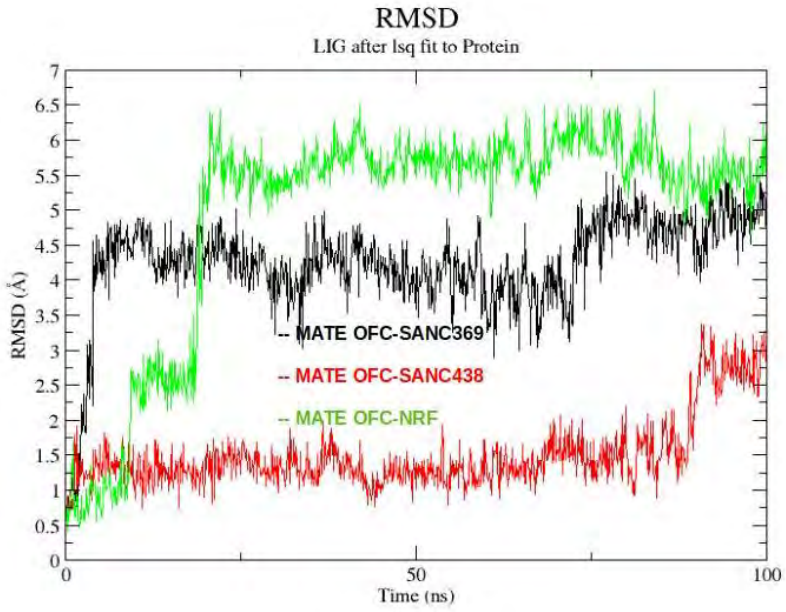


Figure 4.4 | RMSD plot of ligands of MATE OFC complexes during the 100 ns MD simulations: The plot shows the RMSDs (Å) of three different ligands with respect to the MATE OFC protein - SANC369 (black), SANC438 (red), and known drug NRF (green) – in complex with the MATE OFC structure.

While Br-Norfloxacin and Neodiospyrin are of closer molecular weights (380.24 g/mol and 374.34 g/mol, respectively), they exhibited the highest and lowest RMSD values, respectively, throughout the simulation, whereas Parviflorone F, with the highest molecular weight (450.52 g/mol) exhibited RMSD values in mid-range of the other two. As such, it seems that the MWs of the ligands do not influence their interaction with the MATE OFC (and by extension RMSD values). In addition, while the RMSD values may be high, the fluctuations are not extreme, and may indicate that all three ligands remain relatively stable when interacting with the MATE transporter but do undergo some structural change. The results, however, suggest that of the three ligands, the interaction between the MATE OFC protein and the ligand, Neodiospyrin (SANC438), is the most stable, which is followed by Parviflorone F (SANC369), and finally Br-Norfloxacin (OFC-NRF).

4.4.1.3 | RMSD of MATE IFC protein backbone

RMSD plots were also generated for the backbone atoms of the IFC protein structures, shown in Figure 4.5.

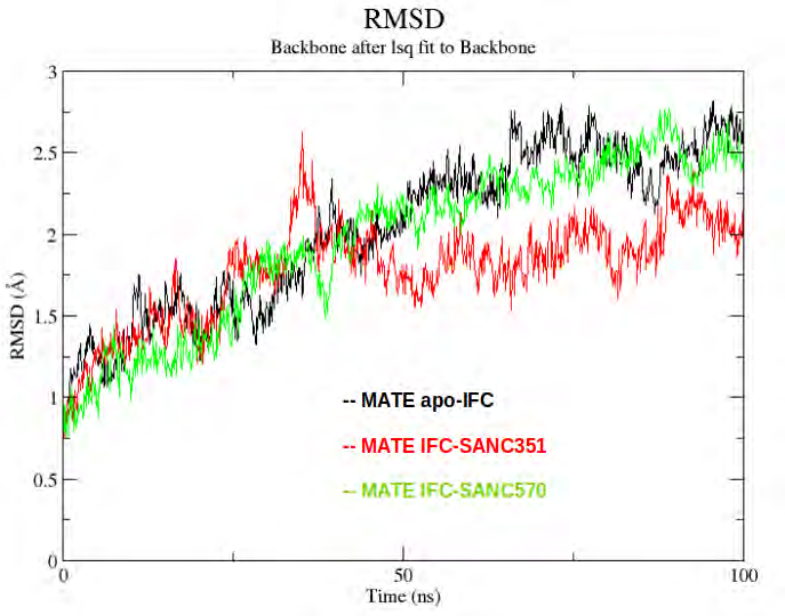


Figure 4.5 | RMSD plot of backbone atoms of MATE IFC structures during the 100 ns MD simulations: The plot shows the RMSD values for the three MATE IFC structures – apo-IFC (black), IFC-SANC00351 (red) and IFC-SANC00570 (green).

As seen in Figure 4.5, the RMSD values for all three structures gradually increased over the simulation period. The RMSD of the apo-IFC (black) increased from 0.86 Å ± 0.01 at 0.1 ns to 2.68 Å ± 0.01 by the end of the simulation, showing an overall increase by 1.82 Å ± 0.01. Following closely, the RMSD of the docked complex, IFC-SANC570 (green) increased from 0.6 Å ± 0.01 to 2.4 Å ± 0.01, showing an overall increase by 1.8 Å ± 0.01. Lastly, the RMSD of the docked complex, IFC-SANC351 (red), increased from 0.76 Å ± 0.01 at 0.1 ns to 2.07 Å ± 0.01 by the end of the simulation, showing an overall increase of 1.31 Å ± 0.01.

4.4.1.4 | RMSD of MATE IFC ligands

The RMSD of the ligands of the IFC docked complexes were also determined, and the plot generated using the software, Grace, is shown in Figure 4.6. The docked complex, IFC-SANC351 (black), comprises the (downloaded) IFC MATE protein (PDB ID: 6FHZ) with the bound SANCDB ligand (5,7-Dihydroxy-3-(3'-hydroxy-4'-methoxybenzyl)-4-chromanone, MW = 316.3 g/mol); the second docked complex, IFC-SANC570 (red) comprises the (downloaded) IFC MATE protein (PDB ID: 6FHZ) with

the bound SANCDB ligand (3,5,7-Trihydroxy-3-(3'-hydroxy-4'methoxybenzyl)-4-chromanone, MW = 332.3 g/mol).

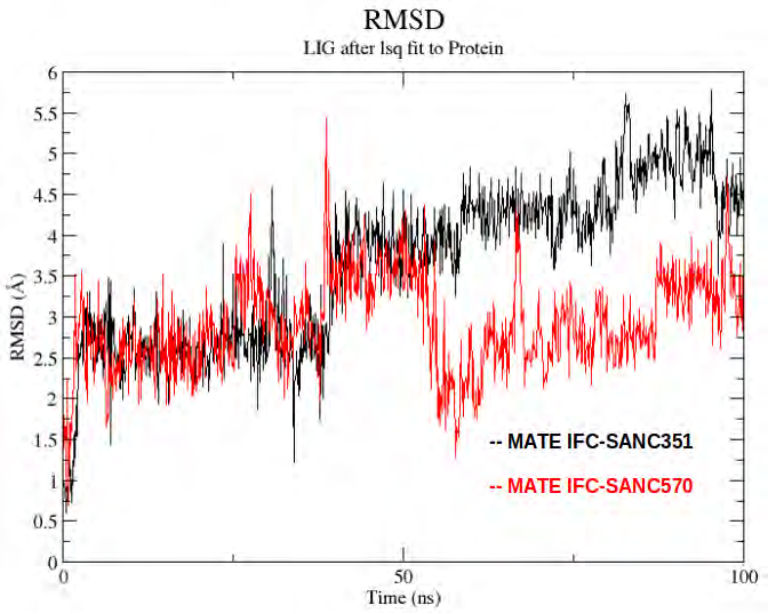


Figure 4.6 | RMSD plot of ligands of MATE IFC docked complexes during the 100 ns MD simulations: The plot shows the RMSDs (Å) of the two SANCDB ligands from the docked complexes with respect to the MATE IFC protein - IFC-SANC351 (black), IFC-SANC570 (red), in complex with the MATE IFC structure.

As seen in Figure 4.6, the RMSD of the IFC-SANC351 complex (black) increases over the 100 ns simulation time. At 0.1 ns, the RMSD is 0.88 Å ± 0.01, peaking to 4.62 Å ± 0.01 at 75.4 ns and to 4.69 Å ± 0.01 at 78.9 ns, eventually reaching an RMSD of 4.5 Å ± 0.01 by the end of the simulation. The RMSD trend suggests that the ligand bound to this complex is somewhat less stable, and highly dynamic throughout the simulation time, while it interacts with the MATE IFC. For the ligand bound to the docked complex IFC-SANC570 (red), the RMSD values overlap for about half the simulation time, after which a slight decrease occurs, and lower RMSD values are maintained for the rest of the simulation time, in comparison to the other ligand. As such it seems more stable and likely undergoes fewer structural changes when interacting with the MATE IFC.

4.4.1.4 | RMSD concluding remarks

The RMSD was used to evaluate the global stability of the MATE IFC and OFC structures, from which the dynamic behaviour of the MATE transporters over the 100 ns simulation was inferred. Table 4.2 gives a summary of the various RMSD values presented and discussed above with the RMSD plots.

Table 4.2 | Summary able of the RMSD values computed from all the seven 100 ns MDS trajectories

Structure	RMSD (Å) at 0.1 ns (± 0.01)	RMSD (Å) at 100 ns (± 0.01)	Change in RMSD (Å)
PROTEIN BACKBONE			
OFCs:			
MATE apo-OFC	0.88	1.80	0.92
MATE OFC-SANC369 docked complex	0.76	1.42	0.66
MATE OFC-SANC438 docked complex	0.76	1.56	0.80
MATE OFC-NRF available complex	0.70	4.0	3.30
IFCs:			
MATE apo-IFC	0.86	2.68	1.82
MATE IFC-SANC351 docked complex	0.76	2.07	1.31
MATE IFC-SANC570 docked complex	0.60	2.40	1.80
LIGAND			
OFCs:			
MATE OFC-SANC369 docked complex	0.57	4.64	4.07
MATE OFC-SANC438 docked complex	0.79	2.75	1.96
MATE OFC-NRF available complex	0.63	5.82	5.19
IFCs:			
MATE IFC-SANC351 docked complex	0.88	4.5	3.62
MATE IFC-SANC570 docked complex	0.98	3.52	2.54

The overall RMSD change that occurred from start to finish was computed for each structure and used to compare conformational changes amongst the structures. With regards to the OFC, the RMSD trend of backbone atoms of the downloaded complex (OFC-NRF), showed a greater overall change – and therefore likely underwent greater structural transitions – in comparison to the *apo*-protein the other two docked complexes, respectively. The RMSD values of the OFC-NRF ligand correlated to the dynamic nature of it the MATE OFC structure it is bound to. Furthermore, the overall RMSD changes were lower in contrast to the three IFC structures, with the exception of OFC-NRF. With regards to the IFC, the backbone RMSD values suggest that the *apo* IFC is slightly more dynamic/less stable than the other two IFC structures.

With the MATE transporters mediating substrate efflux from the cell into the outer environment, they are dependent on structural rearrangements and conformational changes for their function. They employ a rocker-switch model with an alternating access mechanism, continually revoking and exposing their binding site to a substrate (IFC) and undergoing conformational changes that allow it to release the substrate to the extracellular side (OFC) – which is driven via an inward-facing electrochemical gradient. In addition, MATE transporters belong to the *all α-helices* class, in terms of their secondary

structure, known to be more flexible than other classes of proteins, allowing them to undergo the necessary structural changes to perform their function within cell membranes. MATE transporters are thus highly dynamic in nature, and as expected show increasing and in in some cases significant RMSD fluctuations, resulting from structural transitions.

The RMSD, as a means of trajectory analysis, however, has several limitations. As mentioned, it is a global measure, which means that it provides information on the overall structural/conformational changes but gives no indication of the region(s) of the structure that are undergoing these dynamic motions. As such, other complementary analyses measures are required. In the current study, the MD trajectories were also analysed via RMSF and Rg plots (discussed in the following subsections).

4.4.2 | RMSF

The RMSF is another common analysis metric for MD simulation trajectories. Whilst the RMSD is a measure of global stability (governed by structural/conformational changes), the RMSF is used to evaluate the (local) flexibility of the protein residues throughout the simulation time. In other words, it provides insight into the location (i.e. identifies the residues) within the structure that are are more flexible (or rigid) and undergo greater spatial changes (Fraccalvieri *et al.,* 2011). Similar to the RMSD calculation, the RMSF calculates the fluctuation in the position of individual residues in a protein structure from their average positions over time, indicating the degree to which each residue moves during the simulation and can highlight residues that are more flexible or rigid compared to others. The more flexible residues are known to play a role in the function-dynamics link, are able to move and change shape by undergoing conformational changes; Rigid regions, on the other hand, maintain a fixed structure and are not able to move or change shape, and are known to maintain structural integrity of the protein structure and provide stability to the overall protein conformation (Sinha and Smith-Gill, 2002).

The RMSF plots were generated using the software Grace. The residues that showed significant peaks were identified from the plots, following which, their location in terms of their TM helix was determined by analysing the protein amino acid sequence to match the location of the fluctuating residue. This was done using the online tool PDBTM (Protein Data Bank of Transmembrane Proteins, Kozma *et al.,* 2013), which clearly showed the various MATE protein regions (membrane segments and non-membrane segments). In addition, analysis of the RMSF results for the MATE complex structures was complemented by generating 2D diagrams of the ligand-protein interactions, to check for a link between the fluctuating residue with the residues involved in the ligand-protein interaction.

