

Structural studies of proteins conferring nutrient uptake and drug efflux

Feng Qu

Department of Life Sciences

Imperial College London

Thesis submitted for the degree of Doctor of Philosophy

2019

Abstract

Bacterial transporters play fundamental roles in nutrient acquisition and drug efflux. Determining the structures of these transport proteins is essential for understanding the mechanisms bacteria use to ensure cell survival. In *E. coli*, TauABCD is a taurine-sulfur assimilation system upregulated during conditions of sulfur starvation. The substrate-binding protein TauA is required for the initial step of taurine recognition followed by its delivery to the ABC transporter TauBC. In this work, structures of TauA in complex with taurine and three analogues have been determined. Structural comparison shows that the size of the binding site dictates substrate selectivity of TauA. Desolvation of the TauA binding site is likely to be a key determinant in the binding of taurine and its analogues. Under conditions of nutrient starvation, bacteria also produce and release antibacterial peptides like MccC7/C51. MccC7/C51 hijacks the ABC importer YejABEF to enter sensitive bacteria. However, the mechanism for MccC7/C51 internalisation and the physiological function of YejABEF have yet to be defined. In this work, the structure of YejA has been determined in a peptide-bound state, suggesting YejA is likely to be an oligopeptide-binding protein. Furthermore, interrogation of the bound peptide sequence provides a rationale for prodrug design of novel antibiotics. Finally, the AcrAB-TolC efflux system is a major reason for multidrug resistance in *E. coli*. AcrB is a major component of the system, responsible for energy transduction and substrate recognition. However, to date, no inhibitor against bacterial efflux pumps has been clinically licensed as a drug. Here, using a combination of transport and MIC assays, one compound has been highlighted as a potential AcrAB-TolC inhibitor. In addition, well diffracting crystals have been produced for AcrB in the conformation favouring ligand binding. These crystals were applied to the XChem fragment screening platform for ligand soaking. AcrB represents the first example of a membrane protein tested using this platform.

Declaration

I declare that the data described in this thesis has been produced by myself, unless stated otherwise. All sources of information within this thesis have been acknowledged. The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Table of Contents

Abstract	2
Declaration	3
Table of Contents	4
Figure list	8
Table list	11
Abbreviations	12
Acknowledgement	15
Chapter 1 Introduction	16
1.0. Introduction	17
1.0.1. Membrane transporters	17
1.0.2. MFS transporters	20
1.0.3. MATE transporters	23
1.0.4. SMR and PACE transporters	25
1.0.5. ABC transporters	26
1.0.6. RND transporters	53
1.0.7. Aims of the work	56
Chapter 2 Structural and functional studies of <i>E. coli</i> TauA	57
2.1. TauA	58
2.1.1. Nutrient starvation	58
2.1.2. TauA and SsuA	59
2.1.3. Importance and future perspectives	63
2.1.4. Aim of research	64
2.2. Materials and Methods	65
2.2.1. <i>TauA</i> cloning	65
2.2.2. TauA overexpression	66

2.2.3. TauA purification	66
2.2.4. Site-directed mutagenesis	67
2.2.5. Isothermal titration calorimetry	68
2.2.6. Crystallisation	71
2.2.7. Data collection and structure determination	72
2.2.8. TauBC protein production	73
2.2.9. Reconstitution into proteoliposomes	76
2.2.10. ATPase activity assay in liposomes	77
2.3. Results	79
2.3.1. Purification of TauA and mutants	79
2.3.2. TauA crystallisation	82
2.3.3. Structure of TauA in complex with taurine	83
2.3.4. Characterisation of the TauA binding site	92
2.3.5. TauA substrate specificity	96
2.3.6. Comparison with SsuA	103
2.3.7. Purification of TauBC	105
2.3.8. Functional characterisation of TauBC	107
2.4. Discussion	109
2.4.1. Cavity volume may determine substrate specificity	109
2.4.2. Desolvation of the binding site confers promiscuity	111
2.4.3. Is TauABC a type I or type II ABC importer?	113
2.4.4. Conclusion and outlook	114
Chapter 3 Structural studies of <i>E. coli</i> YejA	116
3.0. Introduction	117
3.0.1. Antimicrobial peptides	117
3.0.2. Microcins	118
3.0.3. Microcin genetic systems	121
3.0.4. Microcin biosynthesis and export	123
3.0.5. Microcin uptake and mechanism of action	124
3.0.6. MccJ25: mechanism for hijacking FhuA	126

3.1. YejA	128
3.1.1. MccC7/C51	128
3.1.2. MccC7/C51 hijacks OmpF and YejABEF	133
3.1.3. YejABEF	135
3.1.4. Aim of research	137
3.2. Materials and Methods	138
3.2.1. <i>yejA</i> cloning	138
3.2.2. YejA overexpression	139
3.2.3. YejA purification	139
3.2.4. Isothermal titration calorimetry	140
3.2.5. YejA crystallisation	141
3.2.6. Data collection and structure determination	142
3.2.7. Sample preparation for MS	143
3.3. Results	144
3.3.1. YejA purification	144
3.3.2. Binding of AMP to YejA	146
3.3.3. YejA crystallisation	148
3.3.4. Overall structure of YejA	149
3.3.5. YejA in complex with a polypeptide	157
3.3.6. Characterisation of the bound peptide	160
3.4. Discussion	171
3.4.1. Nature of the bound peptide	171
3.4.2. Phe8' might be a critical residue	172
3.4.3. Binding of the hendcapeptide to YejA	174
3.4.4. The “Venus fly-trap” mechanism of YejA	175
3.4.5. Conclusion and outlook	176
Chapter 4 Inhibitor screening against <i>E. coli</i> AcrAB-TolC	177
4.0. Introduction	178
4.0.1. Bacterial resistance	178
4.0.2. RND-type efflux pumps	183

4.0.3. AcrABZ-TolC multidrug efflux pump	189
4.0.4. AcrB	193
4.0.5. TolC	203
4.0.6. AcrA	206
4.0.7. AcrZ	210
4.1. Efflux pump inhibitor development	212
4.1.1. Efflux pump inhibitors	212
4.1.2. EPIs against the AcrAB-TolC system	213
4.1.3. Structure-based drug design	220
4.1.4. Aims and objectives	221
4.2. Materials and Methods	222
4.2.1. Doxorubicin efflux assay	222
4.2.2. Antibacterial activity assays	224
4.2.3. AcrB protein production	227
4.2.4. AcrB crystallisation	230
4.2.5. AcrB data collection and structure determination	232
4.2.6. XChem fragment screening	233
4.3. Results	237
4.3.1. Doxorubicin efflux assay	237
4.3.2. Screening of the NCI diversity set V library	241
4.3.3. Potentiation of antibiotics	247
4.3.4. AcrB purification	250
4.3.5. AcrB crystallisation and structure	252
4.3.6. Application of XChem for AcrB	258
4.4. Discussion	263
5. Conclusion and outlook	266
6. Reference	267

Figure list

Figure 1.0.1.1.	Structures of bacterial membrane transporters	19
Figure 1.0.2.1.	Structure of <i>E. coli</i> LacY	21
Figure 1.0.2.2.	Alternating access model for the H ⁺ /galactoside symporter LacY	22
Figure 1.0.3.1.	Structure of <i>V. cholerae</i> NorM	24
Figure 1.0.5.1.	Characteristic architecture of ABC transporters	28
Figure 1.0.5.2.	Conserved motifs in the NBDs	30
Figure 1.0.5.3.	Model for MsbA-mediated LPS flipping	35
Figure 1.0.5.4.	Alternating access mechanism for ABC exporters	37
Figure 1.0.5.5.	Crystal structures of representative ABC importers	39
Figure 1.0.5.6.	The “Venus fly-trap” mode of ligand binding	41
Figure 1.0.5.7.	Different clusters of SBPs with their distinct features in orange	45
Figure 1.0.5.8.	Transport mechanism of the Type I ABC importer MalFGK ₂ -MBP	48
Figure 1.0.5.9.	Transport mechanism of the Type II ABC importer BtuCD-F	51
Figure 1.0.6.1.	The AcrA ₆ B ₃ Z ₃ -TolC ₃ efflux pump	55
Figure 2.1.2.1.	The <i>ssuEADCB</i> and <i>tauABCD</i> operons from <i>E. coli</i>	60
Figure 2.1.2.2.	Taurine uptake and oxidation pathway	62
Figure 2.2.1.1.	Schematic representation of the His ₆ -TEV-TauA construct	65
Figure 2.2.5.1.	Taurine and its analogues	69
Figure 2.2.5.2.	Equations for calculation of K _d and ΔS	70
Figure 2.2.8.1.	Schematic representation of the His ₆ -TEV-TauBC construct	74
Figure 2.2.10.1.	Equations for calculation of the rate of ATPase activity	78
Figure 2.3.1.1.	Purification of TauA	80
Figure 2.3.1.2.	Purification of TauA mutants	81
Figure 2.3.2.1.	TauA-taurine crystals	82
Figure 2.3.3.1.	TauA anomalous difference maps	84
Figure 2.3.3.2.	X-ray diffraction pattern for TauA-taurine crystal at 1.31 Å	85
Figure 2.3.3.3.	Stereo view of electron density maps for taurine	86
Figure 2.3.3.4.	Structure of taurine-bound TauA	88
Figure 2.3.3.5.	Structural comparison of TauA and SsuA	90
Figure 2.3.3.6.	Sequence alignment of TauA and SsuA	91
Figure 2.3.4.1.	The taurine-binding site of TauA	93
Figure 2.3.4.2.	Calorimetric titration for taurine binding to TauA and mutants	95
Figure 2.3.5.1.	Taurine and its analogues	97
Figure 2.3.5.2.	Calorimetric titration for the binding of different ligands to TauA	98
Figure 2.3.5.3.	Electron density maps for 2-AEP, ACES and MES	99
Figure 2.3.5.4.	TauA in complex with taurine and its analogues	101
Figure 2.3.6.1.	The alkanesulfonate-binding sites of TauA and SsuA	104
Figure 2.3.7.1.	Purification of TauBC	106
Figure 2.3.8.1.	TauBC ATPase activity in detergent	107

Figure 2.3.8.2.	TauBC ATPase activity in proteoliposomes	108
Figure 2.4.1.1.	Cavity volumes of TauA and SsuA	109
Figure 2.4.1.2.	Comparison between MBP and GBP	110
Figure 2.4.2.1.	Structure of tetrapeptide-bound OppA from <i>S. typhimurium</i>	112
Figure 3.0.2.1.	Representative examples of different microcin classes	120
Figure 3.0.3.1.	Genetic organisation of the microcin systems	122
Figure 3.0.5.1.	MccJ25 and MccB17 uptake and mechanism of action	125
Figure 3.0.6.1.	Structure of MccJ25	126
Figure 3.1.1.1.	Structures of MccC7/C51 and processed MccC7/C51*	129
Figure 3.1.1.2.	MccC7/C51 maturation and export	132
Figure 3.1.2.1.	MccC7/C51: uptake and mechanism of action	134
Figure 3.1.3.1.	Organisation of the <i>yefABEFG</i> operon from <i>E. coli</i>	135
Figure 3.2.1.1.	Schematic representation of the His ₆ -TEV-YejA construct	138
Figure 3.3.1.1.	Purification of YejA	145
Figure 3.3.2.1.	Calorimetric titration for the binding of AMP to YejA	147
Figure 3.3.3.1.	YejA crystals	148
Figure 3.3.4.1.	X-ray diffraction pattern for YejA crystal at 1.5 Å	149
Figure 3.3.4.2.	Structure of YejA in complex with an oligopeptide	150
Figure 3.3.4.3.	Electron density maps for the co-crystallised oligopeptide	151
Figure 3.3.4.4.	Top view of AppA and OppA	154
Figure 3.3.4.5.	Structures of YejA homologues	155
Figure 3.3.4.6.	Sequence alignment of YejA from <i>E. coli</i> and its homologues	156
Figure 3.3.5.1.	Peptide binding to YejA	159
Figure 3.3.6.1.	The MALDI mass spectra of the YejA protein mixture	161
Figure 3.3.6.2.	The MALDI mass spectra of BSA and angiotensin II	162
Figure 3.3.6.3.	The MALDI mass spectra of the matrix	163
Figure 3.3.6.4.	The nLC-ESI mass spectra of the YejA protein mixture	164
Figure 3.3.6.5.	The nLC-ESI mass spectra of BSA and angiotensin II	165
Figure 3.3.6.6.	The MALDI MS/MS fragmentation of the peptide at 1426.7 m/z	166
Figure 3.3.6.7.	Sequence of an identified YejA peptide	167
Figure 3.3.6.8.	The MALDI mass spectra of the YejA sample with spin filtration	168
Figure 3.3.6.9.	The MALDI mass spectra of sample dissolved from YejA crystals	169
Figure 3.4.2.1.	Comparison of peptides binding to YejA, AppA and OppA	173
Figure 4.0.1.1.	Mechanisms of antibiotic resistance	180
Figure 4.0.1.2.	Horizontal gene transfer between bacteria	182
Figure 4.0.2.1.	The AcrA ₆ B ₃ Z ₃ -TolC ₃ efflux pump	186
Figure 4.0.3.1.	Examples of substrates of AcrAB-TolC	192
Figure 4.0.4.1.	Structure of the AcrB monomer	194
Figure 4.0.4.2.	Crystal structure of the asymmetric AcrB in a ligand-bound state	196
Figure 4.0.4.3.	Entrances of the AcrB drug-transport channel	199
Figure 4.0.4.4.	Proximal and distal binding pockets of AcrB	201
Figure 4.0.4.5.	Functional rotation mechanism mediated by AcrB	202
Figure 4.0.5.1.	Structure of the TolC trimer	205

Figure 4.0.6.1.	Structure of the AcrA monomer	208
Figure 4.0.6.2.	Interactions between components of the AcrA ₆ B ₃ -TolC ₃ complex	209
Figure 4.0.7.1.	Interactions between AcrB and AcrZ	211
Figure 4.1.2.1.	EPIs exemplified by PAβN, NMP, ABI-PP and MBX2319	215
Figure 4.1.2.2.	Crystal structure of AcrB in complex with ABI-PP	218
Figure 4.1.2.3.	Binding of MBX2319 to AcrB periplasmic domain	219
Figure 4.2.2.1.	Layout of the MIC assay using the broth microdilution method	225
Figure 4.2.6.1.	Graphic illustration of the XChem fragment screening pipeline	236
Figure 4.3.1.1.	Doxorubicin efflux and accumulation	238
Figure 4.3.1.2.	The effect of AcrB deletion on doxorubicin accumulation	240
Figure 4.3.2.1.	Structures of the identified inhibitors	244
Figure 4.3.2.2.	Compounds 1-6 inhibit doxorubicin efflux	245
Figure 4.3.2.3.	Fluorescence quenching compared with the controls	246
Figure 4.3.4.1.	SDS-PAGE for AcrB purification	251
Figure 4.3.5.1.	Optimised AcrB crystals	253
Figure 4.3.5.2.	X-ray diffraction pattern for AcrB crystal at 2.57 Å	254
Figure 4.3.5.3.	Structure of the apo asymmetric AcrB	256
Figure 4.3.5.4.	Electron density of the AcrB binding pockets	256
Figure 4.3.5.5.	Electron density of the AcrB trimer	257

Table list

Table 1.0.1.1.	Mechanistic comparisons between type I and II ABC importers	52
Table 2.2.1.1.	Primer sequences for <i>tauA</i> gene	65
Table 2.2.4.1.	Primer sequences for TauA mutants	68
Table 2.2.5.1.	Concentrations of TauA, TauA mutants and ligands tested in ITC	70
Table 2.2.8.1.	Primer sequences for <i>tauBC</i> gene	73
Table 2.2.10.1.	Components of the ATPase assay using the EnzChek assay kit	77
Table 2.3.3.1.	Data collection and refinement statistics for TauA crystals	87
Table 2.3.4.1.	Thermodynamic parameters for taurine binding to TauA and mutants	94
Table 2.3.5.1.	Thermodynamic parameters for taurine and analogues binding to TauA	97
Table 2.3.5.2.	MD simulation of water distributions in the binding site of TauA	102
Table 3.0.2.1.	Structural characteristics of secreted microcins	119
Table 3.0.3.1.	Typical gene clusters characterising microcins of different classes	121
Table 3.2.1.1.	Primer sequences for <i>yefA</i> gene	138
Table 3.3.4.1.	Data collection and refinement statistics for YefA crystals	152
Table 4.0.2.1.	The 10 families of the RND superfamily	184
Table 4.2.1.1.	Libraries used in chapter 4	223
Table 4.2.3.1.	Buffers for AcrB purification	229
Table 4.3.2.1.	Chemical names, formulas and MWs of the identified compounds	243
Table 4.3.3.1.	Potentiation effect on different antibiotics	249
Table 4.3.5.1.	Data collection and refinement statistics for AcrB crystals	255
Table 4.3.6.1.	Results from the XChem fragment screening	262

Abbreviations

2-AEP	2-aminoethylphosphonic acid
3D	Three-dimensional
Å	Angstrom
ABC	Adenosine triphosphate-binding cassette
ABI-PP	[[2-({[(3R)-1-{8-{{(4-tert-butyl-1,3-thiazol-2-yl)amino}carbonyl}-4-oxo-3-[(E)-2-(1H-tetrazol-5-yl)vinyl]-4H-pyrido[1,2-a]pyrimidin-2-yl}piperidin-3-yl)oxy}carbonyl}amino)ethyl](dimethyl ammonio)acetate
ACES	N-(2-acetamido)-2-aminoethanesulfonic acid
AcrB	Acriflavine resistance protein B
ADA	N-(2-acetamido)iminodiacetic acid
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
AMP-PNP	Adenosine 5'-(β,γ -imido) triphosphate
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
CA-MHB2	Mueller-Hinton broth 2
CFU	Colony forming units
CMC	Critical micelle concentration
CPM	N-[4-(7-diethylamino-4-methyl-3-coumarinyl) phenyl] maleimide
CV	Column volume
Cymal-6	6-Cyclohexyl-1-Hexyl- β -D-Maltoside
Da	Daltons
DDM	n-Dodecyl- β -D-Maltoside
DDS	Dodecanoyl Sucrose
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
ECF	Energy-coupling factor

EPI	Efflux pump inhibitor
ESI	Electrospray ionization
g	Gravitational force; 9.81 ms^{-2}
GBP	Glucose-binding protein
GPCR	G-protein coupled receptor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMMD	High molecular mass drugs
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
kb	Kilo bases
kDa	Kilodaltons
kpsi	Kilo pound-force per square inch
LB	Lysogeny broth
LMMD	Low molecular mass drug
LPS	Lipopolysaccharide
MALDI	Matrix-assisted laser desorption ionization
MATE	Multidrug and toxic compound extrusion
MBP	Maltose-binding protein
MD	Molecular dynamics
MES	2-(N-morpholino) ethanesulfonic acid
MESG	2-amino-6-mercapto-7-methylpurine riboside
MFP	Membrane fusion protein
MFS	Major facilitator superfamily
MIC	Minimum inhibitory concentration
MOPS	3-(N-morpholino) propanesulfonic acid
MS	Mass spectrometry
MW	Molecular weight
NBD	Nucleotide binding domain
Ni-NTA	Nickel nitrilotriacetic acid
nLC	Nano-scale liquid chromatography
NMP	1-(1-naphthylmethyl)piperazine
NMR	Nuclear magnetic resonance

OD	Optical density
OMF	Outer membrane factor
PAβN	Phe-Arg-β-naphthylamide
PACE	Proteobacterial antimicrobial compound efflux
PanDDA	Pan-Dataset Density Analysis
PBP	Penicillin-binding protein
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Poly(ethylene) glycol
Pi	Inorganic phosphate
PNP	Purine nucleoside phosphorylase
RFU	Relative fluorescent unit
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RND	Resistance nodulation cell division
SAD	Single-wavelength anomalous diffraction
SBD	Substrate-binding domain
SBDD	Structure-based drug design
SBP	Substrate-binding protein
SDS	Sodium dodecylsulfate
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
smFRET	Single-molecule Förster resonance energy transfer
SMR	Small multidrug resistance
SPR	Surface plasmon resonance
MS/MS	Tandem mass spectrometry
TEV	Tobacco Etch Virus
TM	Transmembrane
TSA	Thermal stability assay
TMD	Transmembrane domain
tRNA	Transfer ribonucleic acid

Acknowledgement

I would like to thank my supervisor Dr. Konstantinos Beis for the guidance and supervision during my PhD. I would also like to thank Kiran, Maria, Indran, Jason for being the best colleagues and friends in the lab. I would like to send my regards to Dr. Armin Wagner, Dr. Kamel El Omari and Dr. Ramona Duman for the help with the iodide-SAD experiment, Dr. Alfonso de Simone for the MD simulation, and Prof. Frank von Delft and Dr. Alice Douangamath for the instruction and training on the XChem fragment screening platform.