4.4.2.1 | RMSF of MATE OFCs

Figure 4.7 shows the RMSF plot of the MATE OFC structures over the 100 ns simulation time.

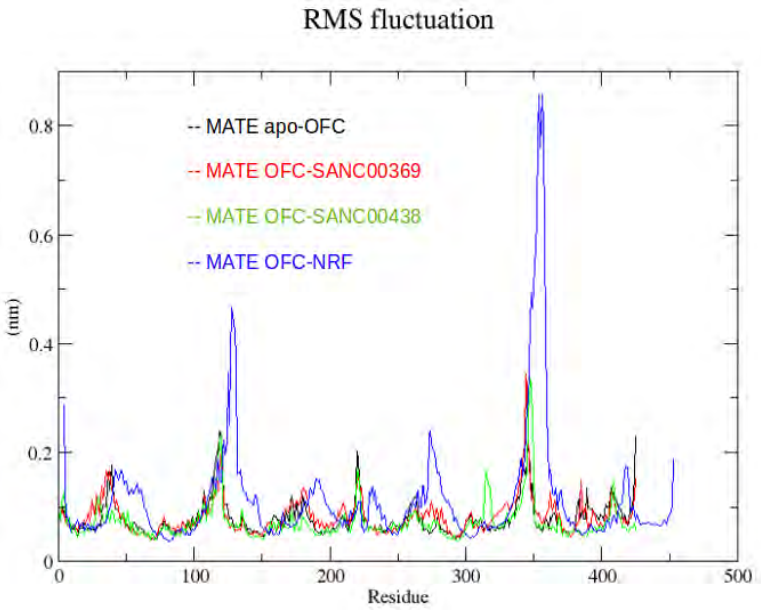


Figure 4.7 | RMSF analysis plot of the MATE OFC structures after the 100 ns MD simulation: The plot shows the RMSF of the protein residues within the OFC of the MATE apo-protein (black), the two docked complexes, OFC-SANC00369 (red) and OFC-SANC00438 (green), and the available complex OFC-NRF (blue).

As seen in Figure 4.7, the residues of the OFC-NRF complex (blue) showed a higher degree of fluctuation in comparison to the other three OFC structures. Specifically, the highest peaks are observed at residues 351 to 358 (0.54 nm ± 0.01 to 0.86 nm ± 0.01, MET 28 on TM9/10, non-membrane segment), 121 – 131 (0.20 nm ± 0.01 to 4.7 nm ± 0.01, ARG 13 on TM3/4, non-membrane segment), 273 – 278 (0.18 nm ± 0.01 to 0.24 nm ± 0.01, LEU 23 on TM7/8, non-membrane segment), and residue 4 (2.9 nm ± 0.01, LYS 4 on TM1, non-membrane segment).

The MATE *apo* OFC (black) shows five predominant peaks at residue 37 and 39 (0.16 and 0.18 nm ± 0.01, respectively, LEU 2 on TM1, membrane segment), 119 and 120 (0.24 and 0.23 nm ± 0.01, respectively, LYS 7, LYS 8 on TM3/4, non-membrane segment), 220 (0.20 nm ± 0.01, LEU 13 on TM6/7, non-membrane segment), 341 – 347 (0.16 nm to 0.21 nm ± 0.01, MET 21 on TM9/10, non-membrane segment), and 424 – 425 (0.16 and 0.23 nm ± 0.01, LEU 26 on TM12, membrane segment). Predominant peaks were observed in the case of the docked compound, OFC-SANC00438 (green), at residues 117 – 120 (0.17 – 0.23 nm ± 0.01, ARG 12 on TM3/4, membrane + non-membrane segment), 220 and 221 (0.17 nm ± 0.01, GLN 24 on TM1), 315 (0.16 ± 0.01, TRP 34 on TM2), and 345 – 350 (0.16 – 0.33 nm ± 0.01, LEU 38 on TM2).

Lastly, for the docked complex, OFC-SANC00369 (red), predominant peaks were observed at residues 119 – 121 (0.16 to 0.21 nm ± 0.01, ARG 12 on TM6/7, non-membrane segment), and 342 – 346 (0.18 to 0.35 nm ± 0.01, GLY 37 and LEU 38 on TM 9/10, non-membrane segment).

In order to investigate the observed RMSF fluctuations, the ligand-protein interactions of the complex structures (OFC-NRF, OFC-SANC00369 and OFC-SANC00438) were inspected using Discovery Studio Visualizer (DSV). By observing which amino acids of the protein interact with the various ligands, possible correlations between more flexible residues (higher peaks) can be revealed. Figure 4.8 shows a 2D diagram of the ligand-protein interactions for the MATE OFC complexes.

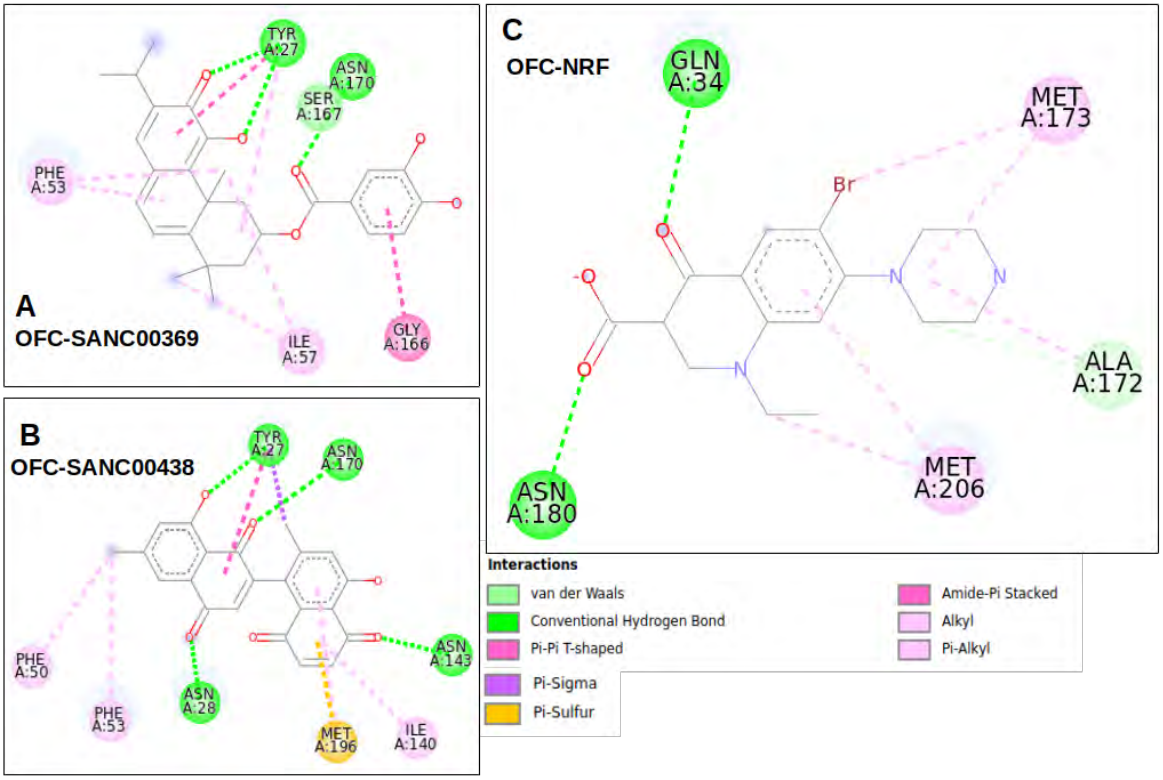


Figure 4.8 | 2D protein-ligand interactions of the MATE OFC complexes: Interacting amino acid residues and bond types are shown for the three OFC complexes – A) OFC-SANC00369, B) OFC-SANC00438 and C) OFC-NRF.

Upon inspection of the protein-ligand interactions in Figure 4.8 for each of the OFC complexes, it was discovered that none of the protein residues involved in ligand interactions generated the above mentioned RMSF peaks, suggesting that the ligand-protein interactions did not influence the dynamic behaviour of these residues; instead it is more likely that these residues are more flexible as they are involved in the functional motions underlying substrate export from the cell.

4.4.2.2 | RMSF of MATE IFCs

RMSF analyses were also performed on the MATE IFC structures; Figure 4.9 shows the plot generated using Grace.

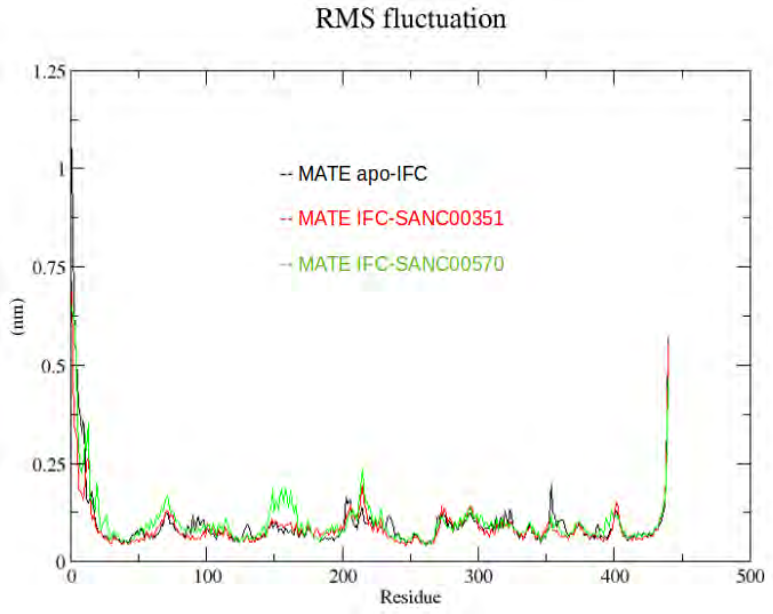


Figure 4.9 | RMSF analysis plot of the MATE IFC structures after the 100 ns MD simulation: The plot shows the RMSF of the protein residues within the IFC of the MATE *apo*-protein (black), and the two docked complexes, OFC-SANC00351 (red) and OFC-SANC00570 (green).

As seen in Figure 4.9, there seems to be significant overlapping of peak residues amongst the three structures. In the case of the MATE apo-IFC (black), peaks were observed at residues 1 – 5 (1.05 nm ± 0.01 to 0.47 nm ± 0.01, LYS 1 on TM1, non-membrane segment), 66 – 79 (0.08 nm ± 0.01 to 0.12 nm ± 0.01, ILE 4, VAL 5 on TM2, membrane segment), 202 – 221 (0.8 nm ± 0.01 to 0.16 nm ± 0.01, MET 12, ILE 13 on TM6/7, membrane + non-membrane segment), 352 – 364 (0.88 nm ± 0.01 to 0.17 nm ± 0.01, TYR 21, ASN 22 on TM9/10, non-membrane segment + membrane segment) and residues 397 – 403 (0.81 nm ± 0.01 to 0.13 nm ± 0.01 ASP 25 on TM11, membrane segment).

In the case of the docked complex, IFC-SANC351 (red), peaks were observed at residues 1 – 5 (0.30 nm to 0.71 nm ± 0.01, LYS 1, TM1 non-membrane segment), 11 – 14 (0.20 nm ± 0.01 to 0.26 nm ± 0.01, LYS 1, TM1 non-membrane segment), 70 – 75 (0.11 nm ± 0.01 to 0.13 nm ± 0.01, ILE 4, TM2, membrane segment), 210 – 217 (0.12 nm ± 0.01 to 0.19 nm ± 0.01, MET 12, ILE 13, TM6, membrane segment), 292 – 298 (0.10 nm ± 0.01 to 0.14 nm ± 0.01, GLN 18, TM8, membrane segment), and 400 – 403 (0.11 nm ± 0.01 – 0.14 nm ± 0.01, ASP 25, TM 11, membrane segment).

In the case of the docked complex, IFC-SANC570 (green), peaks were observed at residues 1 – 5 (0.35 nm to 0.65 nm ± 0.01, TM1 non-membrane segment), 11 – 13 (0.28 nm to 0.35 nm ± 0.01, TM1 non-membrane segment), 66 – 76 (0.12 nm to 0.17 nm ± 0.01, TM2 membrane segment), 146 – 166 (0.12 nm to 0.18 nm ± 0.01, TM4 membrane segment), 206 – 221 (0.10 nm to 0.23 nm ± 0.01, TM6 membrane segment), and residues 438 – 440 (0.19 nm to 0.46 nm ± 0.01, TM12 membrane segment).

As in the case of the OFC, 2D protein-ligand diagrams of the MATE IFC complexes (IFC-SANC351 and IFC-SANC570) were generated using Discovery Studio Visualizer (DSV), in order to determine wheather the fluctuating residues were involved in ligand interactions or instead were possible function related flexible residues. Figure 4.10 shows these 2D diagrams of the ligand-protein interactions for the MATE IFC complexes.

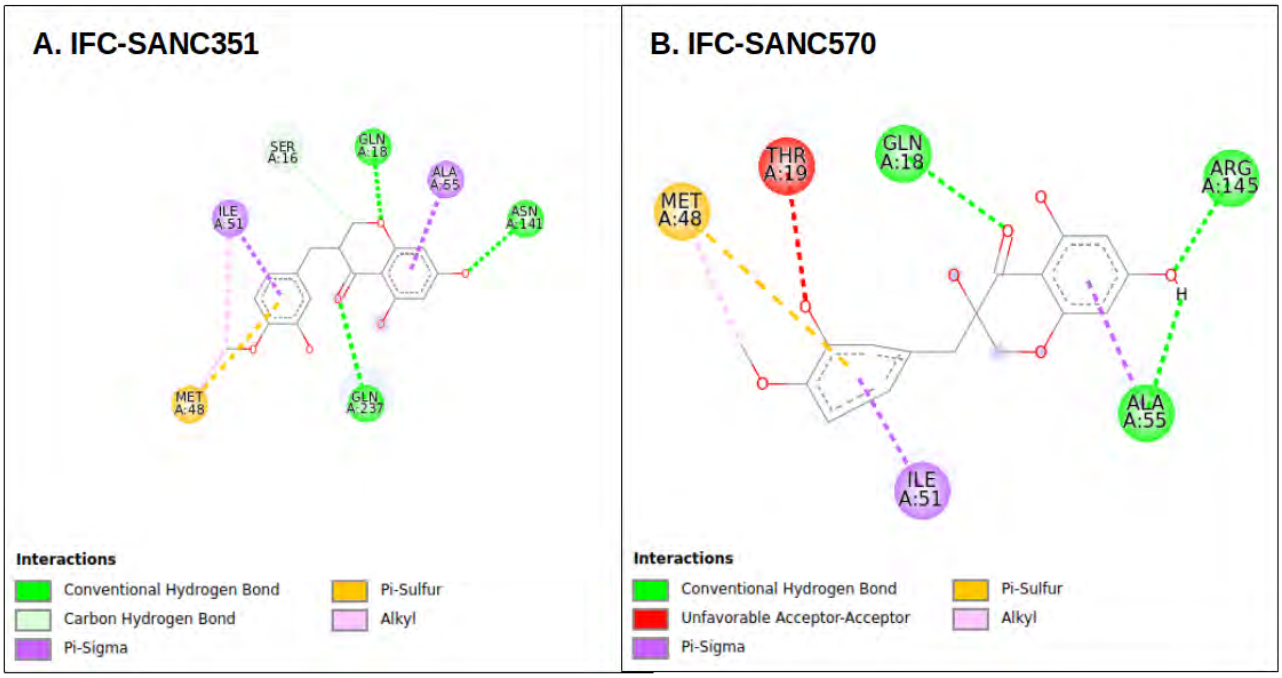


Figure 4.10 | 2D protein-ligand interactions of the MATE IFC complexes: Interacting amino acid residues and bond types are shown for the two IFC complexes – A) IFC-SANC351, B) IFC-SANC570.

Upon visualisation of the protein-ligand interactions (Figure 4.10 B), an unfavourable acceptor-acceptor bond was discovered in the IFC-SANC570 complex structure, indicating that it forms a less energetically stable interaction with the MATE IFC, and that either the compound needs to be optimized to bind to form a more stable interaction with the protein, or that it should be replaced by another (docked) compound altogether, that forms energetically stable bonds. This is discussed further in the next chapter, Chapter 5.

In the case of both IFC complexes, however, the fluctuating residues do not participate in the protein-ligand interaction, implying that these residues are instead involved in the molecular-level mechanism of the substrate-transport function of MATE transporters.

4.4.2.3 | RMSF concluding remarks

As previously explained (in Chapter 1), MATE transporters consist of 12 TM α-helices spanning the membrane bilayer, and although multiple sequence alignments have revealed no consensus sequence amongst MATE transporter subfamilies, various conserved regions have been identified, and are thought to be vital to the function of MATE transporters. These include the perimeters of TM1 and TM7 along with the extracellular loops linking TM1 with TM2 and TM7 with TM8, the cytoplasmic loops that connect TM2 with TM3 and TM8 with TM9, and finally in the loops that link TM4 with TM5 and TM10 with TM11 (Otsuka et al., 2005). Furthermore, in the case of PfMATEs (structures used in this study), previous studies have found TM1 rearrangements (and an Asp 41 residue) are important in the MATE transporter cycle (Tanaka *et al*., 2013). While the Asp 41 residue was not one of the fluctuating residues in either conformation, the fluctuating residues did fall within the regions found to be important for the function of MATE transporters at a molecular level.

4.4.3 | Radii of gyration

The Rg is another standard means of MD simulation trajectory analysis, providing insight into the 'compactness' of the protein. The Rg refers to the spatial positioning of the atoms of a protein around its rotational axis (or center of gravity); it is a measure of the root-mean-square distance of individual atoms of the protein to their center of mass (Huang and Powers, 2001; Sneha and Doss, 2016). A lower Rg is indicative that the atoms of the structure are closer to their centroid, and thus the structure is more compact, and vice versa.

The compactness of a protein is determined by several factors. The first of these factors is the *class* to which the protein belongs – which is based on the secondary structure of the protein. The four main classes include proteins whose secondary structures consist of either: a) all α-helices, b) all β-pleated sheets, c) α-helices/β-pleated sheets, and d) α-helices + β-pleated sheets (Lobanov *et al.,* 2008).

Identifying and investigating the mechanism(s) underlying the dynamic nature of a protein – which is key to its function – provides important insight into the variables that govern their molecular-level behaviour; this opens up the possibility of gaining some control over its functioning. Investigation of the Rg of members from all four structural classes have revealed the distinctive impact of the protein

class on the Rg and thus compactness of the protein; α-proteins exhibit the highest Rg, which is indicative of a more loosely packed structure in contrast to β - and (α +β)-proteins; whereas α /β-proteins exhibit the lowest Rg and more tightly packed structure in comparison to the other classes (Lobanov *et al.,* 2008).

MATE transporters belong to the class of α-proteins, with a unique topology consisting of 12 TM α-helices. As such, they are expected to exhibit higher Rg values with a less compact structure as the atoms would be expected to be further away from their centroid. In addition, α-proteins are known to be more flexible and dynamic in nature than proteins of the other classes and thus are expected to undergo greater conformational changes that are known to be associated with their function (Lobanov *et al.,* 2008).

Rg analysis of MD trajectories for each of the seven structures was performed using GROMACS, and the resulting plots were generated using Grace and are presented and discussed in the following sections.

4.4.3.1 | Radii of gyration of MATE OFC structures

Figure 4.11 shows the Rg plot of the MATE OFC structures generated from the trajectory of the 100 ns MD simulations.

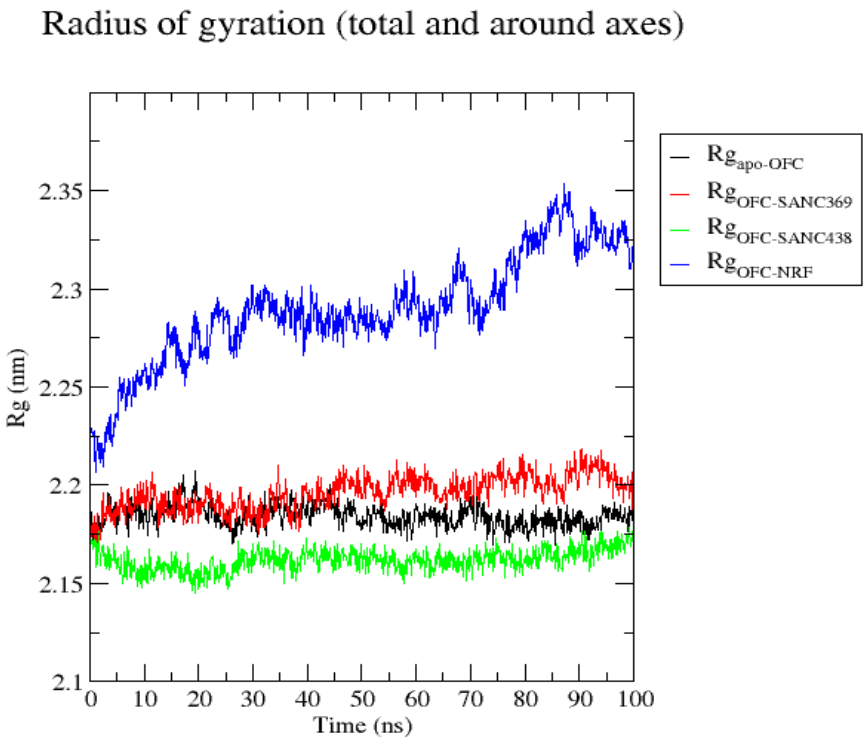


Figure 4.11 | Radius of gyration (Rg) plot of MATE OFCs: Trajectory analysis of the 100 ns MD simulations via Rg calculation of the four MATE OFC structures are shown: apo-OFC (black), OFC-NRF (blue), and the docked complexes, OFC-SANC369 (red) and OFC-SANC438 (green).

Immediately noticeable in Figure 4.11, is that the ORF-NRF complex (blue) exhibits higher Rg values throughout the 100 ns simulation time in comparison to the other three OFC structures. The Rg starts off at 2.23 nm ± 0.01 at 0.1 ns, peaking to 2.34 nm at 92.3 ns, and eventually reaching a value of 2.32 nm ± 0.01 by the end of the simulation time. Of the OFC structures, the Rg suggests that the OFC-NRF complex is the least compact, and likely undergoing greater conformational transitions throughout the simulation time.

The docked complex OFC-SANC369 (red) follows next, showing (slightly) higher Rg values throughout the simulation time than the remaining structures, with an Rg value of 2.18 nm ± 0.01 at 0.1 ns, maintained within a tight range throughout the simulation, reaching an Rg value of 2.20 nm ± 0.01 by the end of the simulation time. The Rg values show that this structure is more compact than the OFC-NRF complex, but less compact than the remaining two structures.

Next, the apo-OFC (black) followed closely, with an Rg value of 2.18 nm ± 0.01 at 0.1 ns, reaching an Rg of nm ± 0.01 by the end of the simulation; the Rg values suggest the apo-OFC structure is slightly more compact than the above mentioned structures.

Lastly, the docked complex, OFC-SANC438 (green), showed the lowest Rg values in comparison to the other OFC structures. At 0.1 ns, the Rg value was 2.17 nm ± 0.01, which was the same at the end of the 100 ns simulation time. As such, the Rg values indicate that this docked complex is the most compact, and undergoes the least amount of structural transitions throughout the simulation time.

4.4.3.1 | Radii of gyration of MATE IFC structures

Figure 4.12 shows the Rg plot of the MATE IFC structures generated from the trajectory of the 100 ns MD simulations.

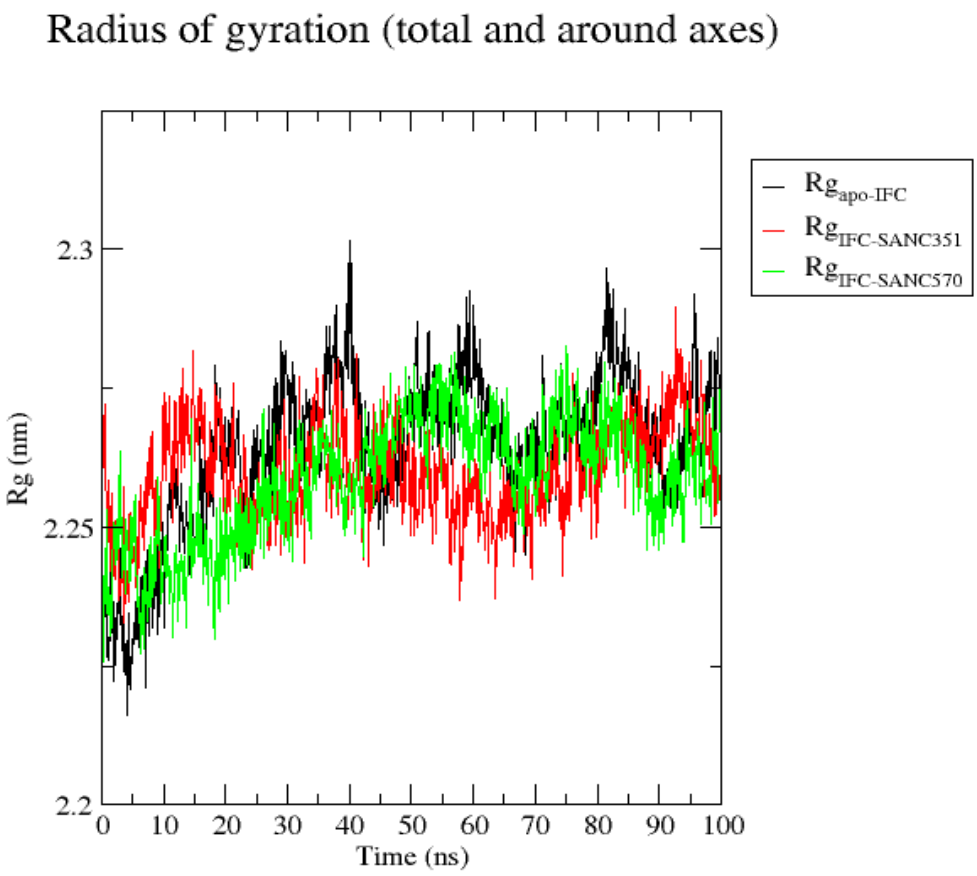


Figure 4.12 | Radius of gyration (Rg) plot of MATE IFCs: Trajectory analysis of the 100 ns MD simulations via Rg calculation of the three MATE IFC structures are shown: apo-IFC (black), and the docked complexes, OFC-SANC351 (red) and OFC-SANC570 (green).