My PhD was funded by China Scholarship Council (CSC).

Chapter 1

Introduction

1.0. Introduction

1.0.1. Membrane transporters

Membrane proteins are large biological molecules that interact with, or form part of cell membranes (Alberts *et al.*, 2013). The cell membrane is a hydrophobic lipid bilayer that physically separates the interior of cells (*i.e.* nucleus and organelles) from the outside environment (Stillwell, 2013). The cell membrane is selectively permeable and only allows the exchange of certain substances and signals inside or outside cells (Yeagle, 2016). Movement of chemicals across cell membranes can occur via either passive diffusion, a process that does not rely on energy input (*e.g.* exchange of CO₂ and O₂), or active transport, which requires specific transporters and expends energy to drive the transport process (*e.g.* transport of sugars, amino acids and ions).

Bacterial membrane transporters enable bacteria to extrude unwanted metabolites or toxins, acquire nutrients, and maintain intracellular levels of salts and protons (Alberts *et al.*, 2013). These substrates are typically hydrophilic and cannot diffuse through the membrane phospholipid bilayer. Substrates of such transport systems show high structural and chemical diversity, including metals, sugars, amino acids, peptides and macromolecules. However, the transport proteins do not display such diversity in their structures or chemistry. In fact, based on their phylogeny and function, these transporters have been classified into three systems: (1) channels and pores, (2) primary and (3) secondary active transporters (Reddy *et al.*, 2016). Each of these systems is composed of different transporter superfamilies or families in which a common structure or mechanism can be observed. Members of the first class, typically α -helical ion channels and β -barrel porins, permit the rapid passage of substrates across the

membrane upon their opening and closing. Membrane transporters can be classified into two mechanistically distinct systems: primary transporters that utilise a primary energy source such as adenosine triphosphate (ATP) and secondary transporters that are energised by a secondary energy source such as a proton or ion gradient (Du *et al.*, 2018). One example of primary active transporters is the ATP-binding cassette (ABC) transport proteins, which are ubiquitously distributed in archaea, prokaryotes and eukaryotes. Members of the ABC superfamily include the bacterial ABC importer MalFGK₂-MBP and exporter Sav1866 (Figure 1.0.1.1). ABC transporters possess diverse physiological roles including nutrient uptake and drug efflux. Secondary transporters can be classified into five distinct families. The resistance nodulation cell division (RND) transporters constitute the major multidrug efflux pumps in Gram-negative bacteria and demonstrate the broadest substrate spectrum. The best studied example of an RND-type transporter is AcrB (Figure 1.0.1.1), which is a major and constitutively expressed multidrug efflux pump in *Escherichia coli*, with its homologues present in all Gram-negative bacteria (*e.g.* *Pseudomonas aeruginosa* MexB). The major facilitator superfamily (MFS) consist of importers for uptake of nutrients (*e.g.* *E. coli* LacY) and exporters for efflux of toxins (*e.g.* *E. coli* EmrD) (Figure 1.0.1.1). The multidrug and toxic compound extrusion (MATE) family exporters confer resistance to cationic drugs, secondary metabolites and xenobiotics (*e.g.* PfMATE of *Pyrococcus furiosus* (Figure 1.0.1.1) and NorM of *Vibrio cholerae*). Members of the MATE family are driven by either the H⁺/drug or Na⁺/drug antiport mechanism, while secondary transporters of other families mostly utilise the proton motive force. The small multidrug resistance (SMR)-type H⁺/drug antiporters contribute to bacterial resistance to lipophilic and cationic drugs (*e.g.* *E. coli* EmrE (Figure 1.0.1.1)). The recently identified proteobacterial antimicrobial compound efflux (PACE) transporters contribute to resistance against a range of different synthetic bactericides such as chlorhexidine and acriflavine in

Gram-negative bacteria (e.g. AceI of *Acinetobacter baumannii* and VP1155 of *Vibrio parahaemolyticus*).

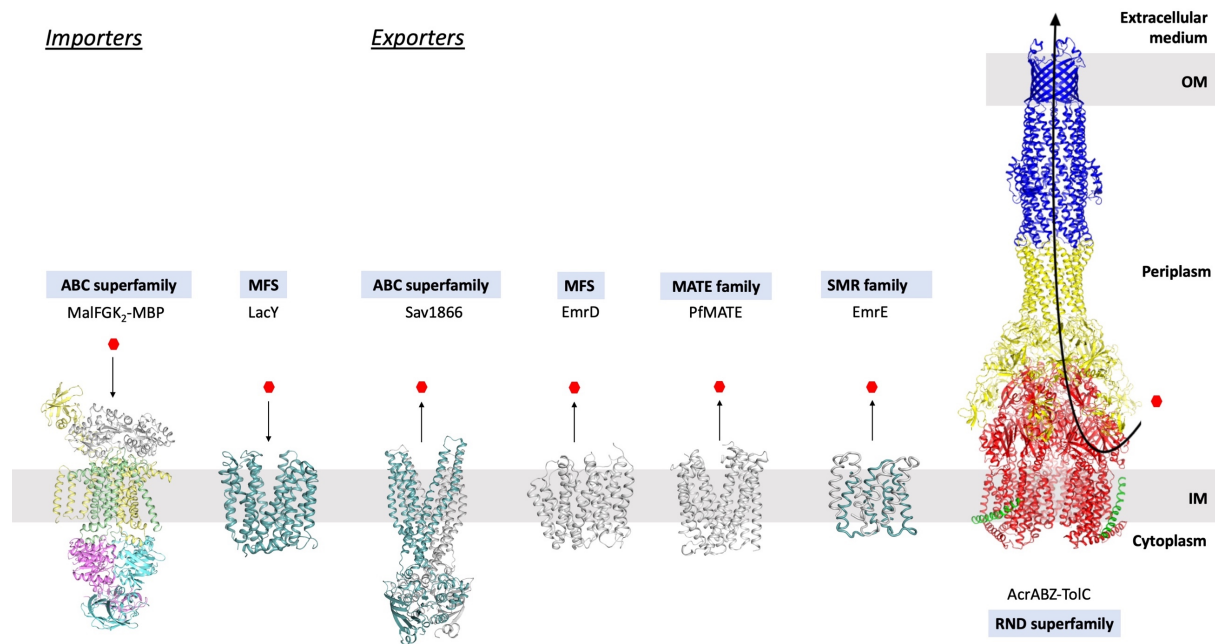


Figure 1.0.1.1. Structures of bacterial membrane transporters. Primary transporters utilise a primary energy source such as ATP, while secondary transporters are energised by a secondary energy source such as a proton or ion gradient. Representatives of each transporter family are presented, including the ABC importer MalFGK₂-MBP (PDB ID: 2R6G, (Oldham *et al.*, 2007)), the MFS importer LacY (PDB ID: 5GXB, (Jiang *et al.*, 2016)), the ABC exporter Sav1866 (2HYD, (Dawson & Locher, 2006)), the MFS exporter EmrD (PDB ID: 2GFP, (Yin *et al.*, 2006)), the MATE transporter PfMATE (PDB ID: 3VVN, (Tanaka *et al.*, 2013)), the SMR transporter EmrE (PDB ID: 3B5D, (Chen *et al.*, 2007)) and the RND efflux pump AcrABZ-TolC (PDB ID: 5O66, (Wang *et al.*, 2017)). Adapted from (Du *et al.*, 2018).

1.0.2. MFS transporters

MFS transporters constitute the largest and most diverse family of secondary transporters (Pao *et al.*, 1998). MFS transporters generally utilise the proton motive force for substrate transport. MFS transporters can be divided into importers (*e.g.* *E. coli* LacY, *Arabidopsis thaliana* NRT1.1, human GLUT1) and exporters (*e.g.* *E. coli* EmrD, MdfA and YajR), with the former involved in uptake of nutrients (*e.g.* sugars, nitrate, organophosphate, oligosaccharide) and the latter efflux of toxins (*e.g.* antibiotics, Krebs cycle metabolites) (Yin *et al.*, 2006; Jiang *et al.*, 2013; Heng *et al.*, 2015; Sun *et al.*, 2014; Kumar *et al.*, 2014; Deng *et al.*, 2014). Members of the MFS family include uniporters (that transport substrates across the membrane without energy coupling), symporters (where substrate transport is coupled to and in the same direction as ion transport) and antiporters (where the coupled transport of substrates and ions occur in opposite directions). MFS transporters predominantly function as monomers. Most members of the MFS family are 400-600 amino acid residues in length and contain 12 or 14 transmembrane (TM) helices. MFS transporters show a high level of structural conservation, with the structural architecture consisting of two domains, each composed of six TM helices. The two domains display pseudo-twofold symmetry and are connected by either a cytoplasmic loop or two TM helices (Sun *et al.*, 2014; Kumar *et al.*, 2014; Deng *et al.*, 2014). Structural and biochemical studies suggest an alternating access mechanism for MFS transporters, where the two halves of the protein undergo conformational changes between inward- and outward-facing states. Conformational changes in most members of the MFS family are associated with protonation and deprotonation (*e.g.* LacY (Guan & Kaback, 2004) and YajR (Jiang *et al.*, 2013)). For these MFS transporters, the outward-facing and inward-facing conformations are generally considered as the ground and excited states, respectively. Energised by protonation of the central cavity, MFS transporter transits from the ground state to the excited state, whereas

the transition from the excited state to the ground state is automatically driven by conformational energy stored in the excited state. The substrate-binding pocket is located halfway into the lipid bilayer and is exposed to either the cytoplasmic or periplasmic side of cell membrane.

LacY is a well-studied prototype of an MFS symporter and is responsible for galactoside transport in *E. coli* (Kaback *et al.*, 2001). It functions as a monomer consisting of 12 TM helices spanning the bacterial inner membrane (Figure 1.0.2.1) (Abramson *et al.*, 2003). The N- and C-terminal halves of LacY constitute two six-TM-helix bundles that are linked by a long cytoplasmic loop between TM6 and TM7. The two domains exhibit the same topology and display a pseudo-twofold symmetry. The internal hydrophilic cavity of LacY is formed between the N-terminal helices TM1, TM2, TM4 and TM5 and the C-terminal helices TM7, TM8, TM10 and TM11, with the remaining helices embedded in the lipid bilayer (Abramson *et al.*, 2003). The substrate-binding pocket lies in the hydrophilic cavity halfway into the inner membrane and is in proximal to the pseudo-twofold axis of LacY (Abramson *et al.*, 2003).

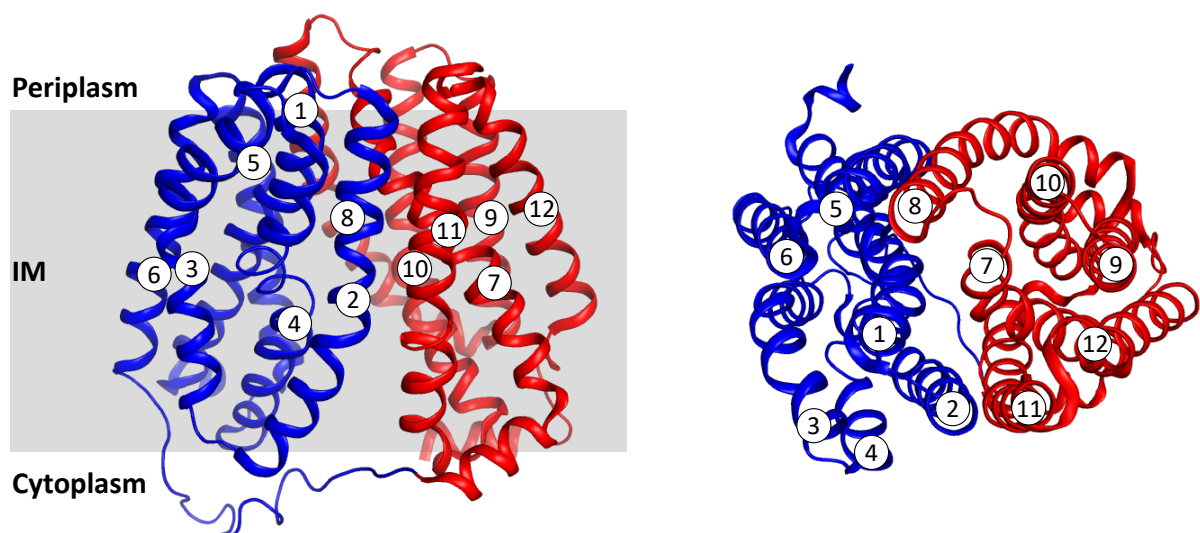


Figure 1.0.2.1. Structure of *E. coli* LacY. The circled numbers indicate the different TM helices TM1-TM12. The N- and C-terminal halves of LacY are coloured in blue and red, respectively. (PDB ID: 1PV6, (Abramson *et al.*, 2003))

LacY is energised by the proton motive force and functions via an alternating access mechanism, where LacY undergoes a conformational change during the transport cycle to expose the substrate-binding site to alternate sides of cell membrane (Figure 1.0.2.2) (Kaback *et al.*, 2007; Smirnova *et al.*, 2007; Majumdar *et al.*, 2007; Zhou *et al.*, 2008; Smirnova *et al.*, 2009). Following protonation of LacY (state 1 to 2), the binding of galactoside (state 2 to 3) induces a large conformational change of LacY and the dimer switches from an outward-facing (states 1-3) to an occluded conformation (state 4). Then, the LacY dimer opens up to the cytoplasm (state 4 to 5), releases the substrate (state 5 to 6), deprotonates (state 6 to 7) and converts to an apo-occluded conformation (state 8), followed by switching to an outward-facing state (state 1).

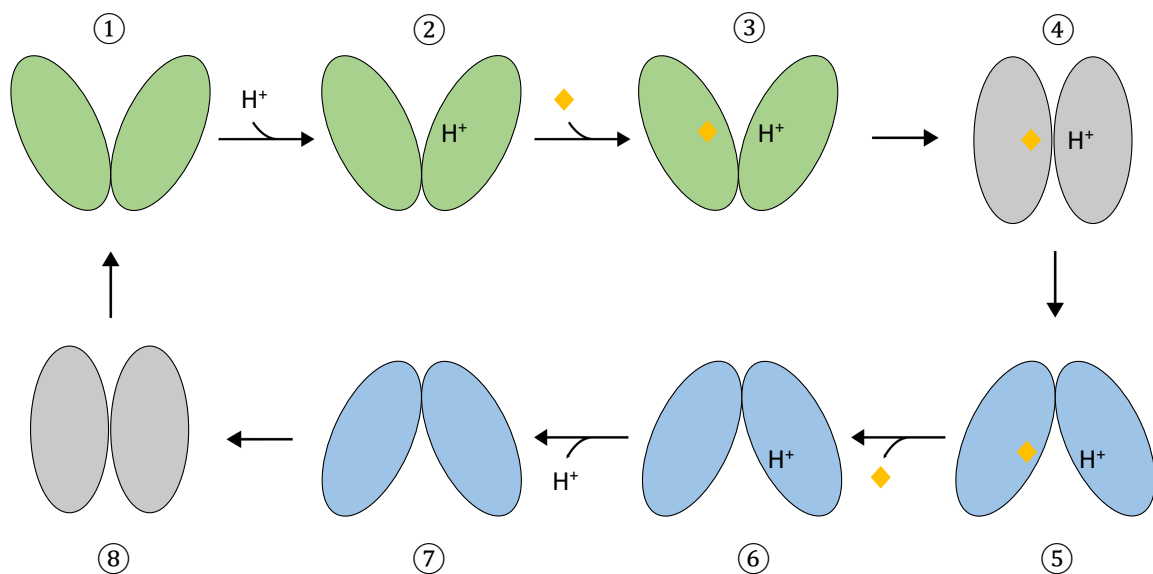


Figure 1.0.2.2. Alternating access model for the H⁺/galactoside symporter LacY. The galactoside molecule is indicated by the yellow diamond. Following protonation of LacY (state 1 to 2), the binding of galactoside (state 2 to 3) induces a conformational change from an outward-facing (states 1-3) to an occluded state (state 4). Then, LacY undergoes a conformational change and switches to an inward-facing state (state 5), followed by release of substrate (state 5 to 6) and deprotonation (state 6 to 7). At this stage, LacY switches to an apo-occluded conformation and then converts to an outward-facing state (state 1), ready for the next cycle. Transition from outward-facing (state 3) to inward-facing (state 5) is a proton motive force-dependent step. Adapted from (Du *et al.*, 2018).

1.0.3. MATE transporters

The MATE family represents a group of secondary antiporters that utilise a Na^+ and/or H^+ gradient for substrate transport (Omote *et al.*, 2006; Kuroda & Tsuchiya, 2009). Members of the MATE family have been found in all kingdoms of life and are involved in the efflux of cationic drugs, secondary metabolites and xenobiotics. Examples of MATE transporters include the bacterial MATEs, NorM of *V. cholerae* (Morita *et al.*, 1998) and PfMATE of *P. furius* (Tanaka *et al.*, 2013), the fungal MATE Erc1 of *Saccharomyces cerevisiae* (Shiomi *et al.*, 1991), the plant MATE ALF5 of *A. thaliana* (Diener *et al.*, 2001) and the mammalian MATE MATE1 of *Homo sapiens* (Otsuka *et al.*, 2005). Bacterial and mammalian MATE transporters also mediate multidrug resistance which reduces the efficacy of clinical drugs used in diseases such as bacterial infections and diabetes (McAleese *et al.*, 2005; Becker *et al.*, 2009).

All the MATE transporters share ~40% sequence similarity, suggesting an overall conserved structural fold and transport model (Omote *et al.*, 2006). MATE transporters range from 400-700 amino acid residues in length and are arranged in 12 TM helices. Similar to MFS transporters, MATE transporters predominantly function as monomers consisting of two six-helix domains. MATE transporters display pseudo-twofold symmetry between the N- and C-terminal domains, with the central cavity located at the domain interface (He *et al.*, 2010; Tanaka *et al.*, 2013). MATE transporters may also follow an alternating access mechanism where the proteins undergo conformational changes during the transport cycle and the substrate-binding site is exposed to alternate sides of cell membrane. MATE transporters couple the import of Na^+ or H^+ to the export of substrates against the electrochemical gradient.

The crystal structure of the MATE transporter NorM from *V. cholerae* has been determined in an outward-facing conformation (Figure 1.0.3.1) (He *et al.*, 2010). NorM is an Na^+ /drug

antiporter that confers resistance to a wide range of antibiotics including norfloxacin, ofloxacin, kanamycin, ciprofloxacin, streptomycin and ethidium bromide. NorM consists of 12 TM helices, which are arranged as two bundles (TM1-TM6 and TM7-TM12) connected by a long cytoplasmic loop between TM6 and TM7 (residues 218-232). The two halves display two-fold symmetry, suggesting that NorM might have evolved from gene duplication (He *et al.*, 2010). An extra helix after TM12 lies horizontally under the cytoplasmic side of TM11 and TM12. The internal cavity is located in between the two TM bundles and is contributed to by all the TM helices except for TM3 and TM9. Amongst all known transporters, topology of MATE transporters is unique as within each domain, none of the TM helices intersect with any three consecutive TM helices (Figure 1.0.3.1) (He *et al.*, 2010).

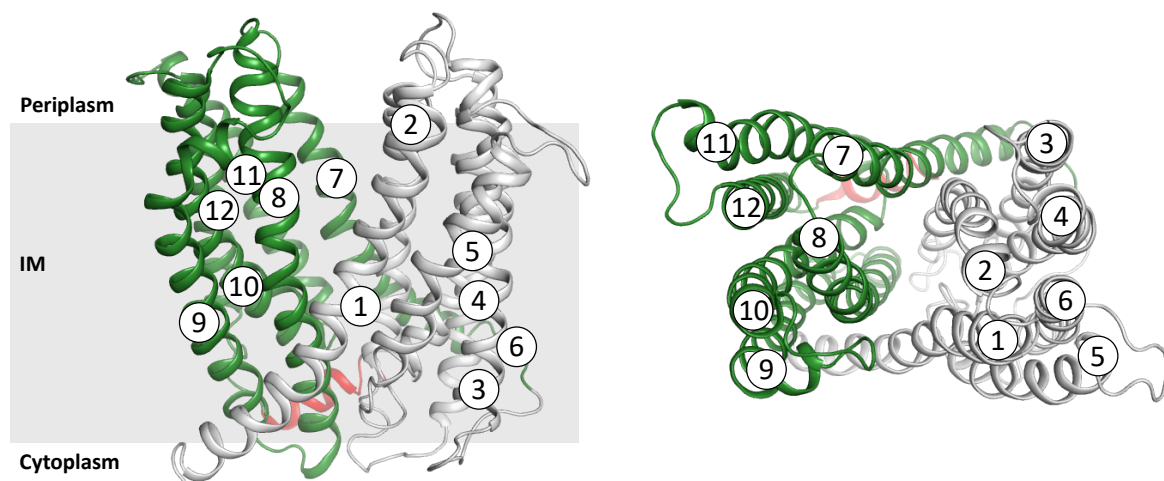


Figure 1.0.3.1. Structure of *V. cholerae* NorM. The circled numbers indicate the different TM helices TM1-TM12. The N- and C-terminal halves of the transporter are coloured in grey and green, respectively. The extra horizontal helix after TM12 is coloured in red. (PDB ID: 3MKT, (He *et al.*, 2010))

1.0.4. SMR and PACE transporters

The SMR family contains inner-membrane transporters that confer resistance to a wide range of toxic compounds including quaternary ammonium and cations in archaea and bacteria (Bay *et al.*, 2008). Members of the SMR family typically range between 100-200 amino acids in length with four TM helices, and function as homodimers or heterodimers (Figure 1.0.1.1).

E. coli EmrE is a H⁺/drug SMR antiporter responsible for the efflux of toxic cations including ethidium, methyl viologen, proflavine and antibiotics such as tetracycline and erythromycin (Yerushalmi *et al.*, 1995). Cryo-EM, X-ray crystallography and NMR have revealed an antiparallel arrangement of EmrE monomers which constitute an asymmetric dimer (Fleishman *et al.*, 2006; Chen *et al.*, 2007; Morrison *et al.*, 2011). Each monomer contributes four TM helices, which are organised in a pairwise manner at the dimeric interface. TM1-TM3 interact with TM1-TM3 from the opposite protomer, forming a substrate-binding cavity, while the two TM4s are mainly involved in dimer formation. During the transport cycle, EmrE alternates between the inward- and outward-facing conformations. EmrE in the two conformations are identical except for their orientation in the inner membrane (Morrison *et al.*, 2011).

Identified as a novel family of multidrug transporters, AceI and its homologue VP1155 have been shown to contribute to resistance against a range of different synthetic bactericides such as chlorhexidine and acriflavine in proteobacteria (Hassan *et al.*, 2013; Hassan *et al.*, 2015). AceI is similar to SMR transporters in terms of size, secondary structure and the presence of a conserved glutamate residue in TM1 (Hassan *et al.*, 2013). PACE transporters are absent in *E. coli*, suggesting that their encoding genes might have been lost early during evolution.

1.0.5. ABC transporters

ABC transporters represent one of the largest transporter superfamilies whose members are found in all kingdoms of life (Ames *et al.*, 1992). These transporters utilise ATP hydrolysis as an energy source to transport molecules across the plasma membrane (Rees *et al.*, 2009). Depending on the direction of substrate transport, ABC transporters can be classified into the exporters that are expressed in all kingdoms and the importers that are present solely in archaea, bacteria and plants (Lee *et al.*, 2008; Kang *et al.*, 2010).

ABC transporters possess diverse physiological roles including uptake of nutrients or vitamins (Oldham *et al.*, 2007; Korkhov *et al.*, 2012; Xu *et al.*, 2013), efflux of drugs or toxins (Dawson & Locher, 2006; Fitzpatrick *et al.*, 2017), mediation of pathogenic virulence (Stergiopoulos *et al.*, 2003) and flipping of bacterial membrane components such as lipopolysaccharide (LPS) (Mi *et al.*, 2017) and oligosaccharide (Perez *et al.*, 2015). Substrates of ABC transporters include antibiotics, ions, vitamins, sugars, drugs, metabolites, amino acids, lipids, peptide and proteins. It has been shown that the dysfunction of certain ABC transporters is associated with multidrug resistance and diseases such as Tangier disease and cystic fibrosis (Thomas & Tampé, 2018). As such, ABC transporters represent an important class of therapeutic targets in drug development.

1.0.5.1. Architecture of ABC transporters

The characteristic architecture of ABC transporters consists of two transmembrane domains (TMDs) that constitute the substrate-binding site and two nucleotide-binding domains (NBDs) for ATP binding and hydrolysis (Rees *et al.*, 2009). The TMDs are embedded in the inner-membrane bilayer and the NBDs are located in the cytoplasm. ABC transporters can exist as homodimers where the NBD and TMD form a single polypeptide (*e.g.* Sav1866 from *S. aureus*, (Dawson & Locher, 2006)) or two polypeptides (*e.g.* BtuCD from *E. coli* (Locher *et al.*, 2002)), or heterodimers where each half of the NBD or TMD domains is encoded by distinct genes (*e.g.* MalFGK₂ from *E. coli* (Oldham *et al.*, 2007)). The NBDs contain highly conserved motifs among different ABC transporters of various origins, while the TMDs differ in sequence and structural topology to ensure substrate selectivity and account for the chemical and structural diversity of substrates. Conformational changes of the NBDs are transmitted to the TMDs by a pair of coupling α -helices in the cytoplasmic loops of the TMDs. ABC importers usually have an additional substrate-binding protein (SBP) or covalently-linked substrate-binding domain (SBD) that captures substrates from the periplasm and delivers them to the TMDs for translocation (Scheepers *et al.*, 2016). The characteristic architecture of ABC transporters is shown in Figure 1.0.5.1.

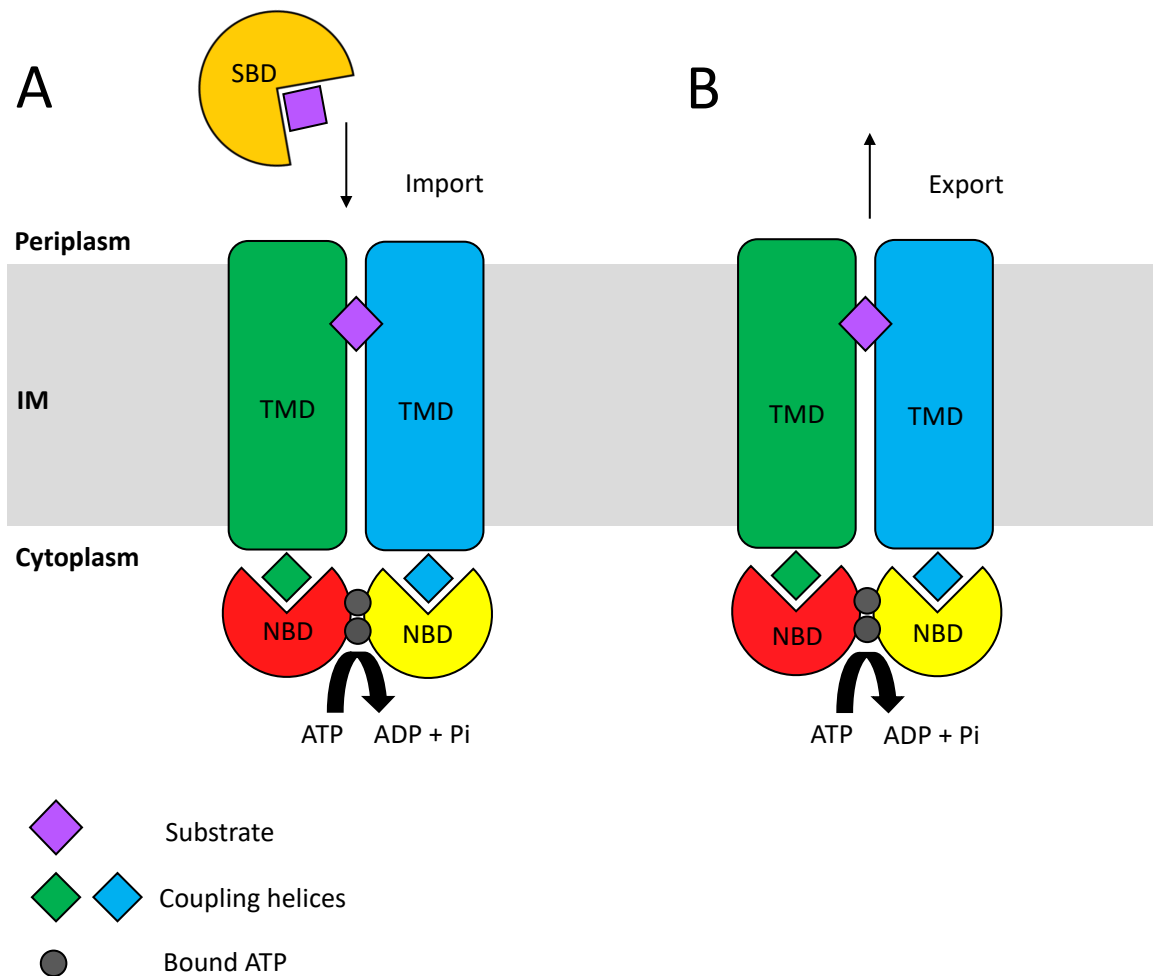


Figure 1.0.5.1. Characteristic architecture of ABC transporters. (A) ABC importers. (B) ABC exporters. The characteristic architecture of ABC transporters consists of two TMDs and two NBDs. Most ABC importers require an additional SBP/SBD to deliver substrate to the TMDs for transport whilst ABC exporters can recruit ligands directly. Communication between the NBDs and TMDs is mediated by a pair of α -helical coupling helices located in the cytoplasmic loops of the TMDs. ABC transporters harness the energy of ATP hydrolysis to translocate substrate across the inner membrane. Adapted from (Locher, 2016).

1.0.5.2. Architecture of the NBDs

The NBDs are highly conserved among ABC transporters of different phylogenetic origins (Figure 1.0.5.2) (Locher, 2016). Each NBD contains a catalytic core domain similar to a RecA-like ATPase and a structurally diverse α -helical domain that is unique to ABC transporters. The NBDs are characterised by seven conserved sequence motifs that are necessary for ATP binding and hydrolysis, as well as networking between different domains of the ABC transporter. The Walker A motif (P loop) (GXXGXGK(S/T), of which X is any amino acid) contains a highly conserved lysine residue which binds to the α - and β -phosphates of ATP (Beis, 2015). The Walker B motif ($\Phi\Phi\Phi\Phi$ DE, where Φ is any hydrophobic residue) contains an aspartate that is immediately followed by a glutamate residue. The aspartate provides coordination of the magnesium ion, whilst the catalytic glutamate residue acts as a general base to activate a water molecule for nucleophilic attack at the γ -phosphate of ATP. The ABC signature motif (LSGGQ) (C loop) binds to the γ -phosphate of ATP through electrostatic interactions and orients ATP in a favourable position for ATP hydrolysis. The A loop provides an aromatic side chain that stabilises ATP by packing against the adenine ring. The H motif (switch histidine) forms a flexible hinge region that stabilises the geometry of the transition state. Contacts between the NBDs and TMDs are mediated by the Q loop, through which the D loop transmits signals of ATP hydrolysis to the TMDs for conformational changes and substrate transport. Both the Walker A and B motifs reside in the catalytic core domain, while the rest are located in the α -helical domain.

The two NBDs are arranged “head-to-tail” to form the composite active site upon dimerization for ATP binding and hydrolysis via an interaction between the adjacent D-loops. Two ATP-binding sites are formed between the Walker A motif of one NBD and the signature motif of the other, so that two ATP molecules can bind and hydrolyse at the same time (Hollenstein *et*

al., 2007). In the absence of ATP, the NBDs are separated, in an open conformation; whereas in the ATP-bound state, two molecules of ATP are sandwiched at the interface of the two NBDs. The NBDs are dimerized upon ATP binding and the dimerization is interrupted by ATP hydrolysis, leading to the release of adenosine diphosphate (ADP) and inorganic phosphate (Pi) (Beis, 2015). There is no fundamental requirement for ATP to be hydrolysed by both ATP-binding sites. Some heterodimeric ABC exporters contain one consensus and one degenerate ATP-binding site in the NBDs. Binding of a single ATP molecule to the consensus site is sufficient for ATP hydrolysis, whereas the degenerate binding site is absent of certain motifs and therefore unable to mediate ATP turnover (Ernst *et al.*, 2008).

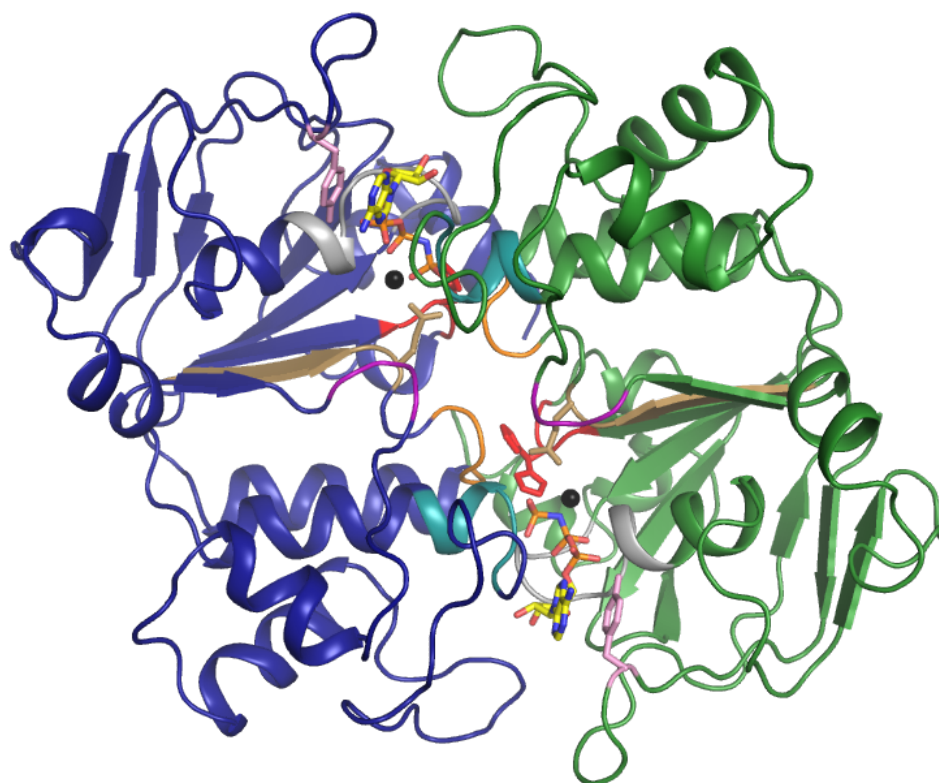


Figure 1.0.5.2. Conserved motifs in the NBDs. The NBDs of the ABC exporter Sav1866 of *S. aureus* in complex with AMP-PNP (PDB ID: 2ONJ, (Dawson & Locher, 2007)). The two NBDs are coloured in blue and green, respectively. Two molecules of AMP-PNP are depicted in stick format and coloured as: carbon in yellow, oxygen in red, nitrogen in blue and phosphorus in orange. The two Mg^{2+} ions are shown as black spheres. Conserved motifs in the NBDs are coloured as: Walker A motif (grey), Walker B motif (sand), ABC-signature motif (cyan), A-loop (pink), H-loop (red), Q-loop (purple) and D-loop (orange).

1.0.5.3. Architecture of the TMDs

The TMDs of ABC transporters are structurally heterogeneous compared to the conserved NBDs to ensure substrate selectivity and account for the chemical and structural diversity of substrates (Locher, 2016). ABC transporters can be divided into seven distinct classes based on sequence and structural homology in the TMDs (Thomas & Tampé, 2018). Types I-III are importers, types IV and V are exporters, type VI are extractors, and members of the type VII subfamily make up components of tripartite efflux pumps.

Type I-II ABC importers contain an arrangement of 5-10 TM helices in each TMD subunit (Oldham *et al.*, 2007; Korkhov *et al.*, 2012). Type III importers, also referred to as the energy-coupling factor (ECF) transporters, display a distinct architecture and lack conventional TMDs. Type IV exporters demonstrate a typical arrangement of six TM helices in each TMD (Dawson & Locher, 2006; Perez *et al.*, 2015; Mi *et al.*, 2017). Type V-VII exporters differ in the numbers of TM helices in each TMD subunit and the structural folds.