As seen in Figure 4.10, the Rg values of the three IFCs overlap throughout the 100 ns simulation time. The apo-IFC (black) starts off with an Rg value of 2.23 nm ± 0.01 at 0.1 ns, reaching an Rg value of 2.28 nm ± 0.01 by the end of the simulation. The docked complex IFC-SANC351 (red) shows an Rg value of 2.26 nm ± 0.01 at 0.1 ns, which is also the Rg value by the end of the 100 ns simulation; suggesting it is slightly more compact than the apo-IFC. The docked complex IFC-SANC570 (green) has an Rg value of 2.24 nm ± 0.01 at 0.1 ns, reaching a value of 2.26 nm ± 0.01 by the end of the 100 ns, showing similar (but slightly less) compactness in comparison the the apo-IFC and the other docked complex, and thus undergoes less conformational transitions during the simulation.

The Rg values overlap significantly for all three IFC structures; it seems as though they all undergo similarly moderate conformational transitions throughout the simulation time and are relatively compact with respect to each other.

4.4.3.2 | Radii of gyration concluding remarks

Simulation trajectories were evaluated via Rg plots – a measure of the size and shape of a protein – to analyse the dynamic behaviour in terms of the structural changes occurring in the MATE transporter structures over the 100 ns time-frame.

In general, an increase in Rg can indicate an expanded and more flexible structure, while a decrease in Rg can indicate a more compact and rigid structure. The MD simulations of the MATE OFC structures showed a significant increase in Rg for the OFC-NRF complex and slight increases for the other three structures. In the case of the IFC, fluctuations and increases in the Rg, especially in the *apo*-protein were observed. It is known that MATE transporters undergo significant conformational changes during their transport cycle; the Rg trends reflect their dynamic and flexible nature, accounting for their substrate transport function across the lipid bilayer.

4.5 | Chapter conclusion

Proteins are dynamic structures, with some requiring more flexibility than others depending on their function. IMPs, especially those involved in substrate transport across the cellular membrane – such as MATE transporters – are highly dynamic and often flexible in nature, which is apparent in their distinctive energy landscapes.

While it is known that MATE transporters function as antiporters under a rocker-switch model with an alternating-access mechanism, the details of the molecular-level dynamics governing their function as transporter proteins remain largely elusive. As such, the current study employed MD simulations to investigate the structure-function-dynamics relationship of MATE transporters in order to study them as drug targets.

The energy landscape of MATE transporter proteins was studied using MD simulations, allowing modelling of the protein's structure and dynamics over a 100 ns time-frame, from which trajectories were generated based on the energy changes that occur as the protein changes conformation. The resulting trajectories provide a dynamic representation of the protein's behaviour, allowing for the exploration of its conformational space. To quantify the structural changes that occurred during the 100 ns MD simulations, several (complementary) metrics were used; the RMSD was used to quantify the protein’s stability, indicative of global conformational changes occurring; the RMSF was used to identify individual residues that are more flexible and move more than others, indicating residues

important for their function – in the case of the complex structures, 2D diagrams of the protein-ligand interactions were also generated to view possible correlation between fluctuating and interacting residues; the Rg was used to determine size and shape of the protein in terms of how compact is it.

The protein backbone RMSD of the MATE *apo*- IFC and OFC as well as the complexes, indicated global structural transitions occurred during the 100 ns simulation. Peaks observed in the RMSF plots revealed the individual protein residues that had moved significantly from their average spatial positions. In order to determine whether these residues were involved in protein-ligand interactions in the complex structures, 2D protein-ligand interaction diagrams were generated to view the residues involved in forming bonds with the SANCDB ligands. It was determined that in the case of both the IFC and OFC structures, the residues involved in protein-ligand interactions did not correlate with those that produced peaks in the RMSF plot. The TM locations of the fluctuating residues were identified and compared to the important functional regions other researchers have found. From this, it was concluded that the fluctuating residues are likely involved in the function of the transporter that allow for structural rearrangements to occur. These residues were found in both the membrane segments – which are expected to be more flexible and dynamic so that the MATE transporter can undergo conformational changes during the transport cycle – as well as in the loop regions of non-membrane segments – also involved in structural rearrangements. Lastly, the Rg of both the *apo* and complex forms of the IFC and OFC indicated that MATE transporters are rather loosely packed, as opposed to being compact structures; belonging to the class of all α-proteins, their secondary structures, as in the case of other proteins belonging to this class, are known to be more flexible to support their function. As such, these proteins are predicted to have greater Rg values and are less compact in nature, which was expected and apparent for the MATE transporters.

Figure 2.5 (Chapter 2) provides a diagram of the proposed ion-driven transport cycle of the MATE transporters. As explained, when in the OFC, both the ion- and substrate- binding sites are unoccupied and exposed to the extracellular side; an inwardly directed ion gradient prompts the binding of an ion located in the extracellular side, which in turn induces the collapse of various TM helices (Claxton et al., 2021). In the case of PfMATE, Tanaka et al. (2013) found, that it is the N-lobe cavity comprising TM1 and TM2 that collapses, releasing the substrate into the extracellular space. As such, during the substrate-release step of the MATE transport cycle, it is expected that the OFC would undergo dynamic motions that allow it to release the substrate, completing the efflux cycle of the transporter. This would suggest a large RMSD to account for the global transitions involved in the collapse of the N-lobe for substrate release; the peak RMSF residues would be expected to correspond to residues in the functional, more flexible regions that support the efflux function; and lastly, the Rg is also expected to be large, as the MATE transporter belongs to the all α-helices class of protein, that are inherently more flexible and

less compact on account of dynamic motions needed to support their function. Together, these metrics speak to the structure-function-dynamics link of the MATE transporter.

Tanaka et al. (2013) also established the successful efflux of Br-Norfloxacin by PfMATE using a fluorescence-based assay. The crystal structure of this complex (PDB ID: 3VVP) was used as a control OFC during the MD simulations – OFC-NRF. Evaluation of the MD trajectory of this complex using each of the evaluation metrics followed the expected trends and exhibited large RMSD values (for both protein backbone and ligand) throughout the simulation, numerous RMSF peaks, and larger Rg values – also confirming the successful efflux of the bound substrate, Br-Norfloxacin.

From comparing these metrics amongst the OFC structures, however, it is apparent that the control complex exhibits significantly higher values than the two docked complexes. The SANCDB ligands bound to the MATE OFC HM seem to form more stable interactions – which is evident in their lower RMSD, RMSF, and Rg values. This indicates that the efflux of these ligands by the MATE transporter may not be successful. However, they seem to bind well to the active site of the protein, and thus have the potential to be optimized into conformation-locking inhibitors that would prevent the conformational switching from the OFC to the IFC.

Evaluation of all three metrics reflected that MATE transporters are a highly dynamic protein, with some regions more flexible than others; this allows them to undergo the necessary structural rearrangements to allow for conformational transitions during the substrate transport cycle. These insights can be useful for developing more effective strategies for inhibiting or regulating the function of MATE transporters. The next chapter, Chapter 5, discusses this study's limitations and future work to achieve this.

CHAPTER 5: CONCLUDING REMARKS AND FUTURE WORK

5.1 | Combining molecular docking and MD simulations

This study aimed to explore the MATE transporters as potential antimicrobial drug targets, owing to their contribution to the global issue of MDR. A two-fold in silico approach – combining molecular docking and MD simulations – was implemented with two main objectives; firstly to identify potential inhibitory compounds via docking and virtual screening of SANCDB compounds against both the IFC and OFC of the MATE protein; and secondly to gain a comprehensive understanding of the molecular interactions between the selected docked complexes, and in doing so to explore the structure-function-dynamics link of MATE transporters via MD simulations. Combining molecular docking and MD simulations is a common approach used in CADD, providing information that complement each other when investigating a drug target.

Molecular docking aims to predict the most energetically favourable binding mode (or spatial orientation) of a ligand interacting with a target protein and is useful in the discovery of potential 'hits' against a drug target; docking provides a more static view of ligand-protein interactions, generating rigid 'poses' of the interaction between a protein and ligand. MD simulations, on the other hand, offer complementary information whereby the dynamics of the protein-ligand interaction can be modelled over a period of time and under set conditions to mimic the protein's physiological environment.

Following up molecular docking with MD simulations allows for the analysis of the stability of the ligand-protein complex, providing a means of docking validation, and, in addition, allowing for the evaluation of the conformational changes and stability of the 3D structure of the protein. As such, the two techniques combined are important in gaining a deeper understanding of the dynamics involved in the binding interaction of a ligand with a drug target as well as the behaviour of the protein and the structural and conformational changes occurring in the protein necessary for its functional mechanism.

In addition to validation of the results of molecular docking, MD simulations provide further information about the interactions between hit molecules and proteins and their mechanism(s) of action. This information can be used to optimize hit compounds into lead compounds (hit-to-lead, discussed briefly in Section 5.3 on Future work and potential scope of this study) and can help guide the design of new, improved drug molecules with increased specificity and efficacy.

5.2 | Limitations of this study

At the later results analysis stage, a limitation was discovered in the flow of the methodology. Firstly, post-docking validation was performed via filtering the lowest energy poses following a visual inspection to ensure that the ligand had bound to the correct substrate-binding site of the MATE transporter. The binding site was known – from available complex structures such as that used as a control (PDB ID: 3VVP, referred to as MATE OFC-NRF) – and ligands were allowed to bind freely, as opposed to defining the binding site. This ensured that the ligands were allowed to bind where they formed energetically stable bonds. As such those that bound to incorrect regions were eliminated. As opposed to inspecting the 2D protein-ligand interaction diagrams at this stage, in the case of the complex structures, the current study generated and viewed these interactions at the MD analysis stage in order to assess whether the fluctuating residues in the RMSF plots were correlated. As such it was discovered at a later stage that one of ligands had formed one unfavourable acceptor-acceptor bond with the MATE IFC in the IFC-SANC570 complex structure, indicating that it forms a less energetically stable interaction and may not be a suitable ligand. Due to time constraints, however, MD simulations were not carried out on another more suitable ligand bound to the IFC from the docking data. However, the current ligand either needs to be optimized at a latter stage, or replaced by the next favourable ligand from the generated docked poses, with inspection of the 2D interaction diagrams at an earlier stage in the future.

In addition, while docking is more time-effective and requires less computational power than MD simulations, virtual screening and docking of all the SANCDB compounds were performed; however, because of the above-mentioned constraints, only two docked complexes, for both the IFC and OFC, were selected to further explore via MD simulations. In addition, simulating a whole system, as in the case of MATE transporters that need to be embedded in a lipid membrane bilayer, contributes to the time and computational costs. As such, not enough data from MD simulations was generated to determine more potential lead inhibitory compounds; in the future, the rest of the docked complexes, with ligands that fit the criteria of binding at the correct site and with the lowest binding energies, can be further studied via MD simulations for their potential to target the MATE transporter. As such this limitation also contributes to the future work.

Lastly, as previously mentioned, the substrate transport of MATE EPs is driven by an inwardly directed ion gradient; the ion preference differs amongst the various MATE sub families. Members of the DinF sub family, including PfMATEs – the structures used in this study – are driven by an H^+ ion gradient; this inwardly directed H^+ gradient would cause the binding of an ion located in the extracellular side to bind at the ion-binding site of the MATE OFC, resulting in the release of the substrate into the extracellular environment. As a limitation of this study, however, an ion gradient was not set up before

conducting the MD simulations and, therefore, the effects of an ion gradient on the dynamic behaviour of the MATE transporters was not investigated; this limitation also contributes to the future work of this study. Simulating an ion gradient would be worthwhile investigating in the future, as targeting the driving source of these transporters may offer another pathway to inhibit or regulate its function. When performing MD simulations, however, there tends to be a compromise between the system size and complexity and the computational and time demand. The presence of an ion gradient requires the calculation of electrostatic interactions between the ions and other components of the system, which can increase the computational cost. Additionally, the movement of ions in response to the gradient can result in changes to the system's overall charge distribution, which must be continuously recalculated throughout the simulation. As such, perhaps a coarse-grained model of the ions, which reduces the number of degrees of freedom that need to be calculated, can be used.

5.3 | Future work and potential scope of this study

Contemporary drug discovery and development involves three main steps – the identification of disease-related biological targets, the identification and optimization of new 'lead' compounds, and finally the development of a novel drug (Liu et al., 2018). As mentioned in the introductory chapter (Chapter 1), the scope of this study lies in the preliminary stages of the drug discovery and design process, as shown in the flowchart in Figure 5.1.