1.0.5.4. The alternating access mechanism

For both ABC importers and exporters, an alternating access mechanism has been proposed for substrate translocation (Locher *et al.*, 2002). The key feature of this model is that ABC transporters alternate between the inward- and outward-facing states, exposing the substrate-binding site to either the cytoplasmic or periplasmic side of cell membrane (Beis, 2015). The net direction of transport is largely determined by the relative binding affinities between the substrate and the two conformations (Locher, 2016). Specifically, for ABC importers, the outward-facing conformation has a higher substrate-binding affinity than the inward-facing conformation, while the opposite situation is expected for ABC exporters. Alternation between the two conformations is energised by ATP binding and turnover. The turnover rates vary greatly among different ABC transporters. For example, the ABC exporter McjD displays a basal activity of 44.4 ± 0.5 nmoles/min/mg in proteoliposomes and is stimulated by its substrate MccJ25 in a concentration-dependent manner (Choudhury *et al.*, 2014). In proteoliposomes, maltose-loaded maltose-binding protein MBP stimulates ATP hydrolysis of the maltose importer MalFGK₂ to ~1200 nmoles/min/mg (> 1000-fold stimulation) (Khare *et al.*, 2009). The vitamin B12 importer BtuCD exhibits a basal ATP turnover rate of 180 nmoles/min/mg in proteoliposomes (Borths *et al.*, 2005). The ATP hydrolysis rate of BtuCD is stimulated by the vitamin B12-binding protein BtuF (20 μ M) to 440 nmoles/min/mg.

MsbA is an essential ABC transporter in Gram-negative bacteria and is required for efflux of LPS precursor from the inner- to outer-leaflet of the inner membrane (Doerrler *et al.*, 2004). LPS is a glycolipid consisting of lipid A, core oligosaccharide and a long-chain O-antigenic polysaccharide. LPS is a major component of the outer-leaflet of the outer membrane in Gram-negative bacteria and is essential for the formation of a proper cell envelope that confers bacterial survival under unfavourable conditions (Ruiz *et al.*, 2009). LPS precursor containing

lipid A and core oligosaccharide is synthesised at the inner-leaflet of the inner membrane, flipped by MsbA, further modified, and transported to the outer membrane. The LPS biosynthesis and transport pathway is an attractive therapeutic target for drug development against Gram-negative bacterial pathogens. MsbA has been shown to share high sequence identity and overlapping substrate spectrum with the ABC multidrug transporters LmrA from *L. lactis* and P-glycoprotein from *H. sapiens* (Veen *et al.*, 1998; Dean & Annilo, 2005). Structures of MsbA have been determined in the apo state (Ward *et al.*, 2007), as well as in the presence of the bound substrate (Mi *et al.*, 2017), adenosine 5'-(β,γ -imido) triphosphate (AMP-PNP) (Ward *et al.*, 2007) and inhibitor (Ho *et al.*, 2018).

The apo structure of MsbA adopts an inward-facing conformation, with each TMD formed by the intertwining of TM1-TM3 and TM6 from one monomer and TM4-TM5 from the opposite monomer (Figure 1.0.5.3A) (Ward *et al.*, 2007). The V-shaped opening of the TMDs is formed between TM4/TM5 and TM3/TM6 and faces towards the cytoplasm for substrate entry. The apo NBDs are disengaged from each other. The cryo-EM structure of LPS-MsbA shows that the LPS molecule inserts deeply into the TMDs (Figure 1.0.5.3B) (Mi *et al.*, 2017). The acyl chains of LPS have reached the periplasmic leaflet of the inner membrane. LPS is yet to be flipped but has been almost transported across the lipid bilayer. The overall structure of LPS-bound MsbA resembles the apo MsbA structure but important differences exist. The TMDs are still open to the cytoplasm, but are positioned closer together, which creates a tighter inner cavity for the bound LPS. The NBDs in the LPS bound structure are arranged in an almost perfectly antiparallel head-to-tail manner. The Walker A and the ABC signature motifs are close to and well aligned to each other across the dimeric interface, which facilitates the binding of ATP. The nucleotide-bound MsbA-AMP-PNP structure exhibits an outward-facing conformation (Figure 1.0.5.3C). Upon nucleotide binding, the NBDs dimerise and form a nucleotide sandwich, where the two nucleotide molecules reside in between the Walker A and

ABC signature motifs. Dimerization of the NBDs is transmitted to the TM helices, pulling TM3/TM6 away from TM1/TM2, and results in a conformational change from the inward- to outward-facing state. This is followed by ATP hydrolysis which releases ADP and Pi and dissociates the NBD dimer. ATP hydrolysis resets the transporter to the inward-facing state (Figure 1.0.5.3A). The transport cycle illustrated by MsbA has been verified by EPR spectroscopy studies (Dong *et al.*, 2005; Zou & Mchaourab, 2009).

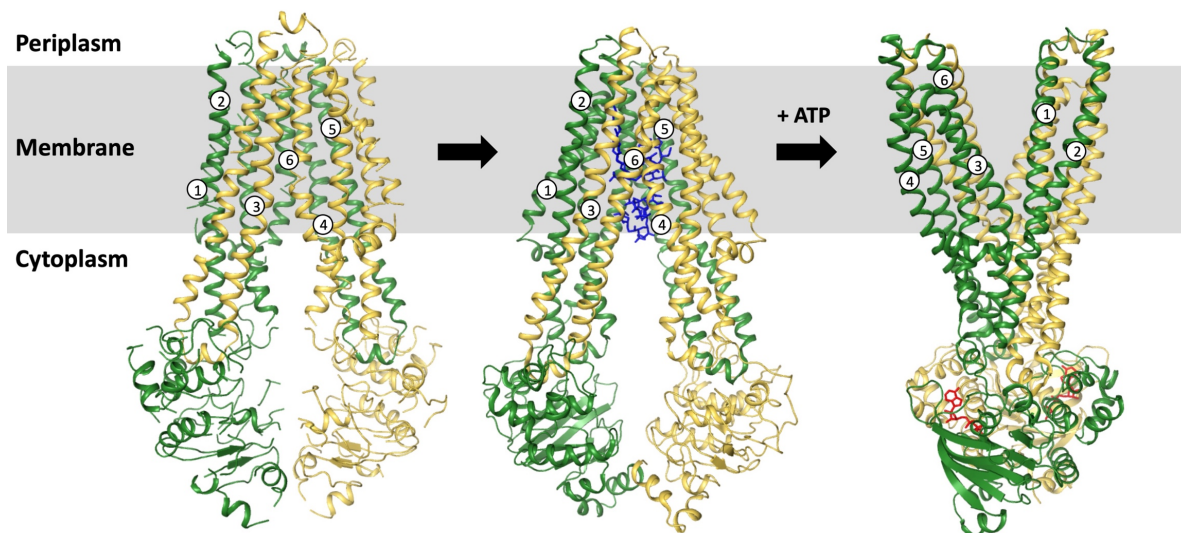


Figure 1.0.5.3. Model for MsbA-mediated LPS flipping. The two monomers of MsbA are coloured in green and yellow, respectively. The bound LPS and AMP-PNP are depicted in blue and red stick format. The circled numbers indicate the different TM helices from one monomer (green). The left panel shows the crystal structure of MsbA in the apo inward-facing state (PDB ID: 3B5X, (Ward *et al.*, 2007)). The TMDs are open to the cytoplasm due to intertwining between TM1-TM3 and TM6 from one monomer and TM4-TM5 from the other monomer. The middle panel is the cryo-EM structure of nucleotide-free MsbA in complex with LPS (PDB ID: 5TV4, (Mi *et al.*, 2017)). The TMDs are still open to the cytoplasm and the NBDs are arranged in an almost perfectly antiparallel head-to-tail fashion. LPS has been almost transported across the bilayer but is yet to be flipped. The right panel shows the crystal structure of MsbA-AMP-PNP in the outward-facing conformation (PDB ID: 3B60, (Ward *et al.*, 2007)). The NBDs are dimerised and the TMDs transit to an outward-facing state owing to the separation between TM3/TM6 and TM1/TM2.

The mechanistic model of type IV ABC exporters has been proposed based on the structural and biochemical studies of MsbA, Sav1866 and McjD (Figure 1.0.5.4) (Dawson & Locher, 2006; Ward *et al.*, 2007; Choudhury *et al.*, 2014; Bountra *et al.*, 2017; Mi *et al.*, 2017). In the absence of ATP, the ABC transporter adopts an inward-facing conformation and the substrate-binding pocket opens to the cytoplasm, which allows substrate binding to the TMDs (state 1). Substrates with high hydrophobicity can be recruited directly from within the inner-membrane bilayer (Locher, 2016). Once the substrate is bound to the TMDs, two ATP molecules bind to the NBDs, causing the NBDs to dimerise (state 1 to state 2). This is followed by the formation of the occluded transporter (state 2) where the TMDs are sealed and the substrate-binding site is inaccessible from either side of the membrane. The occluded conformation (state 2) has been observed in certain ABC exporters, such as McjD and PCAT (Choudhury *et al.*, 2014; Lin *et al.*, 2015). It remains unclear whether the occluded state is necessary for all ABC exporters. The transporter then converts to an outward-facing conformation (state 2 to state 3). In this state, the TMDs form two wings towards the periplasm, which allows the release of substrate. In the last step (state 3 to state 1), the hydrolysis of ATP and release of ADP and Pi dissociate the NBDs and reset the transporter back to the inward-facing state.

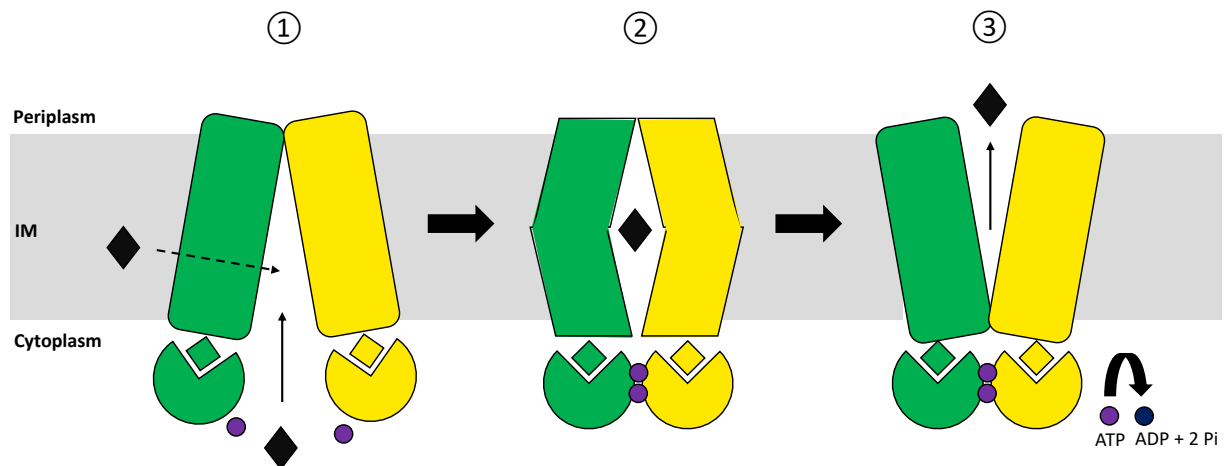


Figure 1.0.5.4. Alternating access mechanism for ABC exporters. The model is proposed based on structures determined for type IV ABC exporters MsbA, Sav1866 and McjD (Dawson & Locher, 2006; Ward *et al.*, 2007; Choudhury *et al.*, 2014; Bountra *et al.*, 2017; Mi *et al.*, 2017). The two half-transporters are coloured in green and yellow, respectively. ATP and ADP molecules are indicated by purple and dark blue circles, respectively. The substrate molecule is shown as the black diamond. The circled numbers indicate the different states during the transport cycle. In the absence of ATP, the transporter adopts an inward-facing conformation and the substrate-binding pocket is open to the cytoplasm, which allows substrate binding to the TMDs (state 1). The horizontal dashed arrow indicates the direct recruit of hydrophobic substrates from within the inner membrane. ATP binding causes the NBDs to dimerise and the ABC transporter converts to the occluded state where the TMDs are not accessible from either side of the inner membrane. ATP binding is followed by the release of substrate upon conversion to the outward-facing conformation (state 3). ATP hydrolysis resets the transporter to an inward-facing state (state 1).

1.0.5.5. ABC importers

Similar to ABC exporters, ABC importers have a characteristic architecture consisting of two conserved NBDs and two less conserved TMDs (Rees *et al.*, 2009). The NBD and TMD subunits of ABC importers are transcribed from individual genes. The NBDs typically form homodimers and the TMDs can assemble to form homo- or hetero-dimers. The NBDs are characterised by a series of conserved motifs, including the Walker A, the Walker B and the ABC signature motifs (Locher, 2016). The TMDs of ABC importers contain a cytoplasmic EAA motif (L-loop), which is composed of two α -helices and is responsible for signal transition from the NBDs to the TMDs (Saurin *et al.*, 1994). Most ABC importers contain an SBP that captures substrates (at nM to mM affinity) from the periplasm and directs them to transporters in the cell membrane (Rice *et al.*, 2014).

Structures of ABC importers have been determined in different conformational states. Based on their topology and transport mechanism, ABC importers are classified into three categories (Figure 1.0.5.5). Type I ABC importers include the maltose transporter MalFGK₂ (Khare *et al.*, 2009), the D-methionine transporter MetNI (Kadaba *et al.*, 2008) and the two molybdate/tungstate transporters ModBC (Gerber *et al.*, 2008; Hollenstein *et al.*, 2007). Generally, type I importers contain fewer TM helices (Rice *et al.*, 2014). TMDs of the maltose importer MalFGK₂ form a heterodimer of MalF and MalG, which contain eight and six TM helices, respectively (Figure 1.0.5.5A). MetNI and ModBC contain five and six TM helices in each TMD, forming homodimers totalling 10 and 12 helices, respectively. In the absence of SBPs, type I importers show negligible ATPase activity, which is stimulated by their cognate ligand-loaded SBPs (Davidson *et al.*, 1992). Type I importers undergo predominant rigid body conformational changes during the transport cycle. Several structurally characterised ABC importers fall into the type II category, which includes the vitamin B₁₂ transporter BtuCD

(Locher *et al.*, 2002), the metal-chelate transporter HI1470/1 (Pinkett *et al.*, 2007) and the heme transporter HmuUV (Woo *et al.*, 2012). Type II importers are generally larger and have more TM helices in the TMDs. Typically, the structurally characterised BtuCD transporter contains 10 TM helices in the TMD subunit (Figure 1.0.5.5B). The conformational changes of type II importers are confined to core helices during substrate transport. The more recently identified ECF transporters have defined characteristic features of the type III ABC importer category, including the folate and the hydroxymethyl pyrimidine ECF transporters (Xu *et al.*, 2013; Wang *et al.*, 2013). Unlike type I and II importers, ECF transporters do not require SBPs for substrate transport (Rodionov *et al.*, 2006; Rodionov *et al.*, 2009). Instead of two related TMDs, ECF transporters contain an “L-shaped” transmembrane coupling subunit (EcfT), an integral substrate-binding subunit (EcfS) and two homologous ATPase subunits (EcfA and EcfA') (Figure 1.0.5.5C) (Xu *et al.*, 2013). The membrane-embedded EcfS protein (S component) captures substrates directly from the periplasm with a pM to nM range affinity (Zhang, 2013).

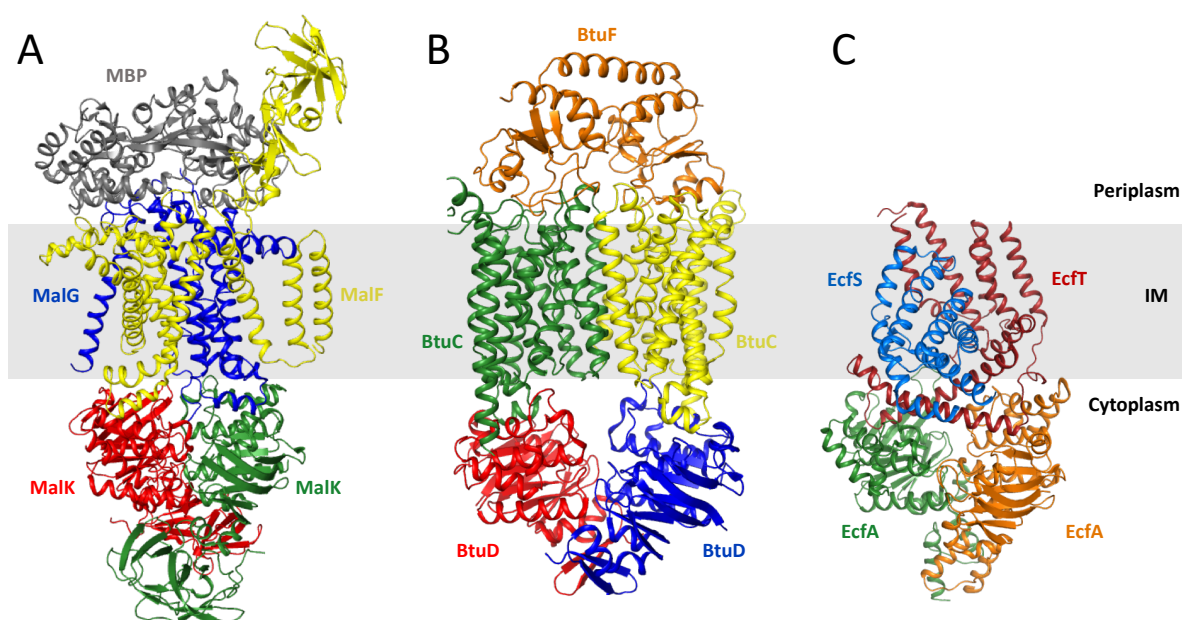


Figure 1.0.5.5. Crystal structures of representative ABC importers. (A) The type I ABC importer MalFGK₂-MBP (PDB ID: 2R6G, (Oldham *et al.*, 2007)). (B) The type II ABC importer BtuCD-F (PDB ID: 2QI9, (Hvorup *et al.*, 2007)). (C) The type III ABC importer folate ECF transporter (PDB ID: 4HUQ, (Xu *et al.*, 2013)).

1.0.5.6. SBPs

Both type I and type II ABC importers contain a fifth subunit, the SBPs for initial substrate recognition and binding (Scheepers *et al.*, 2016). SBPs capture substrates from the periplasm and direct them to the TMDs of ABC transporters, by which the ligands are translocated into the cytoplasm upon ATP hydrolysis at the NBDs (Locher, 2016). In Gram-negative bacteria, SBPs are located in the periplasm. By contrast, in Gram-positive bacteria and archaea, SBPs are fused to the inner membrane via lipid anchors or are directly tethered to their cognate TMDs. In the latter case, the SBPs are referred to as covalently-linked SBDs (Lewinson & Livnat-Levanon, 2017). Although usually one SBD associates with one ABC transporter, there are cases where up to three SBDs in tandem fuse to each TMD subunit, resulting two, four, or six SBDs per assembled ABC transporter, such as LuxPQ from *Vibrio harveyi* for autoinducer-2 signal transduction and the GlnPQ from *Lactococcus lactis* for amino acid uptake (Neiditch *et al.*, 2006; Gouridis *et al.*, 2014). Assembly of multiple SBDs greatly enhances local substrate concentration at the TMDs.