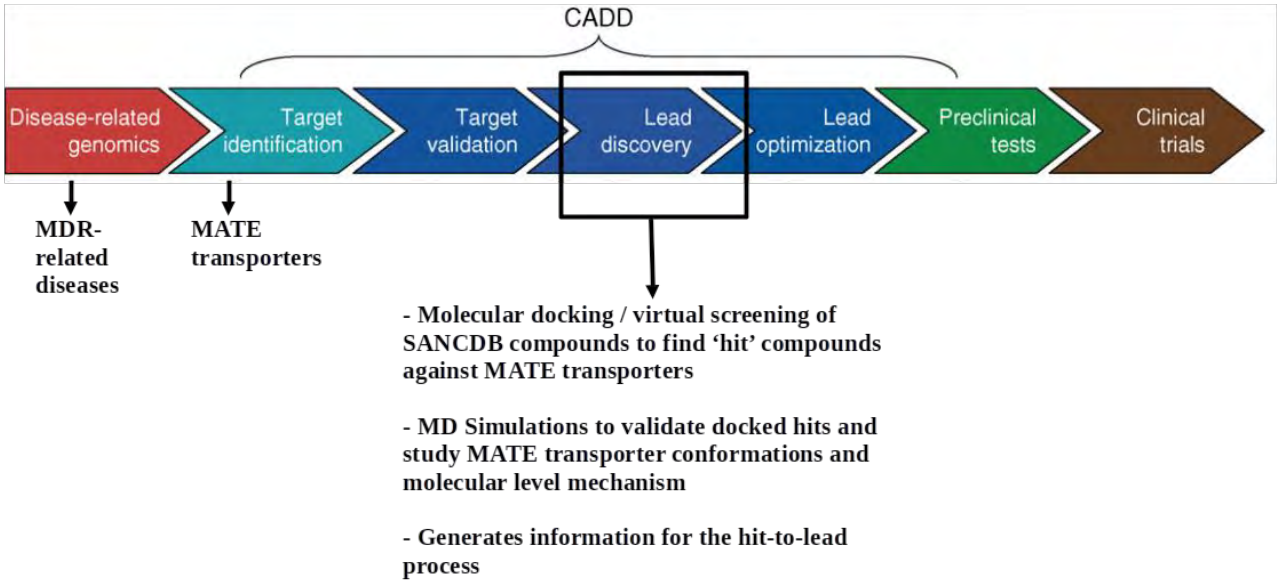


Figure 5.1 | CADD pipeline showing scope of the current study as well as the future scope in drug targeting of MATE transporters (Adapted from Tang et al., 2006).

MATE transporters are crucial drug targets, and achieving cellular control over their efflux of treatment drugs may offer a future solution over MDR, which as mentioned, is a worldwide issue; targeting them requires a deep understanding of their dynamic behaviour as well as the way in which they interact with ligands that have the potential to act as drugs against them at a later stage. The aim to investigate the MATE transporters as potential antimicrobial targets was achieved in via the aforementioned two-fold objectives. The insights generated by this study – both on potential inhibitory ligands against the MATE transporter, as well as a molecular-level view of their underlying mechanism of function by exploring these interactions with the MATE transporter within a membrane bilayer over time – are valuable to the next steps of CADD as shown in Figure 5.1.

REFERENCES

Agarwal S, Chadha D, Mehrotra R (2015) Molecular modeling and spectroscopic studies of semustine binding with DNA and its comparison with lomustine–DNA adduct formation. J Biomol Struct Dyn 33: 1653-1668.

Alav, I., Sutton, J. M., & Rahman, K. M. (2018). Role of bacterial efflux pumps in biofilm formation. The Journal of antimicrobial chemotherapy, 73(8), 2003–2020. https://doi.org/10.1093/jac/dky042.

Alibert, S., N'Gompaza Diarra, J., Hernandez, J., Stutzmann, A., Fouad, M., Boyer, G., Pages, J.M., (2017). Multidrug efflux pumps and their role in antibiotic and antiseptic resistance: a pharmacodynamic perspective. Expert Opin. Drug Metab. Toxicol., 13, 301–309.

Andersen, J., Taboureau, O., Hansen, K. B., Olsen, L., Egebjerg, J., Strømgaard, K., & Kristensen, A. S. (2009). Location of the antidepressant binding site in the serotonin transporter: importance of Ser-438 in recognition of citalopram and tricyclic antidepressants. The Journal of biological chemistry, 284(15), 10276–10284. https://doi.org/10.1074/jbc.M806907200.

Antimicrobial Resistance Collaborators (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet (London, England), 399(10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0.

Appel, M., Hizlan, D., Vinothkumar, K. R., Ziegler, C., & Kühlbrandt, W. (2009). Conformations of NhaA, the Na+/H+ exchanger from Escherichia coli, in the pH-activated and ion-translocating states. Journal of molecular biology, 388(3), 659–672. https://doi.org/10.1016/j.jmb.2009.03.010.

Atanasov, A. G., Waltenberger, B., Pferschy-Wenzig, E. M., Linder, T., Wawrosch, C., Uhrin, P., Temml, V., Wang, L., Schwaiger, S., Heiss, E. H., Rollinger, J. M., Schuster, D., Breuss, J. M., Bochkov, V., Mihovilovic, M. D., Kopp, B., Bauer, R., Dirsch, V. M., & Stuppner, H. (2015). Discovery and resupply of pharmacologically active plant-derived natural products: A review. Biotechnology advances, 33(8), 1582–1614. https://doi.org/10.1016/j.biotechadv.2015.08.001.

Bakheet, T. M., & Doig, A. J. (2010). Properties and identification of antibiotic drug targets. BMC bioinformatics, 11, 195. https://doi.org/10.1186/1471-2105-11-195.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res, 28(1): 235–242.

Beuming, T., Kniazeff, J., Bergmann, M. L., Shi, L., Gracia, L., Raniszewska, K., Newman, A. H., Javitch, J. A., Weinstein, H., Gether, U., & Loland, C. J. (2008). The binding sites for cocaine and dopamine in the dopamine transporter overlap. Nature neuroscience, 11(7), 780–789. https://doi.org/10.1038/nn.2146.

Birch, J., Cheruvara, H., Gamage, N., Harrison, P. J., Lithgo, R., & Quigley, A. (2020). Changes in Membrane Protein Structural Biology. Biology, 9(11), 401. https://doi.org/10.3390/biology9110401.

Blaney J. (2012). A very short history of structure-based design: how did we get here and where do we need to go?. Journal of computer-aided molecular design, 26(1), 13–14. https://doi.org/10.1007/s10822-011-9518-x.

Blanco, P., Hernando-Amado, S., Reales-Calderon, J. A., Corona, F., Lira, F., Alcalde-Rico, M., Bernardini, A., Sanchez, M. B., & Martinez, J. L. (2016). Bacterial Multidrug Efflux Pumps: Much

More Than Antibiotic Resistance Determinants. Microorganisms, 4(1), 14. https://doi.org/10.3390/microorganisms4010014.

Boudker, O., & Verdon, G. (2010). Structural perspectives on secondary active transporters. Trends in pharmacological sciences, 31(9), 418–426. https://doi.org/10.1016/j.tips.2010.06.004.

Boudker, O., Ryan, R. M., Yernool, D., Shimamoto, K., & Gouaux, E. (2007). Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. Nature, 445(7126), 387–393. https://doi.org/10.1038/nature05455.

Bowie J. U. (2004). Membrane proteins: a new method enters the fold. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(12), 3995–3996. https://doi.org/10.1073/pnas.0400671101.

Brett, C. L., Donowitz, M., & Rao, R. (2005). Evolutionary origins of eukaryotic sodium/proton exchangers. American journal of physiology. Cell physiology, 288(2), C223–C239. https://doi.org/10.1152/ajpcell.00360.2004.

Brown, D. & Superti-Furga, G. (2003). Rediscovering the sweet spot in drug discovery. Drug Discov Today. 8: 1067–1077.

Byrne, B., & Iwata, S. (2002). Membrane protein complexes. Current opinion in structural biology, 12(2), 239–243. https://doi.org/10.1016/s0959-440x(02)00316-0.

Campos, S. R., Machuqueiro, M., & Baptista, A. M. (2010). Constant-pH molecular dynamics simulations reveal a β-rich form of the human prion protein. The journal of physical chemistry. B, 114(39), 12692–12700. https://doi.org/10.1021/jp104753t.

Carugo, O. (2003). How root-mean-square distance (r.m.s.d.) values depend on the resolution of protein structures that are compared. Journal of Applied Crystallography, 36(1), 125–128. doi:10.1107/s0021889802020502.

Chen, L., Liu, Y., Liu, H., Kang, L., Geng, J., Gai, Y., Ding, Y., Sun, H., & Li, Y. (2015). Identification and expression analysis of MATE genes involved in flavonoid transport in blueberry plants. PloS one, 10(3), e0118578. https://doi.org/10.1371/journal.pone.0118578.

Chen, W., van der Kamp, M. W., & Daggett, V. (2014). Structural and dynamic properties of the human prion protein. Biophysical journal, 106(5), 1152–1163. https://doi.org/10.1016/j.bpj.2013.12.053.

Claxton, D. P., Jagessar, K. L., & Mchaourab, H. S. (2021). Principles of Alternating Access in Multidrug and Toxin Extrusion (MATE) Transporters. Journal of molecular biology, 433(16), 166959. https://doi.org/10.1016/j.jmb.2021.166959.

Claxton, D. P., Jagessar, K. L., Steed, P. R., Stein, R. A., & Mchaourab, H. S. (2018). Sodium and proton coupling in the conformational cycle of a MATE antiporter from *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(27), E6182–E6190. https://doi.org/10.1073/pnas.1802417115.

Cragg, G.M and Newman, D.J. (2013). Natural products: a continuing source of novel drug leads. Biochim Biophys Acta 1830:3670–3695.

Dar, A. M, Mir S (2017) Molecular Docking: Approaches, Types, Applications and Basic Challenges. J Anal Bioanal Tech 8: 356. doi: 10.4172/2155-9872.1000356.

Dastvan, R., Fischer, A.W., Mishra, S., Meiler, J., McHaourab, H.S., (2016). Protonation-dependent conformational dynamics of the multidrug transporter. EmrE. Proc. Natl. Acad. Sci. U. S. A., 113, 1220–1225.

Dias, D. A, Urban S, Roessner, U. (2012). A historical overview of natural products in drug discovery. Metabolites 2:303–336.

Diener, A. C., Gaxiola, R. A., & Fink, G. R. (2001). Arabidopsis ALF5, a multidrug efflux transporter gene family member, confers resistance to toxins. *The Plant cell*, *13*(7), 1625–1638. https://doi.org/10.1105/tpc.010035.

Dodson, G. G., Lane, D. P., & Verma, C. S. (2008). Molecular simulations of protein dynamics: new windows on mechanisms in biology. EMBO reports, 9(2), 144–150. https://doi.org/10.1038/sj.embor.7401160.

Drew D, North RA, Nagarathinam K, Tanabe M. (2021). Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS). Chemical Reviews. May;121(9):5289-5335. DOI: 10.1021/acs.chemrev.0c00983. PMID: 33886296; PMCID: PMC8154325.

Drew, D., & Boudker, O. (2016). Shared Molecular Mechanisms of Membrane Transporters. *Annual review of biochemistry*, *85*, 543–572. https://doi.org/10.1146/annurev-biochem-060815-014520.

Du, D., van Veen, H. W., Murakami, S., Pos, K. M., & Luisi, B. F. (2015). Structure, mechanism and cooperation of bacterial multidrug transporters. *Current opinion in structural biology*, *33*, 76–91. https://doi.org/10.1016/j.sbi.2015.07.015.

Du, D., Wang-Kan, X., Neuberger, A., van Veen, H. W., Pos, K. M., Piddock, L., & Luisi, B. F. (2018). Multidrug efflux pumps: structure, function and regulation. *Nature reviews. Microbiology*, *16*(9), 523–539. https://doi.org/10.1038/s41579-018-0048-6.

Eisinger, M. L., Nie, L., Dörrbaum, A. R., Langer, J. D., & Michel, H. (2018). The Xenobiotic Extrusion Mechanism of the MATE Transporter NorM_PS from Pseudomonas stutzeri. *Journal of molecular biology*, *430*(9), 1311–1323. https://doi.org/10.1016/j.jmb.2018.03.012.

Ferreira, L. G., Dos Santos, R. N., Oliva, G., & Andricopulo, A. D. (2015). Molecular docking and structure-based drug design strategies. Molecules (Basel, Switzerland), 20(7), 13384–13421. https://doi.org/10.3390/molecules200713384.

Ficici, E., Zhou, W., Castellano, S., & Faraldo-Gómez, J. D. (2018). Broadly conserved Na+-binding site in the N-lobe of prokaryotic multidrug MATE transporters. Proceedings of the National Academy of Sciences of the United States of America, 115(27), E6172–E6181. https://doi.org/10.1073/pnas.1802080115.

Forrest, L. R., Krämer, R., & Ziegler, C. (2011). The structural basis of secondary active transport mechanisms. Biochimica et biophysica acta, 1807(2), 167–188. https://doi.org/10.1016/j.bbabio.2010.10.014.

Forrest, L. R., Zhang, Y. W., Jacobs, M. T., Gesmonde, J., Xie, L., Honig, B. H., & Rudnick, G. (2008). Mechanism for alternating access in neurotransmitter transporters. Proceedings of the National

Academy of Sciences of the United States of America, 105(30), 10338–10343. https://doi.org/10.1073/pnas.0804659105.

Fraccalvieri, D., Pandini, A., Stella, F., & Bonati, L. (2011). Conformational and functional analysis of molecular dynamics trajectories by Self-Organising Maps. BMC Bioinformatics, 12, 1–19.

Gautam, B. (2020). Energy Minimization. In R. T. Maia, R. a. de Moraes Filho, & M. Campos (Eds.), Homology Molecular Modeling - Perspectives and Applications. IntechOpen. https://doi.org/10.5772/intechopen.94809.

Ghosh, S., Bornman, C., & Zafer, M. M. (2021). Antimicrobial Resistance Threats in the emerging COVID-19 pandemic: Where do we stand?. Journal of infection and public health, 14(5), 555–560. https://doi.org/10.1016/j.jiph.2021.02.011.

Giacomini K.M., & Sugiyama Y (2017). Membrane transporters and drug response. Brunton L.L., & Hilal-Dandan R, & Knollmann B.C.(Eds.), Goodman & Gilman's: The Pharmacological Basis of Therapeutics, 13e. McGraw Hill. https://accessmedicine.mhmedical.com/content.aspx?bookid=2189&sectionid=166183302.