Although SBPs were originally found as a part of prokaryotic ABC importers, they have been shown to associate with other membrane transporters (Berntsson *et al.*, 2010). Prokaryotic secondary transporters such as the tripartite tricarboxylate transporters and the tripartite ATP-independent periplasmic transporters use SBPs to capture their substrates (Winnen *et al.*, 2003; Fischer *et al.*, 2010). Some prokaryotic two-component regulatory complexes rely on SBDs for signal recognition (Sun *et al.*, 2002). Certain eukaryotic ligand-gated ion channels and G-protein coupled receptors (GPCRs) are activated by ligand binding of their cognate SBDs (O'Hara *et al.*, 1993). Finally, SBPs are shown to be part of prokaryotic DNA-binding proteins involved in translational regulation (Schumacher *et al.*, 1994).

Despite low sequence similarity (often < 20%), varied sizes (25-70 kDa) and functions, SBPs share a common core architecture consisting of two α/β domains that are connected by a hinge region. The substrate-binding site is located at the interface of two domains. Both domains consist of a central β -sheet flanked by a number of α -helices. In the apo state, the SBP exhibits an open conformation and the two domains freely move around the hinge. Ligand binding induces conformational changes through a hinge-bending movement and the SBP converts to a closed conformation, where the two domains pack tightly around the ligand. SBPs associated with type I ABC importers possess a flexible hinge region and therefore undergo large conformational changes upon ligand binding (Lewinson & Livnat-Levanon, 2017). The ligand-bound SBPs are structurally distinct from their apo forms. This allows the inner-membrane components to distinguish between the substrate-bound and apo SBPs, thereby enabling efficient substrate transport. By contrast, the rigid α -helical hinge of SBPs associated with type II ABC importers limits the relative domain movement, resulting similar apo and ligand-bound conformations. The ligand-induced closure of SBPs of type I ABC importers is typically termed as the “Venus fly-trap” mechanism (Figure 1.0.5.6).

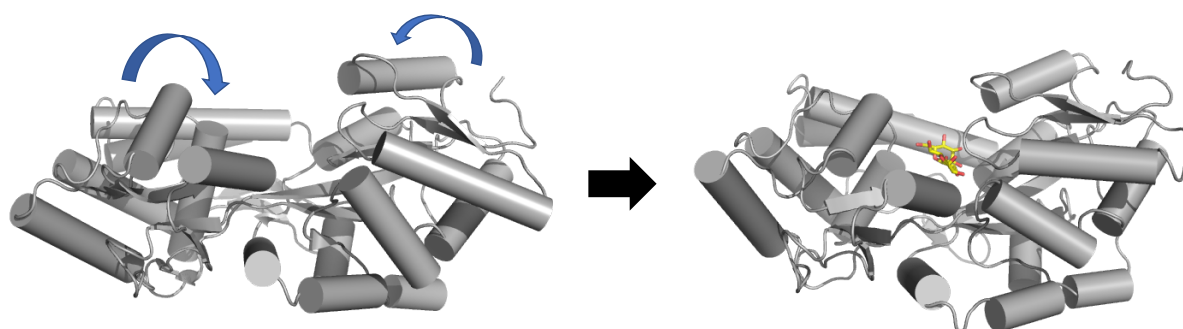


Figure 1.0.5.6. The “Venus fly-trap” mode of ligand binding. The left and right panels show the apo (PDB ID: 1LLS, (Rubin *et al.*, 2002)) and maltose-bound (PDB ID: 1ANF, (Quiocho *et al.*, 1997)) structures of the maltose-binding protein MBP, respectively. Maltose is shown in stick format with carbon and oxygen coloured in yellow and red, respectively. In the apo state, MBP exhibits an open conformation and the cavity is accessible. Upon maltose binding, MBP converts to a closed conformation and the cavity becomes occluded.

The flexible and modular structure of SBPs allows them to associate with a wide range of protein complexes. Based on topology of the central β -sheets and secondary structures, two SBP classes were identified (Fukami-Kobayashi *et al.*, 1999). Class I (*e.g.* the lysine-binding protein LBP (Sack *et al.*, 1989a) and the leucine/isoleucine/valine-binding protein LIVBP (Sack *et al.*, 1989b)) and class II SBPs (*e.g.* the oligopeptide-binding protein OppA (Tame *et al.*, 1994) and the histidine-binding protein HisJ (Oh *et al.*, 1994)) show a β -sheet topology of $\beta_2\beta_1\beta_3\beta_4\beta_5$ and $\beta_2\beta_1\beta_3\beta_n\beta_4$, respectively, where n represents the β -strand after the first crossover from domain I to domain II, or vice versa. The hinge region is formed by three and two segments in class I and class II SBPs, respectively. A third class of SBPs was later defined which contains a single α -helical hinge region and shows a smaller domain movement upon substrate binding (*e.g.* the Zinc-binding protein TroA, (Lee *et al.*, 2002) and the vitamin B₁₂-binding protein BtuF (Borths *et al.*, 2002)). Based on available three-dimensional (3D) structures, SBPs have been later classified into six distinct clusters (A-F) (Berntsson *et al.*, 2010). Proteins within one cluster do not necessarily share the same substrate specificity.

Cluster A

Proteins of cluster A are all class III SBPs, which are associated with type II ABC importers. A distinct feature of cluster A (class III) SBPs is a single α -helix that functions as a hinge region (Figure 1.0.5.7A). The rigidity of this helix results in small motion of the two α/β domains upon substrate binding. All the SBPs in cluster A are associated with metal binding, either by direct interactions or as metal chelates.

Cluster B

Cluster B consists of class I SBPs that contain three segments connecting the two α/β domains (Figure 1.0.5.7B). Most of the cluster B SBPs bind carbohydrates, branched chain amino acids, natriuretic peptides and autoinducer-2. These SBPs are typically part of type I ABC importers, two-component regulatory systems and natriuretic peptide receptors or function as transcriptional regulators and signal transducers.

Cluster C

Cluster C consists exclusively of class II SBPs that interact with ABC transporters. Proteins of cluster C SBPs bind a wide range of substrates, such as di- and oligopeptides, arginine, cellobiose and nickel ions. Specifically, cluster C SBPs are larger in size (53-70 kDa) and possess an additional domain (Figure 1.0.5.7C). For SBPs associated with oligopeptide transport, such as *Bacillus subtilis* AppA and *L. lactis* OppA, this extra domain expands the substrate-binding pocket to accommodate the large oligopeptides.

Cluster D

SBPs of cluster D exclusively belong to class II and contain a hinge region consisting of two short strands (4-5 amino acids in length) (Figure 1.0.5.7D). Proteins of this cluster bind a range of different ligands, including carbohydrates, thiamine, putrescine, ferrous or ferric iron, and tetrahedral oxyanions.

Cluster E

In cluster E, all SBPs belong to class II and are associated to the tripartite ATP-independent periplasmic transporter family. These transporters use an electrochemical gradient (H^+ or Na^+) to energise substrate translocation. SBPs of cluster E feature a large β -strand and a long α -helix spanning both domains (Figure 1.0.5.7E). Substrates of the cluster E SBPs include ectoine, lactate, sialic acid, 2-keto acid and pyroglutamic acid.

Cluster F

Like cluster D, cluster F consists of class II SBPs with a hinge region formed by two segments. However, hinge regions of the cluster F SBPs are typically 8-10 amino acid residues in length (Figure 1.0.5.7F), rather than 4-5 residues as observed in cluster D (Figure 1.0.5.7D). The elongated hinge region enables higher flexibility of cluster F SBPs between the open and closed conformations. Proteins belonging to cluster F bind a wide range of ligands including trigonal planar anions, compatible solutes and amino acids.

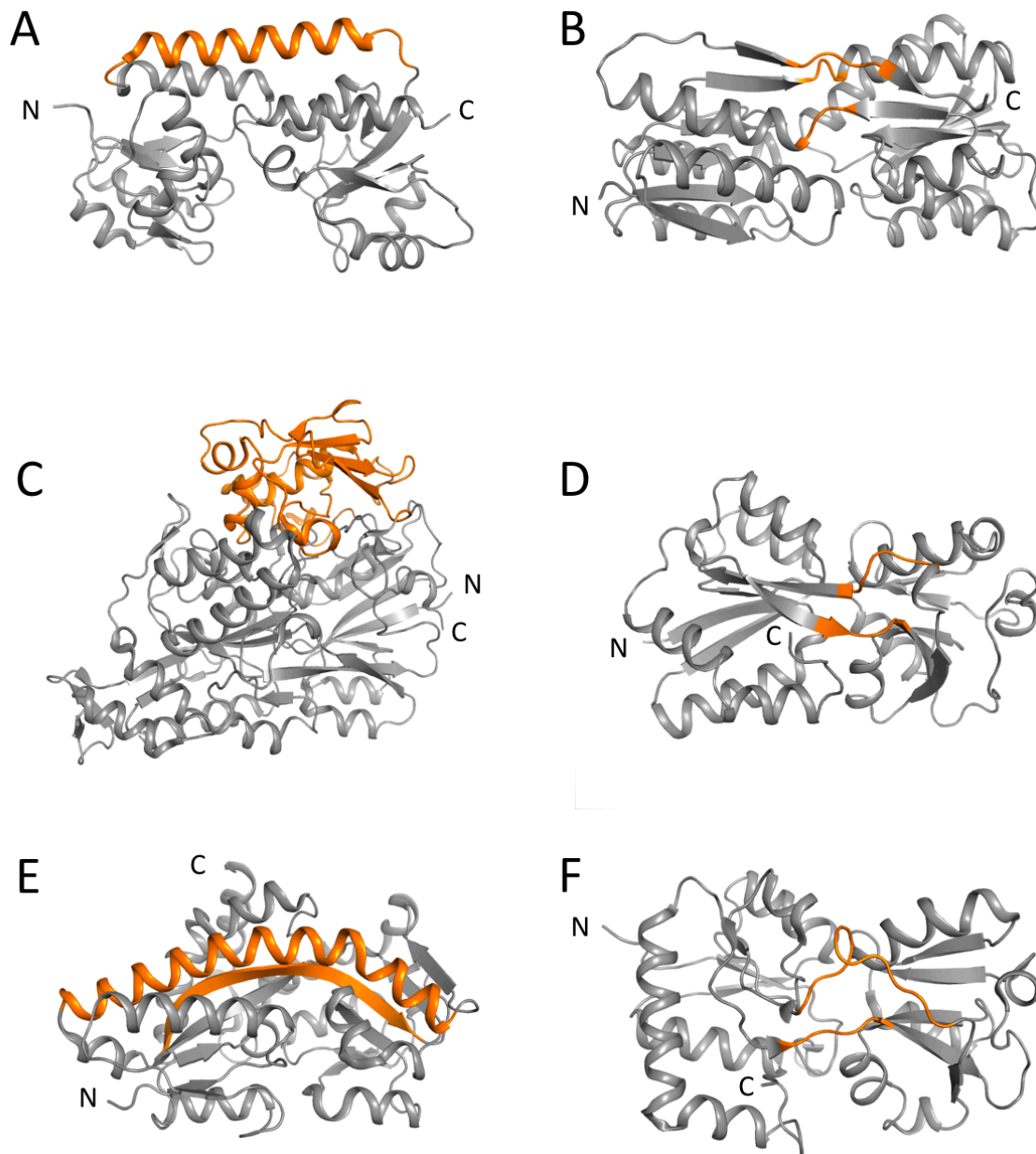


Figure 1.0.5.7. Different clusters of SBPs with their distinct features in orange. (A) In cluster A SBPs, the two domains are connected by a rigid α -helix (*E. coli* BtuF, PDB ID: 1N2Z, (Borths *et al.*, 2002)). (B) Cluster B SBPs contain three segments that connect the two domains (*E. coli* RBP, PDB ID: 1DRJ, (Bjorkmans *et al.*, 1994)). (C) Cluster C SBPs have an additional domain and thus are larger than other SBPs (*L. lactis* OppA, PDB ID: 3DRF, (Berntsson *et al.*, 2009)). (D) Cluster D SBPs have two hinge segments that are 4-5 amino acids in length (*E. coli* ModA, PDB ID: 1AMF, (Hu *et al.*, 1997)). (E) Cluster E SBPs are associated with the tripartite ATP-independent periplasmic transporters. They contain hinge regions formed of a long β -strand and a large α -helix (*Ruegeria pomeroyi* UehA, PDB ID: 3FXB, (Lecher *et al.*, 2009)). (F) Cluster F SBPs contain two hinge segments that are 8-10 amino acids in length, enabling a high flexibility (*E. coli* HisJ, PDB ID: 1HSL, (Yao *et al.*, 1994)). Adapted from: (Berntsson *et al.*, 2010).

1.0.5.7. Type I ABC importers

Type I ABC importers are involved in medium-affinity uptake of a diverse range of nutrients including sugars, ions, amino acids and peptides (Locher, 2016). Substrate binding affinities of type I ABC importers exhibit considerable degree of variation, ranging between tens of nM to several mM. The maltose transporter (MalFGK₂-MBP) from *E. coli* has been extensively studied and reveals a detailed transport mechanism of Type I ABC importers (Figure 1.0.5.8).

The MalFGK₂-MBP system consists of two TMD subunits (MalF and MalG), two NBD subunits (MalK) and an additional SBP (MBP) (Oldham *et al.*, 2007). In the absence of MBP and nucleotide, MalFGK₂ adopts a resting-state inward-facing conformation (Khare *et al.*, 2009). The maltose-binding site is accessible to the cytosol and the two MalK subunits are separated. Type I ABC importers show a very low basal ATPase rate (Davidson *et al.*, 1992). Binding of the cognate SBPs stimulates the ATPase activity of the maltose and alginate importers (Khare *et al.*, 2009; Maruyama *et al.*, 2015), but not that of the molybdate/tungstate importers (Gerber *et al.*, 2008; Hollenstein *et al.*, 2007). In the pre-translocation state, the maltose-loaded MBP forms initial contacts with the TMDs of the nucleotide-free MalFGK₂ (Oldham & Chen, 2011). The two MalK subunits rotate inward and form a partially closed dimer. Consequently, the MalG and MalF subunits rotate against centre of the transporter, which narrows the opening to the cytoplasm while the periplasmic gate remains closed. In type I ABC importers, docking of SBPs to the ABC transporters is transient and exhibits low affinity (Lewinson & Livnat-Levanon, 2017). The large periplasmic P₂ loop of MalF stabilises binding of MBP to the TMDs. Each α/β -domain of MBP interacts with one TMD subunit of MalFGK₂. Binding of ATP to the NBDs in the pre-translocation state further closes the NBDs, which leads to opening of MBP and conformational changes in the TMDs. In the resulting outward-facing state, maltose is released from MBP into the TMDs and two ATP molecules are

positioned at the dimeric interface of the two MalK subunits for hydrolysis (Oldham *et al.*, 2007). The release of maltose from MBP to the TMDs is achieved by two actions: first, MBP is distorted to a lower-affinity open conformation; second, a periplasmic scoop loop from MalG inserts into the substrate-binding site of MBP and causes steric clash with the bound maltose (Oldham *et al.*, 2007). Upon ATP hydrolysis, the NBDs separate from each other and consequently the transporter converts to an inward-facing state. Maltose is then released to the cytoplasm and MBP dissociates from the TMDs. Dissociation of MBP and nucleotide resets the MalFGK₂ transporter.

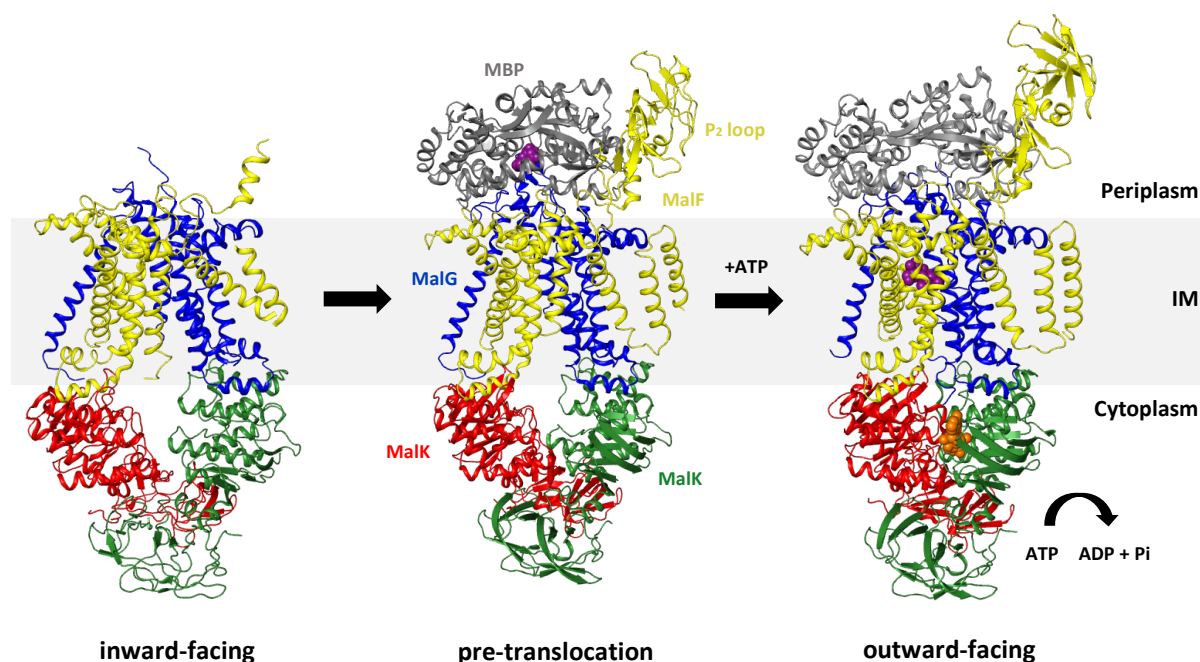


Figure 1.0.5.8. Transport mechanism of the Type I ABC importer MalFGK₂-MBP. The subunits are coloured in yellow (MalF), blue (MalG), red/green (MalK) and grey (MBP). Maltose and nucleotide are shown as purple and orange spheres, respectively. In the absence of MBP and nucleotide, the resting-state MalFGK₂ exists in an inward-facing conformation (PDB ID: 3FH6, (Oldham *et al.*, 2007)). Association of the maltose-bound MBP to the TMDs initiates conformational changes in the TMDs and semi-dimerization of the NBDs, resulting a partially closed conformation (pre-translocation state) (PDB ID: 3PV0, (Oldham & Chen, 2011)). Binding of ATP to the NBDs in the pre-translocation state further closes the NBDs, which leads to the opening of MBP and conformational changes in the TMDs. In the outward-facing state, maltose is released from MBP into the TMDs and two ATP molecules are positioned at the dimeric interface of the NBDs for hydrolysis (PDB ID: 2R6G, (Oldham *et al.*, 2007)). MBP and nucleotide dissociation resets the MalFGK₂ transporter.

1.0.5.8. Type II ABC importer

Type II ABC importers facilitate high-affinity uptake of non-metabolites including vitamin B₁₂, heme and iron-siderophore complexes (Locher, 2016). Compared to substrates of type I ABC importers, these compounds are typically larger, more hydrophobic and are available at low concentrations. The high substrates binding affinity (sub-nM to nM) and specificity of type II ABC importers can be exclusively attributed by their cognate SPBs (Lewinson & Livnat-Levanon, 2017). SBPs of type I systems follow a “Venus fly-trap” mode of ligand binding (Figure 1.0.5.6). Whereas SBPs associated with type II ABC importers are typically connected by a rigid α -helix, resulting in a limited local rearrangement and up to a few degrees of closure (Berntsson *et al.*, 2010). The vitamin B₁₂ transporter BtuCD-F from *E. coli* provides a model of transport mechanism of type II ABC importers (Figure 1.0.5.9).

BtuCD is a homodimer, with each monomer composed of a TMD (BtuC) and an NBD (BtuD) (Locher *et al.*, 2002). The SBP BtuF is required to capture vitamin B₁₂ from the periplasm and direct it to the BtuC subunits. Each of the two TMDs traverse the inner membrane 10 times, totalling 20 TM helices in the BtuCD transporter. The TMDs contain three gates: one on the periplasmic side (periplasmic gate) and two on the cytoplasmic side (cytoplasmic gates 1 and 2) (Locher, 2016). The periplasmic gate and cytoplasmic gate 1 are formed by TM5, whereas the cytoplasmic gate 2 is formed by the loop between TM2 and TM3. The transport cycle of BtuCD-F starts from an outward-facing state (Korkhov *et al.*, 2014). Upon ATP binding, the NBDs dimerize, resulting in opening of the periplasmic gate and closing of the cytoplasmic gate 2. To date, all type II importers display a high degree of basal ATP hydrolysis in the absence of their respective SBPs and a low stimulation by SBPs (Lewinson & Livnat-Levanon, 2017). Docking of vitamin B₁₂-loaded BtuF to BtuC subunits leads to an occluded conformation (Korkhov *et al.*, 2012) and vitamin B₁₂ is released into a hydrophobic (Teflon-

like) cavity in the TMDs, which shows no measurable affinity to its substrate. Type II ABC systems form high-affinity (nM to pM), highly stable complexes with their cognate SBPs. Upon docking to BtuC, the binding pocket of BtuF is distorted and occupied by loops from the periplasmic gate, which promotes the release of vitamin B₁₂ from BtuF to the TMDs. Unlike MalFGK₂ that shows prominent, global conformational rearrangement, TMDs of BtuCD undergoes little structural alternation upon SBP docking (Lewinson & Livnat-Levanon, 2017). The transition to the inward-facing state is triggered by hydrolysis of ATP and release of ADP and Pi, which disrupt the dimeric NBDs, leading to opening of the cytoplasmic gate 2. The large size of vitamin B₁₂ prevents closure of the cytoplasmic gate 1 and the resulting strained pressure promotes substrate dissociation to the cytoplasm. Although no structure is available for BtuCD in an inward-facing state, the crystal structure of its homologous metal-chelate-type ABC transporter Hi1470/71 illustrates an inward-facing conformation (Pinkett *et al.*, 2007). Upon release of substrate, BtuCD converts to an asymmetric conformation that lacks a central cavity since both the periplasmic gate and cytoplasmic gates are closed (Hvorup *et al.*, 2007). The asymmetric state is highly stable which probably prevents ion leakage through the translocation pathway. Before transition to the ATP-bound outward-facing state, an apo symmetrical outward-facing BtuCD transiently occurs upon dissociation of BtuF and rearrangement of cytoplasmic gate 1. ATP binding then resets the BtuCD transporter.

Structural and mechanistic differences between type I and type II ABC importers are summarised in Figure 1.0.5.5 and Table 1.0.5.1.

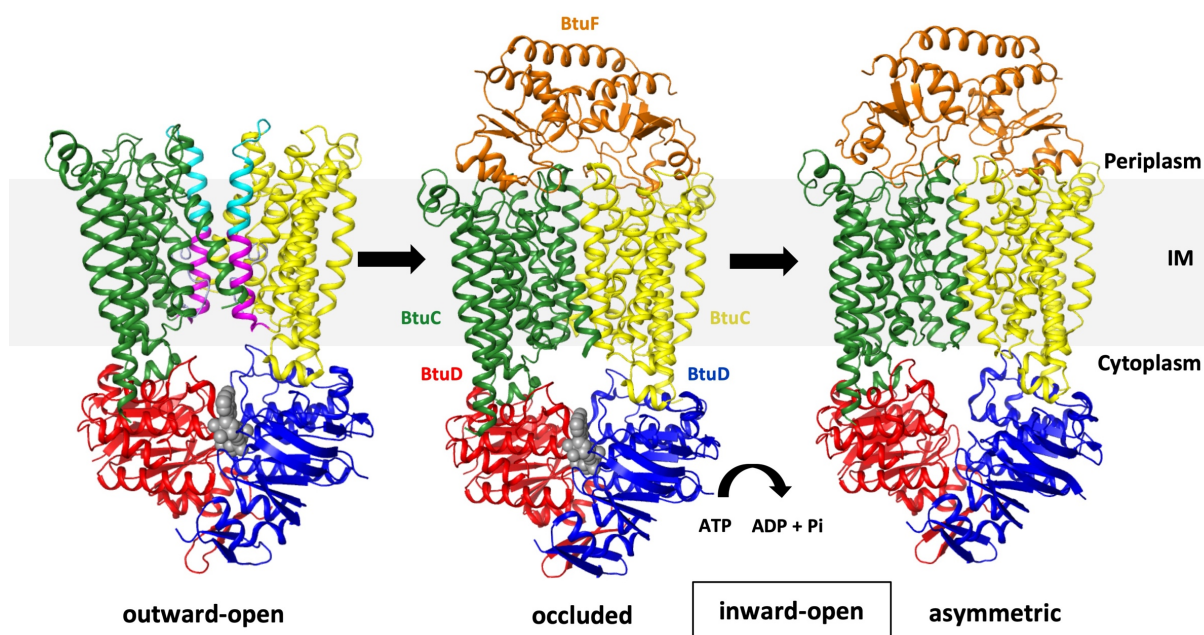


Figure 1.0.5.9. Transport mechanism of the Type II ABC importer BtuCD-F. The subunits of BtuCD-F are coloured in green/yellow (BtuC), red/blue (BtuD) and orange (BtuF). The periplasmic gate, cytoplasmic gates 1 and 2 are coloured in cyan, magentas and light blue, respectively. The ATP analogue AMP-PNP is shown as grey spheres. The resting-state BtuCD exhibits an ATP-bound outward-facing conformation (PDB ID: 4R9U, (Korkhov *et al.*, 2014)). Association of vitamin B₁₂-bound BtuF to the TMDs triggers local conformational changes and BtuCD converts to an occluded state (PDB ID: 4FI3, (Korkhov *et al.*, 2012)). ATP hydrolysis promotes formation of a putative inward-facing state where vitamin B₁₂ is released to the cytoplasm. BtuCD then converts to an asymmetric state (PDB ID: 2QI9, (Hvorup *et al.*, 2007)), followed by dissociation of BtuF and binding of ATP, which reset the transporter to an apo state and then ATP-bound outward-facing state.

	Type I	Type II
Systems	MalFGK ₂ , MetNI, ModBC	BtuCD, ModBC, HmuUV
Transported substrates	Ions, sugars, amino acids, peptides, solutes	Iron-siderophore complexes, vitamin B12, hemin
Substrate affinity	Low to high μM	nM
Structural rearrangements of the SBP upon binding	Large “Venus fly-trap” like closure of the SBP around the substrate	Apo and substrate-bound SBPs are similar
Basal ATPase rate	Low	High
Stimulation of ATPase rate by the SBP	High	Low
Affinity of the transporter-SBP interaction	Variable, tens nM to tens μM	Very high, pM to nM
Effect of ATP on the transporter-SBP interaction	ATP binding stabilises the interaction	ATP binding destabilises the interaction
Effect of substrate on the transporter-SBP interaction	Substrate binding by the SBP promotes its association with the TMDs	High substrate concentrations inhibit the SBP-transporter interaction
Conformation of the ATP-bound transporter	Outward-facing (following SBP docking)	Inward-facing
Conformational coupling between the TMDs and the NBDs	Movements of the TMDs and the NBDs are tightly coupled	In the absence of the SBP, the movements of the TMDs and the NBDs are uncoupled

Table 1.0.5.1. Mechanistic comparisons between type I and II ABC importers. Adapted from (Lewinson & Livnat-Levanon, 2017).

1.0.6. RND transporters

RND transporters have been found in archaea, bacteria and eukaryotes, but only contribute to multidrug resistance in Gram-negative bacteria (Tseng *et al.*, 1999). Compared to other secondary transporters, RND transporters exhibit an atypical structure with a conserved hydrophobic transmembrane domain and a variable hydrophilic periplasmic domain that composes up to 60% of the protein mass (Figure 1.0.6.1).

Gram-negative bacteria tend to have higher resistance to antimicrobial than Gram-positive bacteria. A main reason for this is the presence of RND-type efflux pumps that are overexpressed in the cell membrane (Li *et al.*, 1994). Members of the RND superfamily demonstrate a broad substrate spectrum with typical substrates including antibiotics, dyes, detergents and cytotoxic solvents (Elkins & Nikaido, 2002). RND transporters are also involved in a wide range of physiological functions, such as detoxification of toxic intracellular metabolites, trafficking of quorum sensing molecules, colonisation and host adherence (Yamaguchi *et al.*, 2015). Among all six types of bacterial transporter families, only the RND-type pumps (with a few exceptions) exist predominantly in a tripartite complex composed of an inner-membrane RND transporter, an outer membrane factor (OMF) and a membrane fusion protein (MFP) that connects the RND transporter and the OMF protein (Figure 1.0.6.1). Each of these components is essential for composition of a functional complex. Due to their tripartite composition, RND pumps export substrates directly into the external medium, bypassing the periplasm, after which drugs must traverse the low-permeable outer membrane barrier to re-enter bacterial cells (Nikaido, 1996).

To date, all functionally characterised RND-type pumps function as exporters using a H^+ /substrate antiport mechanism, where the efflux of drugs is coupled to the influx of protons

down the electrochemical gradient (Tseng *et al.*, 1999). From amino acid sequences, all the RND-type transporters are expected to contain 12 TM helices and two large periplasmic domains between TM1-TM2 and TM7-TM8 (Tseng *et al.*, 1999). In the best-characterised RND efflux, the inner-membrane RND transporters all exist as homotrimers pumps (*i.e.* AcrAB-TolC of *E. coli*, CusCBA of *E. coli*, and MexAB-TolC of *P. aeruginosa*). The periplasmic domains contain the ligand-binding pockets, while the transmembrane domains are responsible for energy transduction.

The RND superfamily and the AcrAB-TolC system will be further discussed in chapter 4.

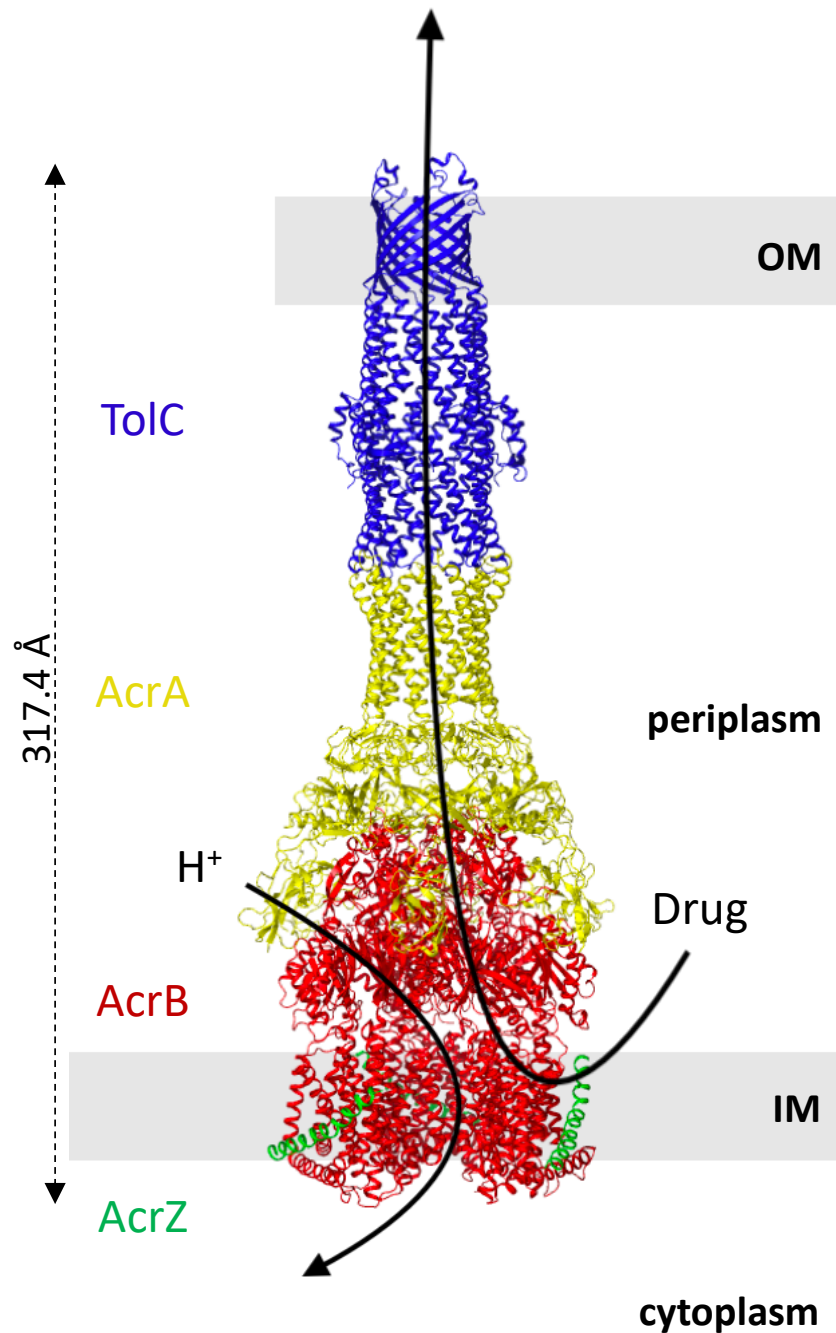


Figure 1.0.6.1. The AcrA₆B₃Z₃-TolC₃ efflux pump. The assembly is composed of the homotrimers AcrBZ-TolC and the hexamer AcrA, and traverses both the inner and outer membranes of *E. coli* (PDB ID: 5O66, (Wang *et al.*, 2017)). The overall mass of the AcrABZ-TolC complex is 771 kDa. The inner-membrane transporter AcrB captures substrates from either within the inner membrane or the periplasm. The drugs are then transported to the outer membrane protein TolC and extruded to the extracellular medium. AcrB is the major component of the system, responsible for substrate specificity and energy transduction. The drug efflux process is mediated by AcrB through a functional rotation mechanism coupled to proton influx. AcrA, AcrB, AcrZ and TolC are coloured in yellow, red, green and blue, respectively.

1.0.7. Aims of the work

Members of the RND, MATE, SMR and PACE families are exclusively exporters, while the ABC and MFS superfamilies contain both exporters and importers (Du *et al.*, 2018). Exporters of the six families have been found to contribute to bacterial resistance, while ABC and MFS importers also confer bacterial nutrient uptake (Piddock, 2006; Du *et al.*, 2018). For example, the ABC exporter Sav1866 and the MFS exporter EmrD are involved in bacterial multidrug resistance, whereas the ABC importer BtuCD-F and the MFS permease LacY are responsible for the uptake of vitamin B₁₂ and lactose, respectively. These six families of transporters together constitute a network conferring bacterial growth and survival. Therefore, they represent good targets for therapeutic intervention.

Although transport models have been proposed for different classes of transporters, the mechanistical details of substrate transport remain elusive. ABC importers play an essential role in nutrient uptake and usually show specific substrate spectrum. However, it is unclear how these transporters specifically recognise and transport certain substrates. We want to investigate the molecular details that underpin substrate recognition and transport of ABC importers and whether they can transport other compounds. The RND superfamily is a major reason for multidrug resistance in Gram-negative bacteria. Although there are structures available, there is a lack of inhibitor-bound structures for RND transporters. Inhibitors of RND transporters can be used in combination therapy to restore efficacy of conventional antibiotics. To date, no inhibitor against bacterial efflux systems has been clinically licensed as a drug for bacterial infections. We are interested in identifying inhibitors of RND transporters and determine structures of the protein in complex with inhibitors.

Chapter 2

Structural and functional studies of

***E. coli* TauA**

Manuscript in preparation:

Qu, F., El Omari, K., Wagner, A., De Simone, S. & Beis, K. Desolvation of the substrate binding protein TauA dictates ligand specificity for the alkanesulfonate ABC importer TauABC.

2.1. TauA

2.1.1. Nutrient starvation

Bacteria in their natural habitat are frequently subjected to nutrient starvation. In certain natural environments, such as open oceans or volcanos, nutrient-limiting conditions may sometimes last for up to several months (Jannasch & Taylor, 1984). Many bacteria have developed a series of survival mechanisms during nutrient starvation. In Gram-negative bacteria, nutrient starvation can limit cell division and alter cell morphology (*e.g.* change in surface/volume ratio) (Kjelleberg *et al.*, 1993; Kolter *et al.*, 1993). In starved *E. coli*, synthesis of bulk proteins and proteins for exponential growth is significantly reduced, whilst degradation of existing proteins is increased, providing amino acids to synthesise proteins that are more essential for cell survival (Reeve *et al.*, 1984). Cryptic growth also plays an importance role in starvation survival, where the surviving cells recycle nutrients released from dead cells (Postgate, 1976). Finally, a vital strategy employed by starved bacteria is the expression of previously unused proteins (*e.g.* transporters) to sequester and assimilate nutrients from a wide range of compounds (Roszak & Colwell, 1987).

Studies on starving bacteria are of great interest for bacterial ecophysiology and its application (Roszak & Colwell, 1987). Nutrient-starving bacteria play an essential role in secondary production and pioneering new environment. Furthermore, nutrient limitation has been shown to be an efficient approach for cell growth control in immobilised cell bioreactors (Inloes *et al.*, 1985). Stabilisation of starving bacterial cells may be a crucial step to new fermentation technology.

2.1.2. TauA and SsuA

Sulfur is an essential element for bacterial survival and growth. Living organisms utilise sulfur to synthesise amino acids (*i.e.* methionine and cysteine) and cofactors (*e.g.* molybdopterin and coenzyme A) (Kertesz, 2001). Under normal conditions, bacteria tend to assimilate sulfur from either the amino acid cysteine or inorganic sulfate compounds (Ploeg *et al.*, 1996). Under sulfur starvation, bacteria rely on organic aliphatic sulfonates as sulfur sources for growth. Aliphatic sulfonates that can be utilised by bacteria include sulfonates, sulfamates and sulfate esters. Bacteria have developed several systems to utilise alkanesulfonates for sulfur acquisition.