Giacomini, K. M., Huang, S. M., Tweedie, D. J., Benet, L. Z., Brouwer, K. L., Chu, X., Dahlin, A., Evers, R., Fischer, V., Hillgren, K. M., Hoffmaster, K. A., Ishikawa, T., Keppler, D., Kim, R. B., Lee, C. A., Niemi, M., Polli, J. W., Sugiyama, Y., Swaan, P. W., … Zhang, L. (2010). Membrane transporters in drug development. Nature reviews. Drug discovery, 9(3), 215–236. https://doi.org/10.1038/nrd3028.

Gomeni, R., Bani, M., D'Angeli, C., Corsi, M. & Bye, A. (2001). Computer-assisted drug development (CADD): an emerging technology for designing first-time-in-man and proof-of-concept studies from preclinical experiments. Eur J Pharm Sci. 13:261–270.

Gottesman, M.M., Fojo, T., Bates, S.E., (2002). Multidrug resistance in cancer: role of ATP-dependent transporters. Nature Rev. Cancer, 2, 48–58.

Groeneveld, M., & Slotboom, D. J. (2010). Na(+):aspartate coupling stoichiometry in the glutamate transporter homologue Glt(Ph). Biochemistry, 49(17), 3511–3513. https://doi.org/10.1021/bi100430s.

Gu, J., Gui, Y., Chen, L., Yuan, G., Lu, H.Z., Xu, X. (2013). Use of natural products as chemical library for drug discovery and network pharmacology. PLoS one 8:e62839.

Guan L. (2022). Structure and mechanism of membrane transporters. Scientific reports, 12(1), 13248. https://doi.org/10.1038/s41598-022-17524-1.

Guan, L., & Kaback, H. R. (2006). Lessons from lactose permease. Annual review of biophysics and biomolecular structure, 35, 67–91. https://doi.org/10.1146/annurev.biophys.35.040405.102005.

Guedes, I. A, de Magalhães CS, Dardenne LE (2014) Receptor-ligand molecular docking. Biophysical Reviews 6: 75-87.

Guo Y. (2020). Be Cautious with Crystal Structures of Membrane Proteins or Complexes Prepared in Detergents. Crystals, 10(2), 86. https://doi.org/10.3390/cryst10020086.

Guskov, A., Jensen, S., Faustino, I., Marrink, S. J., & Slotboom, D. J. (2016). Coupled binding mechanism of three sodium ions and aspartate in the glutamate transporter homologue GltTk. Nature communications, 7, 13420. https://doi.org/10.1038/ncomms13420.

Harvey, A. L., Edrada- Ebel, R. & Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. Nat. Rev. Drug Discov. 14, 111–129 (2015).

Hatherley R, Brown DK, Musyoka TM, et al. SANCDB: a South African natural compound database. J Cheminform. 2015;7:29. Published 2015 Jun 19. doi:10.1186/s13321-015-0080-8.

Hatherley, R., Brown, D. K., Glenister, M., & Tastan Bishop, Ö. (2016). PRIMO: An Interactive Homology Modeling Pipeline. PloS one, 11(11), e0166698. https://doi.org/10.1371/journal.pone.0166698.

Hashimoto, K., Ogawa, W., Nishioka, T., Tsuchiya, T., & Kuroda, T. (2013). Functionally cloned pdrM from Streptococcus pneumoniae encodes a Na+ coupled multidrug efflux pump. PLoS One 8:e59525.

Hediger, M. A., Romero, M. F., Peng, J. B., Rolfs, A., Takanaga, H., & Bruford, E. A. (2004). The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. Pflugers Archiv : European journal of physiology, 447(5), 465–468. https://doi.org/10.1007/s00424-003-1192-y.

Hiasa, M., Matsumoto, T., Komatsu, T., & Moriyama, Y. (2006). Wide variety of locations for rodent MATE1, a transporter protein that mediates the final excretion step for toxic organic cations. American journal of physiology. Cell physiology, 291(4), C678–C686. https://doi.org/10.1152/ajpcell.00090.2006.

Higgins C. F. (2007). Multiple molecular mechanisms for multidrug resistance transporters. Nature, 446(7137), 749–757. https://doi.org/10.1038/nature05630.

Hillgren, K M; Keppler, D; Zur, A A; Giacomini, K M; Stieger, B; Cass, C E; Zhang, L (2013). Emerging transporters of clinical importance: an update from the International Transporter Consortium. Clinical Pharmacology and Therapeutics, 94(1):52-63. https://doi.org/10.1038/clpt.2013.74.

Hirayama, B. A., Díez-Sampedro, A., & Wright, E. M. (2001). Common mechanisms of inhibition for the Na+/glucose (hSGLT1) and Na+/Cl-/GABA (hGAT1) cotransporters. British journal of pharmacology, 134(3), 484–495. https://doi.org/10.1038/sj.bjp.0704274.

Hu, L., Benson, M.L., Smith, R.D., Lerner, M.G., and Carlson, H.A. (2005). Binding MOAD (Mother Of All Databases). Proteins, 60(3): 333–340.

Huang, X., & Powers, R. (2001). Validity of using the radius of gyration as a restraint in NMR protein structure determination. Journal of the American Chemical Society, 123(16), 3834–3835. https://doi.org/10.1021/ja005770p.

Huda MN, Morita Y, Kuroda T, Mizushima T, Tsuchiya T (2001) Na+-driven multidrug efflux pump VcmA from Vibrio cholerae non-O1, a non-halophilic bacterium. FEMS Microbiol Lett 203:235–239.

Hvorup, R. N., Winnen, B., Chang, A. B., Jiang, Y., Zhou, X. F., & Saier, M. H., Jr (2003). The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. *European journal of biochemistry*, *270*(5), 799–813. https://doi.org/10.1046/j.1432-1033.2003.03418.x.

Irwin, J. J. and Shoichet, B.K. (2005). ZINC—a free database of commercially available compounds for virtual screening. J Chem Inf Model, 45(1): 177–182.

Issa, N. T., Badiavas, E. V., & Schürer, S. (2019). Research Techniques Made Simple: Molecular Docking in Dermatology - A Foray into In Silico Drug Discovery. The Journal of investigative dermatology, 139(12), 2400–2408.e1. https://doi.org/10.1016/j.jid.2019.06.129.

Jagessar, K. L., Mchaourab, H. S., & Claxton, D. P. (2019). The N-terminal domain of an archaeal multidrug and toxin extrusion (MATE) transporter mediates proton coupling required for prokaryotic drug resistance. The Journal of biological chemistry, 294(34), 12807–12814. https://doi.org/10.1074/jbc.RA119.009195.

Jaidhan, B. J., Rao, P. S., & Apparao, A. (2014). Energy minimization and conformation analysis of molecules using steepest descent method. International Journal of Computer Science Information Technologies, 5(3), 3525–3528.

Jia, R., Martens, C., Shekhar, M., Pant, S., Pellowe, G.A., Lau, A.M., Findlay, H.E., Harris, N.J., et al., (2020). Hydrogen-deuterium exchange mass spectrometry captures distinct dynamics upon substrate and inhibitor binding to a transporter. Nature Commun., 11, 6162.

Jin, X., Shao, Y., Bai, Q., Xue, W., Liu, H., & Yao, X. (2016). Insights into conformational regulation of PfMATE transporter from Pyrococcus furiosus induced by alternating protonation state of Asp41 residue: A molecular dynamics simulation study. *Biochimica et biophysica acta*, *1860*(6), 1173–1180. https://doi.org/10.1016/j.bbagen.2016.02.007.

Jo, S., Kim, T., & Im, W. (2007). Automated builder and database of protein/membrane complexes for molecular dynamics simulations. PloS one, 2(9), e880. https://doi.org/10.1371/journal.pone.0000880.

Jo, S., Lim, J. B., Klauda, J. B., & Im, W. (2009). CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. Biophysical journal, 97(1), 50–58. https://doi.org/10.1016/j.bpj.2009.04.013.

Jomain, J. B., Tallet, E., Broutin, I., Hoos, S., van Agthoven, J., Ducruix, A., Kelly, P. A., Kragelund, B. B., England, P., & Goffin, V. (2007). Structural and thermodynamic bases for the design of pure prolactin receptor antagonists: X-ray structure of Del1-9-G129R-hPRL. The Journal of biological chemistry, 282(45), 33118–33131. https://doi.org/10.1074/jbc.M704364200.

Kaatz GW, McAleese F, Seo SM (2005) Multidrug resistance in Staphylococcus aureus due to overexpression of a novel multidrug and toxin extrusion (MATE) transport protein. Antimicrob Agents Chemother 49:1857–1864.

Kalra S. (2014). Sodium Glucose Co-Transporter-2 (SGLT2) Inhibitors: A Review of Their Basic and Clinical Pharmacology. Diabetes therapy : research, treatment and education of diabetes and related disorders, 5(2), 355–366. https://doi.org/10.1007/s13300-014-0089-4.

Kapp E, Malan SF, Joubert J, Sampson SL. Small molecule efflux pump inhibitors in Mycobacterium tuberculosis: a rational drug design perspective. Mini-Rev Med Chem. 2018;18:72-86.

Kathleen M. Hillgren, Dietrich Keppler, Arik Zur, Kathleen M. Giacomini, Bruno Stieger, Carol E. Cass, and Lei Zhang, Emerging Transporters of Clinical Importance: An Update from the International Transporter Consortium, Clinical Pharmacology & Therapeutics, accepted article preview online 8 April 2013; doi:10.1038/clpt.2013.74.

Kazmier, K., Sharma, S., Quick, M., Islam, S.M., Roux, B., Weinstein, H., Javitch, J.A., McHaourab, H.S., (2014). Conformational dynamics of ligand-dependent alternating access in LeuT. Nature Struct. Mol. Biol., 21, 472–479.

Khazir, Jabeena & Mir, Bilal & Mir, Shabir Ahmad & Cowan, Donald. (2013). Natural products as lead compounds in drug discovery. Journal of Asian natural products research. 15. 10.1080/10286020.2013.798314.

Kermani A. A. (2021). A guide to membrane protein X-ray crystallography. The FEBS journal, 288(20), 5788–5804. https://doi.org/10.1111/febs.15676.

Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov. 2004; 3(11):935–949. [PubMed: 15520816].

Klebe G. Virtual ligand screening: strategies, perspectives and limitations. Drug Discov. Today 11(13–14), 580–594 (2006).

Kobilka, B. K., & Deupi, X. (2007). Conformational complexity of G-protein-coupled receptors. Trends in pharmacological sciences, 28(8), 397–406. https://doi.org/10.1016/j.tips.2007.06.003.

Koehn, F. E., and Carter, G. T. (2005). The evolving role of natural products in drug discovery. Nature Reviews Drug Discovery, 4(3), 206–220. doi:10.1038/nrd1657.

Kozma, D., Simon, I., & Tusnády, G. E. (2013). PDBTM: Protein Data Bank of transmembrane proteins after 8 yezhars. Nucleic acids research, 41(Database issue), D524–D529. https://doi.org/10.1093/nar/gks1169.

Krah, A., Huber, R. G., Zachariae, U., & Bond, P. J. (2020). On the ion coupling mechanism of the MATE transporter ClbM. Biochimica et biophysica acta. Biomembranes, 1862(2), 183137. https://doi.org/10.1016/j.bbamem.2019.183137.

Krishnamurthy, H., Piscitelli, C. L., & Gouaux, E. (2009). Unlocking the molecular secrets of sodium-coupled transporters. Nature, 459(7245), 347–355. https://doi.org/10.1038/nature08143.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. Journal of molecular biology, 305(3), 567–580. https://doi.org/10.1006/jmbi.2000.4315.

Kufareva, I., Rueda, M., Katritch, V., Dock, G.P.C.R., Stevens, R.C., & Abagyan, R. (2011). Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. Structure (London, England: 1993), 19, 1108–1126.

Kumar A, Schweizer HP. Bacterial resistance to antibiotics: active efflux and reduced uptake. Adv Drug Deliver Rev. 2005;57:1486-1513.

Kuroda, T., & Tsuchiya, T. (2009). Multidrug efflux transporters in the MATE family. Biochimica et biophysica acta, 1794(5), 763–768. https://doi.org/10.1016/j.bbapap.2008.11.012.

Kusakizako, T., Miyauchi, H., Ishitani, R., & Nureki, O. (2020). Structural biology of the multidrug and toxic compound extrusion superfamily transporters. *Biochimica Et Biophysica Acta. Biomembranes*, *1862*(12), 183154. https://doi.org/10.1016/j.bbamem.2019.183154.

Lamut, A., Peterlin Mašič, L., Kikelj, D., & Tomašič, T. (2019). Efflux pump inhibitors of clinically relevant multidrug resistant bacteria. Medicinal research reviews, 39(6), 2460–2504. https://doi.org/10.1002/med.21591.

Lappano, R., & Maggiolini, M. (2011). G protein-coupled receptors: novel targets for drug discovery in cancer. Nature reviews. Drug discovery, 10(1), 47–60. https://doi.org/10.1038/nrd3320.

Larsen, A. K., Escargueil, A. E. & Skladanowski, A. Resistance mechanisms associated with altered intracellular distribution of anticancer agents. Pharmacol Ther 85, 217–229 (2000).