In *E. coli*, two sets of genes have been identified that encode sulfur assimilation systems, *tauABCD* and *SsuEADCB* (Figure 2.1.2.1) (Ploeg *et al.*, 1996; Ploeg *et al.*, 1999). The operons of *tauABCD* and *SsuEADCB* are upregulated only during cell growth in the absence of cysteine or sulfate. The *tauABCD* cluster is located at 8.5 min on the *E. coli* chromosomal DNA and the TauABCD system shows high specificity to 2-aminoethanesulfonic acid (taurine) (Ploeg *et al.*, 1996; Eichhorn *et al.*, 1999). The *ssuEADCB* gene cluster (21.4 min on the chromosome) encodes a sulfur utilisation system that allows *E. coli* to assimilate sulfur from a broad range of alkanesulfonates except for taurine, such as 2-(N-morpholino) ethanesulfonic acid (MES), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and 3-(N-morpholino) propanesulfonic acid (MOPS).

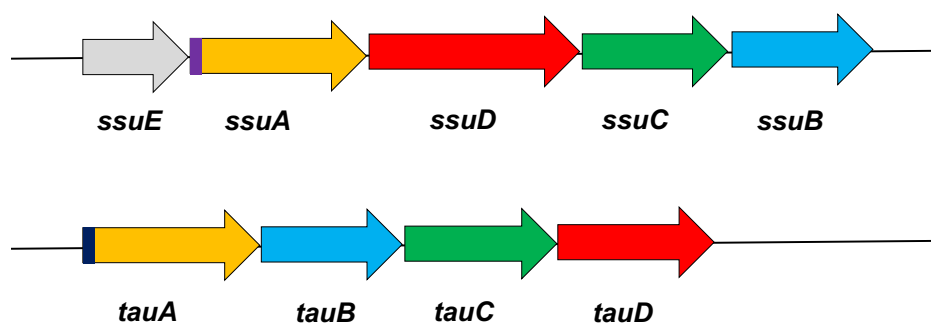


Figure 2.1.2.1. The *ssuEADCB* and *tauABCD* operons from *E. coli*. Arrows indicate the direction of gene transcription. Accession codes: P75853 (SsuA), P0AAI1 (SsuB), P75851 (SsuC), P80645 (SsuD), P80644 (SsuE), Q47537 (TauA), Q47538 (TauB), Q47539 (TauC), P37610 (TauD). The purple and dark blue boxes indicate sequences encoding the signal peptides of SsuA and TauA, respectively.

Taurine is one of the few naturally occurring alkanesulfonates, accounting for ~0.1% of total human body weight (Huxtable, 1992). Taurine forms the main component of bile and is widely distributed in animal tissues such as the large intestine, which constitutes the natural habitat of *E. coli* (Ploeg *et al.*, 1996). Taurine plays fundamental roles in animals, which include synthesis of bile acids and antioxidation. Animals are unable to metabolise taurine (Huxtable, 1992). Excess taurine produced by animals is excreted in the form of taurine or bile salts (*e.g.* taurocholate) and oxidised by bacteria to complete the natural redox cycle of sulfur. Utilisation of taurine is of great importance for the survival of bacteria and balance of the sulfur cycle. *E. coli* employs the TauABC transporter to import taurine, and the dioxygenase TauD to release sulfite (Figure 2.1.2.2). Disruption of each of the *tauABCD* genes results in a loss of the ability to utilise taurine as a sulfur source but does not affect the utilisation of a range of other aliphatic sulfonates (Eichhorn *et al.*, 2000). Taurine utilisation was restored when complementing *tauABC* mutants with a clone of the *tauABC* locus (Ploeg *et al.*, 1996).

In *E. coli*, *tauABC* and *ssuACB* encode ABC importers for the uptake of taurine and other alkanesulfonates, respectively (Ploeg *et al.*, 1996; Ploeg *et al.*, 1999). TauD is an α -ketoglutarate-dependent dioxygenase that oxidises taurine and releases sulfite (Eichhorn *et al.*, 1997). SsuD is a reduced Flavin mononucleotide (FMN)-dependent monooxygenase that desulfonates a wide range of substrates (Eichhorn *et al.*, 1999). FMN is supplied by an NAD(P)H: FMN oxidoreductase, SsuE. TauA and SsuA both contain a N-terminal signal sequence, indicating that both proteins are located in the periplasm and function as periplasmic substrate-binding proteins (SBPs). TauB and SsuB function as the cytoplasmic ATP-binding domains (NBDs), while TauC and SsuC are the cognate transmembrane domains (TMDs). TauBC and SsuBC are both homodimeric ABC transporters that utilise the energy of ATP hydrolysis to translocate substrates across the inner membrane. The subunit compositions of these ABC importers are TauA(BC)₂ and SsuA(CB)₂. It is proposed that TauA and SsuA capture their substrates from the periplasm and direct them to TauC and SsuC, respectively. Upon ATP binding and hydrolysis by TauB and SsuB, the substrates are translocated across the inner membrane to the cytoplasm, where they are oxidised by the oxygenases, leading to the release of sulfite. Deletion studies have revealed that *E. coli* TauABC and SsuABC are complementary in substrate transport (Ploeg *et al.*, 1996).

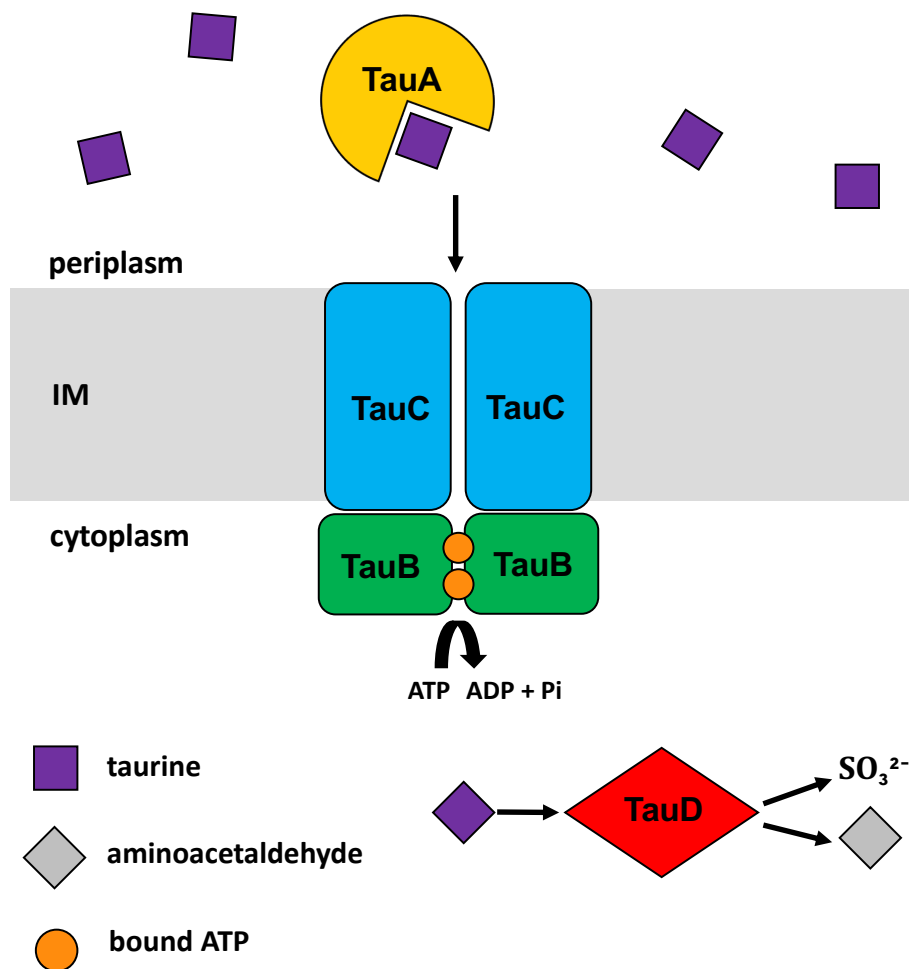


Figure 2.1.2.2. Taurine uptake and oxidation pathway. It has been proposed that TauA captures taurine with high affinity from the periplasm and directs it to TauC. Upon ATP binding and hydrolysis by TauB, taurine is translocated across the inner membrane to the cytoplasm where it is subsequently oxidised by TauD, leading to the release of sulfite and aminoacetaldehyde.

2.1.3. Importance and future perspectives

A diverse range of bacteria and fungi in soil and freshwater can degrade taurine and use it as their sole source of sulfur, carbon and energy (King & Quinn, 1997). In addition to *E. coli*, the taurine-sulfur assimilation system is also present in pathogenic bacteria such as *Yersinia pestis* and *Shigella dysenteriae*, as well as nitrogen-fixing bacteria such as *Rhizobium etli* and *Azospirillum brasilense*. The wide distribution of the Tau system among bacterial species in different natural environments indicates its importance in bacterial survival and physiology.

Structural and biochemical studies have provided insights into the architecture and mechanism of nutrient importers in bacteria. Given the large number of ABC transporters and their broad spectrum of substrates, it is likely that our knowledge of ABC transporters is still limited to more representative members. Further information is needed to validate existing transport models proposed for ABC transporters. Nutrient binding to SBPs is the first step in the substrate translocation process mediated by ABC importers and is the key determinant for substrate recognition and specificity (Locher, 2016). Until now, structures and functions of SBPs of various classes have been revealed (Berntsson *et al.*, 2010). TauA is an SBP found in a wide range of bacterial species and is involved in taurine recognition and cellular internalisation (Ploeg *et al.*, 1996; Eichhorn *et al.*, 2000). The crystal structures of apo SsuA from *E. coli* (Beale *et al.*, 2010) and alkanesulfonate-bound SsuA from *Xanthomonas axonopodis* pv. *citri* (de Araújo *et al.*, 2013) have validated the “Venus fly-trap” mechanism of substrate binding in SsuA, a close homologue of TauA. However, no structural or direct biochemical information is available for TauA. The binding mode between TauA and its physiological substrate taurine is unknown. The molecular details underlying substrate selectivity of the Tau and Ssu systems remain elusive. Finally, further investigation is required to improve our understanding of the TauABC architecture, and how the SBP TauA is coupled to the NBDs and TMDs

2.1.4. Aim of research

Deletion analysis shows that the Ssu system is responsible for uptake of a range of different aliphatic sulfonates while the Tau system has a higher specificity for taurine (Eichhorn *et al.*, 2000). SBPs have been extensively studied, which provides some indications into the structure and mechanism of TauA. However, no direct data regarding how the Tau system has such as strict substrate preference are available. The aim of this research is to provide structural and functional information into the molecular mechanism governing substrate recognition and binding of TauA.

The specific objectives of this project are as follows:

1. To determine structure of TauA in complex with taurine.
2. To determine the binding affinity between TauA and taurine using ITC.
3. To validate key residues in taurine binding using site-directed mutagenesis.
4. To illustrate the molecular mechanism that underpins substrate selectivity of TauA.
5. To understand mechanism of the different substrate specificities between TauA and SsuA.
6. To functionally characterise the cognate TauBC transporter.

2.2. Materials and Methods

2.2.1. *TauA* cloning

The gene fragment encoding the mature TauA of *E. coli* (accession number: Q47537) was amplified using *E. coli* chromosomal DNA as a template. Primers *tauAFw* and *tauARv* were designed and purchased from Invitrogen (Table 2.2.1.1). The amplified *tauA* gene was inserted into the pEHisTEV vector (Liu & Naismith, 2009) through the NcoI and XhoI restriction sites. This resulted in a construct encoding an N-terminal His₆-tagged TauA protein with a Tobacco Etch Virus (TEV) protease cleavage site (Figure 2.2.1.1). Three additional amino acids remain at the N-terminus after the mature protein owing to the linker and the TEV recognition protease site. The pEHisTEV-*tauA* plasmid was verified by DNA sequencing (Source BioScience).

Primer	Sequence
<i>tauAFw</i>	5'-CATGCCATGGTGAACGTCACCGTGGCG-3'
<i>tauARv</i>	5'-CCGCTCGAGTTATTGCACGAAGCGCGAG-3'

Table 2.2.1.1. Primer sequences for *tauA* gene.

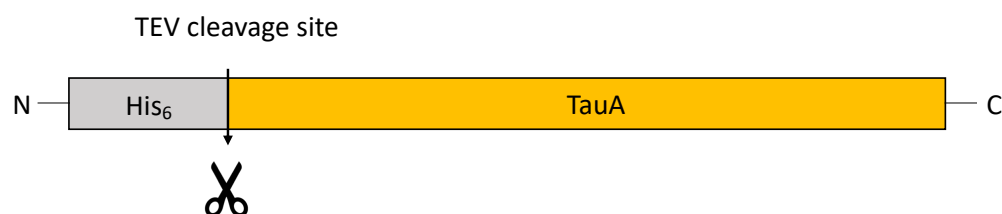


Figure 2.2.1.1. Schematic representation of the His₆-TEV-TauA construct. The construct encodes an N-terminal His₆-tag, a TEV cleavage site and the TauA protein. The signal peptide of TauA was removed to ensure expression of TauA in the cytoplasm.

2.2.2. TauA overexpression

E. coli BL21 (DE3) cells transformed with pEHisTEV-*tauA* plasmid were plated onto Lysogeny broth (LB)-agar plate supplemented with 50 µg/mL kanamycin and incubated overnight (16-18 hours) at 37 °C. A single colony was picked and used to inoculate 100 mL of LB media supplemented with 50 µg/mL kanamycin. The starter culture was grown at 37 °C with 220 rpm shaking overnight. 10 mL of the overnight starter culture was used to inoculate 1 L of LB media supplemented with 50 µg/mL kanamycin. The culture was incubated at 37 °C with 220 rpm shaking. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM when the culture reached an OD₆₀₀ of ~0.6, and cell growth continued at 22 °C for 18 hours. Cells were harvested by centrifugation at 6,000 g (Beckman Coulter) for 10 minutes and stored at -80 °C until further use.

2.2.3. TauA purification

The whole cell pellet was resuspended in phosphate-buffered saline (PBS) buffer supplemented with 20 mM imidazole (pH 7.5), 3 mM MgCl₂, deoxyribonuclease (50 unit/mL), lysozyme (0.1 mg/mL) and PefaBloc (0.1 mg/mL). Cell-free lysate was prepared by sonication (10 cycles; 10-second per cycle; 10-second interval) (Soniprep150 Plus ultrasonic disintegrator (MSE)) followed by centrifugation (200,000 g for 1 hour) and filtration (0.22 µm) (Millipore). Sonication was performed on ice and the following purification steps were performed at 4 °C. For the immobilised metal affinity chromatography (IMAC) step, the supernatant was loaded on a His-Trap column (column volume (CV) of 5 mL) (GE Healthcare) pre-equilibrated with PBS containing 20 mM imidazole (pH 7.5). The column was washed with 20 CV of PBS

containing a linear imidazole concentration from 20 mM to 250 mM (pH 7.5). The ÄKTApurifier (GE Healthcare) system was used to monitor the UV_{280} absorbance of the flow through. The His₆-tagged TauA protein was eluted from PBS buffer containing 125 mM imidazole. The purity of different fractions was assessed by sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) with a 4-12% Bis-Tris protein gel (Invitrogen NuPAGE). Elution fractions with high purity were collected and TEV protease was added to the eluate with a protein-to-TEV ratio of 10:1 to cleave the His-tag. The sample was dialysed overnight in 2 L of buffer A (20 mM Tris (pH 7.5) and 150 mM NaCl).

The overnight cleaved sample was loaded on a His-Trap column pre-equilibrated with buffer A (reverse-IMAC). The flow through and 2 CV of buffer A wash were collected and concentrated to a final volume of 7 mL, and injected to the Superdex 200 prep grade HiLoad 16/60 (GE Healthcare) column, pre-equilibrated with buffer A. The ÄKTApurifier (GE Healthcare) system was used to monitor the UV_{280} absorbance of the fractions. Fractions containing TauA protein were collected and concentrated to 40 mg/mL using a 30 kDa molecular weight (MW) cut-off concentrator (Thermo Fisher Scientific).

2.2.4. Site-directed mutagenesis

With reference to the crystal structure (see later), residues of TauA that form important hydrogen bonds with taurine were identified by inspection of the taurine-binding site. These residues were mutated in order to better characterise taurine binding. Alanine-substituted TauA mutants (Gln30Ala, Glu106Ala, Thr132Ala and Asp205Ala) were generated by PCR using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies). The primers designed for mutagenesis are shown in Table 2.2.4.1. Mutations were confirmed by DNA

sequencing (Source Bioscience). Mutants of TauA were expressed and purified as described for wild-type TauA (section 2.2.3).

Mutation	Forward primer	Reverse primer
N30A	5'-GTTCGGCTGAGGTTGCAT ACGCCACGGTGACG-3'	5'-CGTCACCGTGGCGTATG CAACCTCAGCCGAAC-3'
E106A	5'-CTTTACCACCAGCGCTGC GGAGTTACCCAGTTT-3'	5'-AAACTGGGTAAGTCCGCA GCGCTGGTGGTAAAG-3'
T132A	5'-GGCTGTAGTGGGTGGCG GAGATAAACGGTAC-3'	5'-GTACCGTTTATCTCCGCC ACCCACTACAGCC-3'
D205A	5'-ACCACCCAGACGGCCAG CGTTGGCG-3'	5'-CGCCAACGCTGGCCGTC TGGGTGGT-3'

Table 2.2.4.1. Primer sequences for TauA mutants.

2.2.5. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was performed to characterise the binding of taurine to wild-type TauA and mutants as well as the binding of taurine analogues to TauA (MicroCal iTC₂₀₀ microcalorimeter, Malvern Instruments). Compounds 2-aminoethylphosphonic acid (2-AEP), N-(2-acetamido)-2-aminoethanesulfonic acid (ACES) and MES were selected as structural analogues of taurine for measurements (Figure 2.2.5.1).

Binding of ligand to protein was measured at 25 °C. The protein buffer contained 20 mM Tris (pH7.5) and 150 mM NaCl (ITC buffer). Protein and ligands were diluted to desired concentrations (Table 2.2.5.1) in ITC buffer. A protein concentration of 17.8 µM was used to test binding between TauA/TauA mutants and taurine, whereas a protein concentration of 450 µM was used to examine TauA binding to 2-AEP, ACES and MES. The pH of the protein and

ligand was adjusted at 4 °C until identical. 200 μL of protein and 60 μL of ligand were added into the sample cell and injection syringe, respectively. Seventeen 2.4 μL injections were performed with a 0.5 μL pre-injection, a 180-second interval and a 0.5 $\mu\text{L}/\text{second}$ injection speed. Control experiments were performed where ligand was injected into the ITC buffer, with the same experimental parameters. The heat of interaction was obtained by subtracting control measurements (heat associated with the dilution of the ligand) from the experimental measurements. The baseline was generated, and the concentration-normalized binding isotherm was integrated using the MicroCal Origin software (version 7) (Origin, 2002). During the initial baseline equilibration, the titrant can be diluted by the cell material and therefore the first integrated injection heat can be inaccurate. The first data point was removed from data fitting to generate a more accurate model. The isotherm was fitted to a single site binding model. Values of binding constant (K_a , M^{-1}), enthalpy change (ΔH , cal/mol) and binding stoichiometry (n) were allowed to vary (unless mentioned specifically) during data fitting. The dissociation constant (K_d , M) and entropy change (ΔS , $\text{cal/mol}/^\circ\text{C}$) are derived from K_a and ΔH according to the equations shown in Figure 2.2.5.2.