Law, C. J., Maloney, P. C., & Wang, D. N. (2008). Ins and outs of major facilitator superfamily antiporters. *Annual review of microbiology*, *62*, 289–305. https://doi.org/10.1146/annurev.micro.61.080706.093329.

Leone V, Pogoryelov D, Meier T, Faraldo-Gómez JD (2015) On the principle of ion selectivity in Na+/H+-coupled membrane proteins: Experimental and theoretical studies of an ATP synthase rotor. Proc Natl Acad Sci USA 112:E1057–E1066.

Li, N., Meng, H., Xing, H., Liang, L., Zhao, X., & Luo, K. (2017). Genome-wide analysis of MATE transporters and molecular characterization of aluminum resistance in Populus. Journal of experimental botany, 68(20), 5669–5683. https://doi.org/10.1093/jxb/erx370.

Liao, C., Peach, M. L., Yao, R., & Nicklaus, M. C. (2013). Molecular docking and structure-based virtual screening. In SilicoDrug Discovery and Design, 6–20. doi:10.4155/ebo.13.181.

Lipinski, C.A., Lombardo, F, Dominy, B.W, Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3–26.

Liu, J., Li, Y., Wang, W., Junyi, G., & Li, Y. (2016). Genome-wide analysis of MATE transporters and expression patterns of a subgroup of MATE genes in response to aluminum toxicity in soybean. BMC Genomics. 17. 10.1186/s12864-016-2559-8.

Liu, X., & Pan, G. (Eds.). (2019). Drug Transporters in Drug Disposition, Effects and Toxicity. Advances in Experimental Medicine and Biology. Doi:10.1007/978-981-13-7647-4.

Liu, X., Shi, D., Zhou, S., Liu, H., Liu, H., & Yao, X. (2018). Molecular dynamics simulations and novel drug discovery. Expert opinion on drug discovery, 13(1), 23–37. https://doi.org/10.1080/17460441.2018.1403419.

Lobanov, M.Y., Bogatyreva, N.S. & Galzitskaya, O.V. Radius of gyration as an indicator of protein structure compactness. Mol Biol 42, 623–628 (2008). https://doi.org/10.1134/S0026893308040195.

Lolkema, J. S., & Slotboom, D. J. (2003). Classification of 29 families of secondary transport proteins into a single structural class using hydropathy profile analysis. Journal of molecular biology, 327(5), 901–909. https://doi.org/10.1016/s0022-2836(03)00214-6.

Lolkema, J. S., & Slotboom, D. J. (2008). The major amino acid transporter superfamily has a similar core structure as Na+-galactose and Na+-leucine transporters. Molecular membrane biology, 25(6-7), 567–570. https://doi.org/10.1080/09687680802541177.

Lomize, M. A., Lomize, A. L., Pogozheva, I. D., & Mosberg, H. I. (2006). OPM: orientations of proteins in membranes database. Bioinformatics (Oxford, England), 22(5), 623–625. https://doi.org/10.1093/bioinformatics/btk023.

Lomize, M.A., Pogozheva, I.D., Joo H., Mosberg, H.I., Lomize, A.L. (2012). OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res. 40 (Database issue), D370-D376. PDF PubMed.

Long F, Rouquette-Loughlin C, Shafer WM, Yu EW (2008). Functional cloning and characterization of the multidrug efflux pumps NorM from Neisseria gonorrhoeae and YdhE from Escherichia coli. Antimicrob Agents Chemother 52:3052–3060.

Loo, D. D., Hirayama, B. A., Sala-Rabanal, M., & Wright, E. M. (2008). How drugs interact with transporters: SGLT1 as a model. The Journal of membrane biology, 223(2), 87–106. https://doi.org/10.1007/s00232-008-9116-6.

Lopes-Rodrigues, V., Seca, H., Sousa, D., Sousa, E., Lima, R. T., & Vasconcelos, M. H. (2014). The network of P-glycoprotein and microRNAs interactions. International journal of cancer, 135(2), 253–263. https://doi.org/10.1002/ijc.28500.

Lopes-Rodrigues, V., Di Luca, A., Mleczko, J., Meleady, P., Henry, M., Pesic, M., Cabrera, D., van Liempd, S., Lima, R. T., O'Connor, R., Falcon-Perez, J. M., & Vasconcelos, M. H. (2017). Identification of the metabolic alterations associated with the multidrug resistant phenotype in cancer and their intercellular transfer mediated by extracellular vesicles. Scientific reports, 7, 44541. https://doi.org/10.1038/srep44541.

Lu M, Radchenko M, Symersky J, Nie R & Guo Y (2013a) Structural insights into H+-coupled multidrug extrusion by a MATE transporter. Nat. Struct. Mol. Biol. 20, 1310–1317.

Lu, M., Symersky, J., Radchenko, M., Koide, A., Guo, Y., Nie, R., and Koide, S. (2013b). Structures of a Na+-coupled, substrate-bound MATE multidrug transporter. Proc. Natl. Acad. Sci. U S A 110, 2099–2104.

Mandal, S.; Moudgil, M.N.; Mandal, S.K. Rational drug design. Eur. J. Pharmacol. 2009, 625, 90–100.

Martinez J.L., Sánchez M.B., Martínez-Solano L., Hernandez A., Garmendia L., Fajardo A., Alvarez-Ortega C. Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. FEMS Microbiol. Rev. 2009;33:430–449. doi: 10.1111/j.1574-6976.2008.00157.x.

Masuda, S., Terada, T., Yonezawa, A., Tanihara, Y., Kishimoto, K., Katsura, T., Ogawa, O., & Inui, K. (2006). Identification and functional characterization of a new human kidney-specific H+/organic cation antiporter, kidney-specific multidrug and toxin extrusion 2. Journal of the American Society of Nephrology : JASN, 17(8), 2127–2135. https://doi.org/10.1681/ASN.2006030205.

McAleese, F., Petersen, P., Ruzin, A., Dunman, P. M., Murphy, E., Projan, S. J., & Bradford, P. A. (2005). A novel MATE family efflux pump contributes to the reduced susceptibility of laboratory-derived Staphylococcus aureus mutants to tigecycline. *Antimicrobial agents and chemotherapy*, *49*(5), 1865–1871. https://doi.org/10.1128/AAC.49.5.1865-1871.2005.

McConkey, BJ, Sobolev, V., & Edelman, M. (2002). The performance of current methods in ligand-protein docking. Current Science, 83 , 845-856.

McMurry L, Petrucci RE Jr., and Levy SB (1980). Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 77: 3974–3977.

Meng, X. Y., Zhang, H. X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. Current computer-aided drug design, 7(2), 146–157. https://doi.org/10.2174/157340911795677602.

Mitchell, P. (1967) Translocations through natural membranes. Adv Enzymol 29:33–85.

Morita Y, Kataoka A, Shiota S, Mizushima T, Tsuchiya T (2000). NorM of vibrio parahaemolyticus is an Na+-driven multidrug efflux pump. J Bacteriol 182:6694–6697.

Morita, M., Shitan, N., Sawada, K., Van Montagu, M. C., Inzé, D., Rischer, H., Goossens, A., Oksman-Caldentey, K. M., Moriyama, Y., & Yazaki, K. (2009). Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in Nicotiana tabacum. Proceedings of the National Academy of Sciences of the United States of America, 106(7), 2447–2452. https://doi.org/10.1073/pnas.0812512106.

Morris, G.M., Lim-Wilby, M. (2008). Molecular Docking. In: Kukol, A. (eds) Molecular Modeling of Proteins. Methods Molecular Biology™, vol 443. Humana Press. https://doi.org/10.1007/978-1-59745-177-2_19.

Mousa, J. J., Newsome, R. C., Yang, Y., Jobin, C., & Bruner, S. D. (2017). ClbM is a versatile, cation-promiscuous MATE transporter found in the colibactin biosynthetic gene cluster. Biochemical and biophysical research communications, 482(4), 1233–1239. https://doi.org/10.1016/j.bbrc.2016.12.018.

Nikaido H. (2009). Multidrug resistance in bacteria. Annual review of biochemistry, 78, 119–146. https://doi.org/10.1146/annurev.biochem.78.082907.145923.

Ntie-Kang, F., Onguéné, P.A., Scharfe, M., Owono Owono, L.C., Megnassan, E., Mbaze, L.M. et al (2014). ConMedNP: a natural product library from Central African medicinal plants for drug discovery. RSC Adv 4:409.

O'Neill, J. (2014) Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations.https://www.who.int/news/item/29-04-2019-new-report-calls-for-urgent-action-to-avert-antimicrobial-resistance-crisis.

Oberai, A., Ihm, Y., Kim, S., & Bowie, J. U. (2006). A limited universe of membrane protein families and folds. Protein science : a publication of the Protein Society, 15(7), 1723–1734. https://doi.org/10.1110/ps.062109706.

Omote, H., Hiasa, M., Matsumoto, T., Otsuka, M., & Moriyama, Y. (2006). The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. Trends in pharmacological sciences, 27(11), 587–593. https://doi.org/10.1016/j.tips.2006.09.001.

Otsuka, M., Yasuda, M., Morita, Y., Otsuka, C., Tsuchiya, T., Omote, H., & Moriyama, Y. (2005). Identification of essential amino acid residues of the NorM Na+/multidrug antiporter in Vibrio parahaemolyticus. Journal of bacteriology, 187(5), 1552–1558. https://doi.org/10.1128/JB.187.5.1552-1558.2005.

Padan, E., Kozachkov, L., Herz, K., & Rimon, A. (2009). NhaA crystal structure: functional-structural insights. The Journal of experimental biology, 212(Pt 11), 1593–1603. https://doi.org/10.1242/jeb.026708.

Pavić, A., Holmes, A., Postis, V., & Goldman, A. (2019). Glutamate transporters: a broad review of the most recent archaeal and human structures. Biochemical Society transactions, 47(4), 1197–1207. https://doi.org/10.1042/BST20190316.

Petrache, H. I. (2012). "5.2 Lipid Bilayer Structure," in Comprehensive Biophysics. Editor E. H. Egelman (Amsterdam: Elsevier), 3–15. doi:10.1016/b978-0-12-374920-8.00502-6.

Piddock L. J. (2006). Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. Clinical microbiology reviews, 19(2), 382–402. https://doi.org/10.1128/CMR.19.2.382-402.2006.

Poole K, Lomovskaya O. Can efflux inhibitors really counter resistance? Drug Discov Today: Therap Strategies. 2006;3:145-152.

Pule, C. M., Sampson, S. L., Warren, R. M., Black, P. A., van Helden, P. D., Victor, T. C., & Louw, G. E. (2016). Efflux pump inhibitors: targeting mycobacterial efflux systems to enhance TB therapy. The Journal of antimicrobial chemotherapy, 71(1), 17–26. https://doi.org/10.1093/jac/dkv316.

Putman, M., van Veen, H. W., & Konings, W. N. (2000). Molecular properties of bacterial multidrug transporters. Microbiology and molecular biology reviews : MMBR, 64(4), 672–693. https://doi.org/10.1128/MMBR.64.4.672-693.2000.

Rackovsky, S., & Scheraga, H. A. (2020). The structure of protein dynamic space. Proceedings of the National Academy of Sciences of the United States of America, 117(33), 19938–19942. https://doi.org/10.1073/pnas.2008873117.

Reading, E., Hall, Z., Martens, C., Haghighi, T., Findlay, H., Ahdash, Z., Politis, A., Booth, P.J., (2017). Interrogating membrane protein conformational dynamics within native lipid compositions. Angew. Chem. Int. Ed. Engl., 56,15654–15657.

Reid, L. M., Guzzetti, I., Svensson, T., Carlsson, A. C., Su, W., Leek, T., von Sydow, L., Czechtizky, W., Miljak, M., Verma, C., De Maria, L., & Essex, J. W. (2022). How well does molecular simulation reproduce environment-specific conformations of the intrinsically disordered peptides PLP, TP2 and ONEG?. Chemical science, 13(7), 1957–1971. https://doi.org/10.1039/d1sc03496k.

Reza, A., Sutton, J. M., & Rahman, K. M. (2019). Effectiveness of Efflux Pump Inhibitors as Biofilm Disruptors and Resistance Breakers in Gram-Negative (ESKAPEE) Bacteria. Antibiotics (Basel, Switzerland), 8(4), 229. https://doi.org/10.3390/antibiotics8040229.

Rohs R, Bloch I, Sklenar H, Shakked Z. (2005). Molecular flexibility in ab-initio drug docking to DNA: binding-site and binding-mode transitions in all-atom Monte Carlo simulations. Nucl Acids Res 33: 7048-7057.

Saier, M. H., Jr, Tran, C. V., & Barabote, R. D. (2006). TCDB: the Transporter Classification Database for membrane transport protein analyses and information. Nucleic acids research, 34(Database issue), D181–D186. https://doi.org/10.1093/nar/gkj001.

Salo-Ahen OMH, Alanko I, Bhadane R, Bonvin AMJJ, Honorato RV, Hossain S, Juffer AH, Kabedev A, Lahtela-Kakkonen M, Larsen AS, Lescrinier E, Marimuthu P, Mirza MU, Mustafa G, Nunes-Alves A, Pantsar T, Saadabadi A, Singaravelu K, Vanmeert M. (2021). Molecular Dynamics Simulations in Drug Discovery and Pharmaceutical Development. Processes; 9(1):71. https://doi.org/10.3390/pr9010071.

Salo-Ahen, O. M. H., Alanko, I., Bhadane, R., Bonvin, A. M. J. J., Honorato, R. V., Hossain, S., Juffer, A. H., et al. (2020). Molecular Dynamics Simulations in Drug Discovery and Pharmaceutical Development. Processes, 9(1), 71. MDPI AG. Retrieved from http://dx.doi.org/10.3390/pr9010071.

Santini, S., & Derreumaux, P. (2004). Helix H1 of the prion protein is rather stable against environmental perturbations: molecular dynamics of mutation and deletion variants of PrP(90-231). Cellular and molecular life sciences : CMLS, 61(7-8), 951–960. https://doi.org/10.1007/s00018-003-3455-3.

Santos, A.L., Chaves-Silva, S., Yang, L. et al. (2017). Global analysis of the MATE gene family of metabolite transporters in tomato. BMC Plant Biol 17, 185. https://doi.org/10.1186/s12870-017-1115-2.

Sargsyan, K., Grauffel, C., & Lim, C. (2017). How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. Journal of chemical theory and computation, 13(4), 1518–1524. https://doi.org/10.1021/acs.jctc.7b00028.

Schindler, B. D. & Kaatz, G. W. (2016). Multidrug efflux pumps of Gram- positive bacteria. Drug Resist. Updat. 27, 1–13.

Sharma, A., Gupta, V. K., & Pathania, R. (2019). Efflux pump inhibitors for bacterial pathogens: From bench to bedside. The Indian journal of medical research, 149(2), 129–145. https://doi.org/10.4103/ijmr.IJMR_2079_17.

Shimada, I., Ueda, T., Kofuku, Y., Eddy, M. T., & Wüthrich, K. (2019). GPCR drug discovery: integrating solution NMR data with crystal and cryo-EM structures. Nature reviews. Drug discovery, 18(1), 59–82. https://doi.org/10.1038/nrd.2018.180.

Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N. G., Hadden, J. M., Sharples, D., Sansom, M. S., Iwata, S., Henderson, P. J., & Cameron, A. D. (2010). Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. Science (New York, N.Y.), 328(5977), 470–473. https://doi.org/10.1126/science.1186303.

Shitan, N., & Yazaki, K. (2007). Accumulation and membrane transport of plant alkaloids. Current pharmaceutical biotechnology, 8(4), 244–252. https://doi.org/10.2174/138920107781387429.

Shoji, T., Inai, K., Yazaki, Y., Sato, Y., Takase, H., Shitan, N., Yazaki, K., Goto, Y., Toyooka, K., Matsuoka, K., & Hashimoto, T. (2009). Multidrug and toxic compound extrusion-type transporters implicated in vacuolar sequestration of nicotine in tobacco roots. Plant physiology, 149(2), 708–718. https://doi.org/10.1104/pp.108.132811.

Schlegel, K., Leone, V., Faraldo-Gómez, J. D., & Müller, V. (2012). Promiscuous archaeal ATP synthase concurrently coupled to Na+ and H+ translocation. Proceedings of the National Academy of Sciences of the United States of America, 109(3), 947–952. https://doi.org/10.1073/pnas.1115796109.

Schulz, S., Iglesias-Cans, M., Krah, A., Yildiz, O., Leone, V., Matthies, D., Cook, G. M., Faraldo-Gómez, J. D., & Meier, T. (2013). A new type of Na(+)-driven ATP synthase membrane rotor with a two-carboxylate ion-coupling motif. PLoS biology, 11(6), e1001596. https://doi.org/10.1371/journal.pbio.1001596.

Singh, S. K., Piscitelli, C. L., Yamashita, A., & Gouaux, E. (2008). A competitive inhibitor traps LeuT in an open-to-out conformation. Science (New York, N.Y.), 322(5908), 1655–1661. https://doi.org/10.1126/science.1166777.

Sinha, N., & Smith-Gill, S. J. (2002). Protein structure to function via dynamics. Protein and peptide letters, 9(5), 367–377. https://doi.org/10.2174/0929866023408508.

Sneha, P., & Doss, C. G. (2016). Molecular Dynamics: New Frontier in Personalized Medicine. Advances in protein chemistry and structural biology, 102, 181–224. https://doi.org/10.1016/bs.apcsb.2015.09.004.

Soto S. M. (2013). Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. Virulence, 4(3), 223–229. https://doi.org/10.4161/viru.23724.

Sousa SF, Cerqueira NM, Fernandes PA, Ramos MJ. Virtual screening in drug design and development. Comb. Chem. High Throughput Screen. 13(5), 442–453 (2010).

Spengler G, Kincses A, Gajdacs M, Amaral L. New roads leading to old destinations: efflux pumps as targets to reverse multidrug resistance in bacteria. Molecules. 2017;22:468.

Steiner-Mordoch, S., Soskine, M., Solomon, D., Rotem, D., Gold, A., Yechieli, M., Adam, Y., & Schuldiner, S. (2008). Parallel topology of genetically fused EmrE homodimers. The EMBO journal, 27(1), 17–26. https://doi.org/10.1038/sj.emboj.7601951.

Sun, J., Deng, Z., & Yan, A. (2014). Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. Biochemical and biophysical research communications, 453(2), 254–267. https://doi.org/10.1016/j.bbrc.2014.05.090.

Symersky, J., Guo, Y., Wang, J., & Lu, M. (2015). Crystallographic study of a MATE transporter presents a difficult case in structure determination with low-resolution, anisotropic data and crystal twinning. *Acta crystallographica. Section D, Biological crystallography*, *71*(Pt 11), 2287–2296. https://doi.org/10.1107/S1399004715016995.

Szewczyk P, Karyakin A, Evin M, Hong W-X, Zhang Q & Chang G (2010) Structure of a cation-bound multidrug and toxic compound extrusion transporter. Nature 467, 991–994.

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., Pulcini, C., Kahlmeter, G., Kluytmans, J., Carmeli, Y., Ouellette, M., Outterson, K., Patel, J., Cavaleri, M., Cox, E. M., Houchens, C. R., Grayson, M. L., Hansen, P., Singh, N., Theuretzbacher, U., … WHO Pathogens Priority List Working Group (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. The Lancet. Infectious diseases, 18(3), 318–327. https://doi.org/10.1016/S1473-3099(17)30753-3.

Takanashi K, Shitan N and Yazaki K (2014) The multidrug and toxic compound extrusion (MATE) family in plants. Plant Biotechnol 31, 417–430.

Takanashi, Kojiro & Shitan, Nobukazu & Yazaki, Kazufumi. (2014). The multidrug and toxic compound extrusion (MATE) family in plants. Plant Biotechnology. 31. 417-430. 10.5511/plantbiotechnology.14.0904a.

Tanaka Y, Iwaki S & Tsukazaki T (2017) Crystal Structure of a Plant Multidrug and Toxic Compound Extrusion Family Protein. Structure 25, 1455–1460.e2.

Tanaka, Y., Hipolito, C. J., Maturana, A. D., Ito, K., Kuroda, T., Higuchi, T., Katoh, T., Kato, H. E., Hattori, M., Kumazaki, K., Tsukazaki, T., Ishitani, R., Suga, H., & Nureki, O. (2013). Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. Nature, 496(7444), 247–251. https://doi.org/10.1038/nature12014.

Tanaka, Y., Iwaki, S., Sasaki, A., & Tsukazaki, T. (2021). Crystal structures of a nicotine MATE transporter provide insight into its mechanism of substrate transport. FEBS letters, 595(14), 1902–1913. https://doi.org/10.1002/1873-3468.14136.

Tang, Y., Zhu, W., Chen, K., & Jiang, H. (2006). New technologies in computer-aided drug design: Toward target identification and new chemical entity discovery. Drug discovery today. Technologies, 3(3), 307–313. https://doi.org/10.1016/j.ddtec.2006.09.004.

Tanwar, J., Das, S., Fatima, Z., & Hameed, S. (2014). Multidrug resistance: an emerging crisis. *Interdisciplinary perspectives on infectious diseases*, *2014*, 541340. https://doi.org/10.1155/2014/541340.

Tautermann C. S. (2014). GPCR structures in drug design, emerging opportunities with new structures. Bioorganic & medicinal chemistry letters, 24(17), 4073–4079. https://doi.org/10.1016/j.bmcl.2014.07.009.

Tavoulari, S., Forrest, L. R., & Rudnick, G. (2009). Fluoxetine (Prozac) binding to serotonin transporter is modulated by chloride and conformational changes. The Journal of neuroscience : the official journal of the Society for Neuroscience, 29(30), 9635–9643. https://doi.org/10.1523/JNEUROSCI.0440-09.2009.

Teilum, K., Olsen, J. G., & Kragelund, B. B. (2009). Functional aspects of protein flexibility. Cellular and molecular life sciences : CMLS, 66(14), 2231–2247. https://doi.org/10.1007/s00018-009-0014-6.

Theobald, D. L., & Miller, C. (2010). Membrane transport proteins: surprises in structural sameness. Nature structural & molecular biology, 17(1), 2–3. https://doi.org/10.1038/nsmb0110-2.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry, 31(2), 455–461. https://doi.org/10.1002/jcc.21334.

Turner, P.J. (2005). XMGRACE, Version 2.3.7. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR.

Upadhyay N, Kar D, Deepak Mahajan B, Nanda S, Rahiman R, Panchakshari N, Bhagavatula L and Datta S (2019) The multitasking abilities of MATE transporters in plants. J Exp Bot 70, 4643–4656.

Valli, M., Dos Santos RN, Figueira LD, Nakajima CH, Castro-Gamboa I, Andricopulo AD et al (2013) Development of a natural products database from the biodiversity of Brazil. J Nat Prod 76:439–444.

Vanni, S., Campomanes, P., Marcia, M., & Rothlisberger, U. (2012). Ion binding and internal hydration in the multidrug resistance secondary active transporter NorM investigated by molecular dynamics simulations. Biochemistry, 51(6), 1281–1287. https://doi.org/10.1021/bi2015184.

Verhalen, B., Dastvan, R., Thangapandian, S., Peskova, Y., Koteiche, H.A., Nakamoto, R.K., Tajkhorshid, E., McHaourab, H.S., (2017). Energy transduction and alternating access of the mammalian ABC transporter Pglycoprotein. Nature, 543, 738–741.

Veselovsky, A. V. & Ivanov, A. S. (2003). Strategy of computer-aided drug design. Curr Drug Targets Infect Disord. 3: 33–40.

von Heijne G. (2007). The membrane protein universe: what's out there and why bother?. Journal of internal medicine, 261(6), 543–557. https://doi.org/10.1111/j.1365-2796.2007.01792.x.

Whited, A. M., & Johs, A. (2015). The interactions of peripheral membrane proteins with biological membranes. Chemistry and physics of lipids, 192, 51–59. https://doi.org/10.1016/j.chemphyslip.2015.07.015.

WHO. Global Priority List of Antibiotic-Resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics. 2017. Geneva, Switzerland:WHO.

Workman, P. (2003). How much gets there and what does it do?: The need for better pharmacokinetic and pharmacodynamic endpoints in contemporary drug discovery and development. Curr Pharm Des. 9: 891–902.

Wu, E. L., Cheng, X., Jo, S., Rui, H., Song, K. C., Dávila-Contreras, E. M., … Im, W. (2014). CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. Journal of Computational Chemistry, 35(27), 1997–2004. doi:10.1002/jcc.23702.

Xu, X. J., Su, X. Z., Morita, Y., Kuroda, T., Mizushima, T., & Tsuchiya, T. (2003). Molecular cloning and characterization of the HmrM multidrug efflux pump from Haemophilus influenzae Rd. *Microbiology and immunology*, *47*(12), 937–943. https://doi.org/10.1111/j.1348-0421.2003.tb03467.x

Yazaki, K. (2005). Transporters of secondary metabolites. Curr. Opin. Plant Biol. 8, 301–307.

Yeh, V., Goode, A., & Bonev, B. B. (2020). Membrane Protein Structure Determination and Characterisation by Solution and Solid-State NMR. Biology, 9(11), 396. https://doi.org/10.3390/biology9110396.

Zakrzewska, S., Mehdipour, A. R., Malviya, V. N., Nonaka, T., Koepke, J., Muenke, C., Hausner, W., Hummer, G., Safarian, S., & Michel, H. (2019). Inward-facing conformation of a multidrug resistance MATE family transporter. Proceedings of the National Academy of Sciences of the United States of America, 116(25), 12275–12284. https://doi.org/10.1073/pnas.1904210116.

Zhang, M. Q., & Wilkinson, B. (2007). Drug discovery beyond the 'rule-of-five'. Current opinion in biotechnology, 18(6), 478–488. https://doi.org/10.1016/j.copbio.2007.10.005.

Zhang, S. (2011). Computer-aided drug discovery and development. Methods in molecular biology (Clifton, N.J.), 716, 23–38. https://doi.org/10.1007/978-1-61779-012-6_2.

Zhao, J., Xie, H., Mehdipour, A. R., Safarian, S., Ermler, U., Münke, C., Thielmann, Y., Hummer, G., Ebersberger, I., Wang, J., & Michel, H. (2021). The structure of the Aquifex aeolicus MATE family multidrug resistance transporter and sequence comparisons suggest the existence of a new subfamily.

Proceedings of the National Academy of Sciences of the United States of America, 118(46), e2107335118. https://doi.org/10.1073/pnas.2107335118.

Zhu, H.; Wu, J.; Jiang, Y.; Jin, J.; Zhou, W.; Wang, Y.; Han, G.; Zhao, Y.; Cheng, B. Genomewide analysis of MATE-type gene family in maize reveals microsynteny and their expression patterns under aluminum treatment. J. Genet. 2016, 95, 691–704.