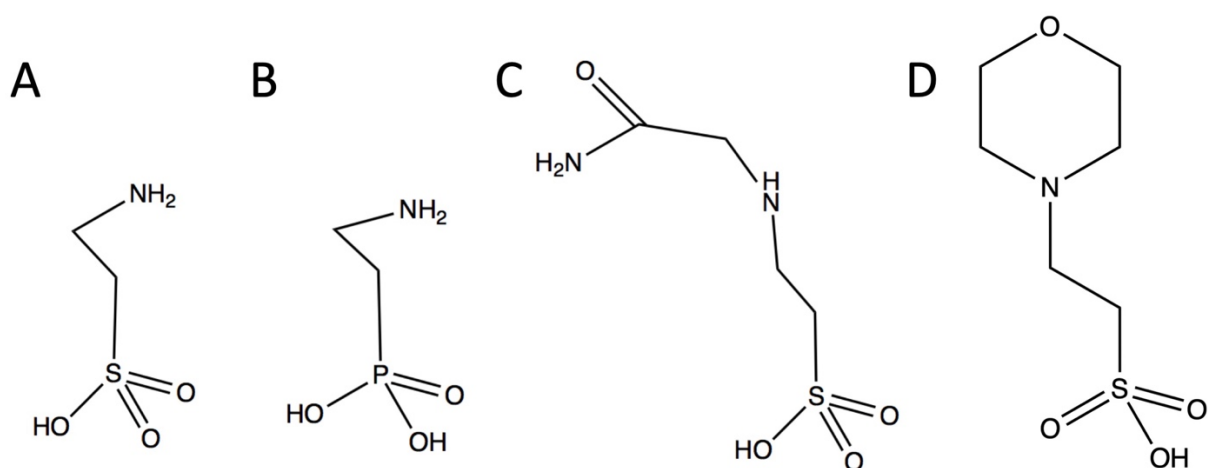


Figure 2.2.5.1. Taurine and its analogues. (A) taurine, (B) 2-AEP, (C) ACES, (D) MES.

		TauA^{WT}	TauA^{Q30A}	TauA^{E106A}	TauA^{T132A}	TauA^{D205A}
Taurine	[Protein] (μM)	17.8	17.8	17.8	17.8	17.8
	[Taurine] (μM)	120	360	720	720	720
2-AEP	[Protein] (μM)	450				
	[2-AEP] (μM)	4500				
ACES	[Protein] (μM)	450				
	[ACES] (μM)	4500				
MES	[Protein] (μM)	450				
	[ACES] (μM)	4500				

Table 2.2.5.1. Concentrations of TauA, TauA mutants and ligands tested in ITC.

$$K_d = \frac{1}{K_a}$$

$$\Delta G = -RT \ln K_a = \Delta H - T \Delta S$$

Figure 2.2.5.2. Equations for calculation of K_a and ΔS . K_d : the dissociation constant (M); K_a : the binding constant (M^{-1}); ΔG : the Gibbs free energy change (J); R : the gas constant; T : the absolute temperature. ΔH : the enthalpy change (cal/mol); ΔS : the entropy change (cal/mol/°C).

2.2.6. Crystallisation

Crystallisation trials were carried out using the vapour diffusion sitting drop method on 96-well CrystalQuick plates (Grenier). Apo TauA as well as TauA in complex with taurine were both tested. For apo TauA, the purified protein at a concentration of 40 mg/mL (1.2 mM) was used for crystallisation directly. For the TauA-aurine complex, purified TauA at a concentration of 40 mg/mL (1.2 mM) was incubated with 1.5 mM taurine at 20 °C for 1 hour prior to crystallisation. Crystallisation conditions were screened against the JCSG-plus (Molecular Dimensions), PACT-premier (Molecular Dimensions) and Index (Hampton Research) screening plates at 20 °C and 4 °C. The reservoir volume was 100 µL and the sitting drop consisting of 100 nL of protein and 100 nL of reservoir solution was dispensed by the Mosquito (TPP Labtech), a robotic dispensing system. Apo TauA did not crystallise. TauA-aurine crystals grew in condition containing 0.2 M sodium iodide, 0.1 M Bis-Tris propane (pH 8.5) and 20% (w/v) PEG 3350. Crystallisation conditions were optimised on 96-well CrystalQuick low profile plates. Different pH (± 0.5), temperature (20 °C and 4 °C) as well as different protein (5 mg/mL, 20 mg/mL and 40 mg/mL) and PEG concentrations ($\pm 4\%$ (w/v) with 2% (w/v) interval) were screened against the original condition. Crystallisation optimisation was also performed on hanging drop 24-well XRL plates by mixing protein and reservoir solution at a ratio of 1:1.

Compounds 2-AEP, ACES and MES are structural analogues of taurine (Figure 2.2.5.1). Co-crystallisation of TauA (20 mg/mL) with 10 mM of 2-AEP, ACES and MES were performed as described for TauA-aurine crystals.

2.2.7. Data collection and structure determination

Crystals were cryoprotected by transferring crystals to crystallisation buffer supplemented with 20% glycerol (v/v) and flash-cooled in liquid nitrogen. X-ray data were collected on the I03, I23 and I24 beamlines at the Diamond Light Source. The taurine-bound TauA crystal diffracted X-rays to 1.31 Å resolution. TauA-aurine crystals were formed in the space group $P2_12_12_1$. In order to obtain the phases, molecular replacement was performed using apo SsuA from *E. coli* (PDB ID: 2X26, (Beale *et al.*, 2010)) and MES/MOPS/HEPES-bound SsuA from *X. axonopodis pv. citri* (PDB ID: 3KSJ, 3KSX and 3E4R, (de Araújo *et al.*, 2013)) as search models. However, no structure solution was found by molecular replacement. A single-wavelength anomalous diffraction (SAD) experiment was performed. Although TauA does not contain a sulfur-containing residue (*i.e.* methionine or cysteine), 1.5 mM of taurine was added during crystallisation and it was highly likely that taurine is bound to the binding site of TauA. Therefore, sulfur-SAD was initially attempted, with the wavelength tuned to 2.7552 Å.

The crystal diffracted X-rays to the edge of the detector (1.77 Å resolution). In total 3×360° of data were collected using the inverse-beam method (10° wedges) from a single crystal. Data were indexed, integrated and scaled using XDS and XSCALE (Kabsch, 2010). The heavy atom substructure and initial electron density map were obtained with the program suite SHELEXC/D/E (Sheldrick, 2009). Since the TauA crystals grew in high iodide salt conditions (0.2 M sodium iodide), a strong anomalous signal for the iodide ions was observed during data collection. In fact, the anomalous scattering of iodide ions (10.9 e⁻) was ~seven times stronger than obtained for sulfur atoms (1.5 e⁻) at 2.7552 Å wavelength. Iodide-SAD was therefore performed. The initial phases were fed to Buccaneer (Cowtan, 2006) and 83.4% of the initial model was built. The initial model was used as a search model for molecular replacement of the TauA-aurine structure in Phaser (McCoy *et al.*, 2007). Iterative cycles of model building

and refinement were completed using Refmac5 (Murshudov *et al.*, 2011) and COOT (Emsley & Cowtan, 2004). The orientation of taurine in the binding pocket is unambiguously defined by the clear electron density maps of both TauA molecules in the asymmetric unit. Water molecules were added by ARP/Warp (Langer *et al.*, 2008) and manually checked in COOT. The final TauA-aurine structure was refined at 1.31 Å resolution with an R-factor of 13.1% and R-free of 17.6%. Structures of TauA-2AEP (1.62 Å), TauA-ACES (1.5 Å) and TauA-MES (1.5 Å) were solved by molecular replacement using the refined taurine-bound TauA structure as a search model.

2.2.8. TauBC protein production

2.2.8.1. *tauBC* cloning

The gene fragments encoding TauBC of *E. coli* (accession numbers: Q47538 (TauB) and Q47539 (TauC)) were amplified using *E. coli* chromosomal DNA as a template. Primers *tauBC*-Fw and *tauBC*-Rv were designed and purchased from Invitrogen (Table 2.2.8.1). The amplified *tauBC* gene was inserted into the pEHsTEV vector (Liu & Naismith, 2009) through the NcoI and XhoI restriction sites. This resulted in a construct of an N-terminal His₆-tagged TauBC protein with a TEV protease cleavage site (Figure 2.2.8.1). The pEHsTEV-*tauBC* plasmid was verified by DNA sequencing (Source BioScience).

Primer	Sequence
<i>tauBC</i> -Fw	5'- CATGCCATGGCTCTGCAAATCTCTCATC -3'
<i>tauBC</i> -Rv	5'- CCGCTCGAGTCATTGTACTTCTCCATGCCAG -3'

Table 2.2.8.1. Primer sequences for *tauBC* gene.

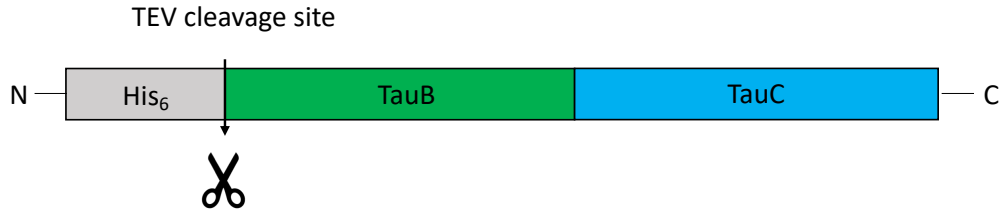


Figure 2.2.8.1. Schematic representation of the His₆-TEV-TauBC construct.

2.2.8.2. TauBC overexpression

E. coli C43 (DE3) cells transformed with pEHisTEV-*tauBC* plasmid were plated onto LB-agar plate supplemented with 50 µg/mL kanamycin and incubated overnight at 37 °C. A single colony was picked and used to inoculate 100 mL of LB media supplemented with 50 µg/mL kanamycin. The starter culture was grown at 37 °C with 220 rpm shaking overnight. 10 mL of the overnight starter culture was used to inoculate 1 L of LB media supplemented with 50 µg/mL kanamycin. The culture was incubated at 37 °C with 220 rpm shaking. IPTG was added to a final concentration of 1 mM when the culture reached an OD₆₀₀ of ~0.6, and cell growth continued at 25 °C for 18 hours. Cells were harvested by centrifugation at 6,000 g for 10 minutes and stored at -80 °C until further use.

2.2.8.3. TauBC membrane preparation

The whole cell pellet was resuspended in PBS buffer supplemented with 3 mM MgCl₂, deoxyribonuclease (50 unit/mL), and PefaBloc (0.1 mg/mL). The cell resuspension was passed through a cell disruptor (Constant Systems) for two cycles (22/25 kpsi) and centrifuged at 15,000 g for 10 minutes to remove unbroken cells and cellular debris. During cell breakage,

the sample was kept on ice. Supernatant containing cell membranes was recovered and subjected to an ultracentrifuge step at 180,000 g for 1 hour. Membrane pellets were resuspended in PBS buffer with a Dounce glass homogeniser (Fisher Scientific) on ice.

2.2.8.4. TauBC purification

Following the preparation of total membranes, TauBC was solubilised with 1% (w/v) n-Dodecyl- β -D-Maltoside (DDM) (Anatrace) by 150 rpm stirring for 1 hour at 4 °C. After an ultracentrifugation step (180,000 g for 1 hour at 4 °C), supernatant containing TauBC was recovered for further purification. For IMAC, the sample was incubated with 10 mL of Ni-NTA SuperFlow resin (Qiagen) and 40 mM imidazole (pH 7.5), stirring gently at 150 rpm for 4 hours at 4 °C. The protein-nickel slurry was then loaded onto a 50-mL glass polyprep column (Bio-Rad) and washed with 10 CV of wash buffer A (40 mM imidazole (pH 7.5), 0.1% DDM, PBS) and 10 CV of wash buffer B (60 mM imidazole (pH 7.5), 0.1% DDM, PBS), followed by protein elution with elution buffer (250 mM imidazole (pH 7.5), 0.1% DDM, PBS). The IMAC was performed at 4 °C. TEV protease was added to the eluate with a protein-to-TEV ratio of 1:1 to cleave the His-tag. The sample was dialysed overnight at 4 °C in 2 L of dialysis buffer (20 mM Tris (pH 7.5), 150 mM NaCl and 0.03 % DDM).

The overnight cleaved sample was loaded on a His-Trap column pre-equilibrated with dialysis buffer. The cleaved TauBC protein was eluted with dialysis buffer containing 40 mM imidazole (pH 7.5). For the size exclusion chromatography (SEC) step, fractions containing TauBC were collected and concentrated to a final volume of 500 μ L and injected to the Superdex 200 10/300 GL (GE Healthcare) column, pre-equilibrated with dialysis buffer. The ÄKTApurifier system was used to monitor the UV₂₈₀ absorbance of the fractions. Fractions containing TauBC protein

were collected and concentrated to 15 mg/mL using a 100 kDa MW cut-off concentrator (Thermo Fisher Scientific).

2.2.9. Reconstitution into proteoliposomes

For the basal ATPase activity test, the TauBC protein was diluted in a buffer containing 20 mM Tris (pH 7.5), 150 mM NaCl and 0.03 % DDM to a final concentration of 0.5 mg/ml (8.62 μ M) (final volume of 150 μ L). To test the stimulated ATPase activity, TauBC, TauA and taurine were diluted in buffer containing 20 mM Tris (pH 7.5), 150 mM NaCl and 0.03 % DDM to final concentrations of 8.62 μ M, 330 μ M and 3.3 mM, respectively (final volume of 150 μ L). A 20 mg/mL *E. coli* polar lipid (Avanti polar lipids) was prepared by dissolving the lipids in H₂O followed by water bath ultrasonication for 10 min. 30 μ L of the phospholipid was then added to the protein and incubated for 5 min on ice. The solution was further diluted in 6 mL of liposome buffer (50 mM Tris (pH 7.5) and 50 mM KCl) and incubated again for 5 min on ice. Proteoliposomes were pelleted by 45,000 g ultracentrifugation for 30 min at 4°C. The supernatant was removed, and the pellets were resuspended in 100 μ L liposome buffer. The total amount of reconstituted TauBC was calculated using the Pierce Bicinchoninic Acid (BCA) Protein Assay Kit (Thermo Scientific) as described in the manufacturer's instructions. The proteoliposomes were immediately used in ATPase assays.

2.2.10. ATPase activity assay in liposomes

The ATPase assays were performed as described in the EnzChek Phosphate Assay Kit (Molecular Probes) according to manufacturer's instructions. The ATPase assay using this kit measures production of Pi during ATP hydrolysis. The kit consists of a substrate, 2-amino-6-mercapto-7-methylpurine riboside (MESG), a purine nucleoside phosphorylase (PNP) and a reaction buffer. PNP catalyses the reaction between the produced Pi and MESG, leading to the formation of ribose 1-phosphate and 2-amino-6-mercapto-7-methyl-purine. The enzymatic conversion from MESG to 2-amino-6-mercapto-7-methyl-purine results in a spectrophotometric shift in absorbance from 330 nm for the substrate to 360 nm for the product. Each reaction was done in a total reaction volume of 100 μ L with a final TauBC concentration of 1 μ M (5.8 μ g total protein). ATPase assay components are shown in Table 2.2.10.1.

Components	In detergent		In proteoliposomes	
	Without TauA/taurine	With TauA/taurine	Without TauA/taurine	With TauA/taurine
TauBC (10 μ M)	10 μ L	10 μ L	10 μ L	10 μ L*
TauA (300 μ M)	3.3 μ L	0	0	0
Taurine (5 mM)	2 μ L	0	0	0
MESG	20 μ L	20 μ L	20 μ L	20 μ L
PNP	1 μ L	1 μ L	1 μ L	1 μ L
Reaction buffer (20x)	5 μ L	5 μ L	5 μ L	5 μ L
DDM (1%)	3 μ L	3 μ L	0	0
H ₂ O	55.7 μ L	61 μ L	64 μ L	64 μ L

Table 2.2.10.1. Components of the ATPase assay using the EnzChek assay kit. *: TauA and taurine were added to TauBC prior to proteoliposome reconstitution.

Absorbance at 360 nm (A_{360}) was monitored in a 96-well black clear-bottom plate (Greiner) at 25 °C using a SpectraMax microplate spectrofluorometer (Molecular Devices). The reaction mixture was incubated at 25 °C for 10 minutes, allowing the components to equilibrate. The reaction was initiated by adding 1 μ l of 100 mM ATP solution (final concentration of 1 mM) using a multichannel pipette. A standard curve was prepared previously in the Beis group where KH_2PO_4 solution was used as the source of Pi. A linear line was plotted where $A_{360}=0.0214 \times \text{Pi } (\mu\text{M})$. Changes in A_{360} during the linear part of the reaction were equated to liberation of phosphate (derived from the standard curve gradient) to determine the overall rate of ATP hydrolysis (V_{max}). The equations used to determine the rate of ATPase activity (nmol/min/mg) are shown in Figure 2.2.10.1. The calculation is based on each reaction mixture containing 5.8 μ g of TauBC.

$$\frac{\text{Change in Absorbance (360 nm)}}{0.0214} = \text{Liberated Pi (nmoles)}$$

$$\frac{\text{Liberated Pi (nmoles)}}{\text{Time (min)} \times \frac{5.8}{1000} \text{ mg}} = \text{Rate of ATPase activity (nmoles/min/mg)}$$

Figure 2.2.10.1. Equations for calculation of the rate of ATPase activity.

2.3. Results

2.3.1. Purification of TauA and mutants

The sequence of TauA from *E. coli* possesses a signal peptide that is 22 amino acid residues in length. Signal peptides are N-terminal peptides (usually 16-30 residues long) of newly synthesised secretory and membrane proteins, which direct the proteins to their destinations (Kapp *et al.*, 2013). The presence of the signal peptide in TauA indicates that TauA functions as a periplasmic binding protein. The signal peptide was removed to ensure expression of the recombinant TauA in the cytoplasm. Following the IMAC and reverse-IMAC steps, SEC was performed for wild-type TauA using the Superdex 200 prep grade HiLoad 16/600 column. TauA was eluted at 90 mL as a monodisperse peak (Figure 2.3.1.1A). The theoretical MW of mature TauA is 31.975 kDa. According to the column data specification, myoglobin (17 kDa) elutes at a volume of 95 mL and ovalbumin (44 kDa) elutes at a volume of 82 mL. Based on these protein standards, the observed elution volume of TauA is consistent with a monomeric species, in agreement with the oligomeric states of SBPs. The single band from SDS-PAGE indicated that pure TauA was obtained (Figure 2.3.1.1B).

TauA mutants were expressed and purified as described for wild-type TauA. Due to the lower overexpression level, TauA mutants were purified using the Superdex 200 10/300 GL column, which has a smaller working volume. The chromatographic profiles show that the purified TauA mutants were monodisperse and stable (Figure 2.3.1.2A). An almost identical monodisperse peak at ~16 mL was observed for all four TauA mutants. The data file for the Superdex S200 10/300 GL column shows that an elution peak at 16 mL corresponds to ~32 kDa in MW, which is in agree with the expected size of TauA mutants. Figure 2.3.1.2A shows

the chromatographic profile for the Gln30Ala mutant. SDS-PAGE showed that all four TauA mutants have been purified to a level comparable to wild-type TauA (Figure 2.3.1.2B).

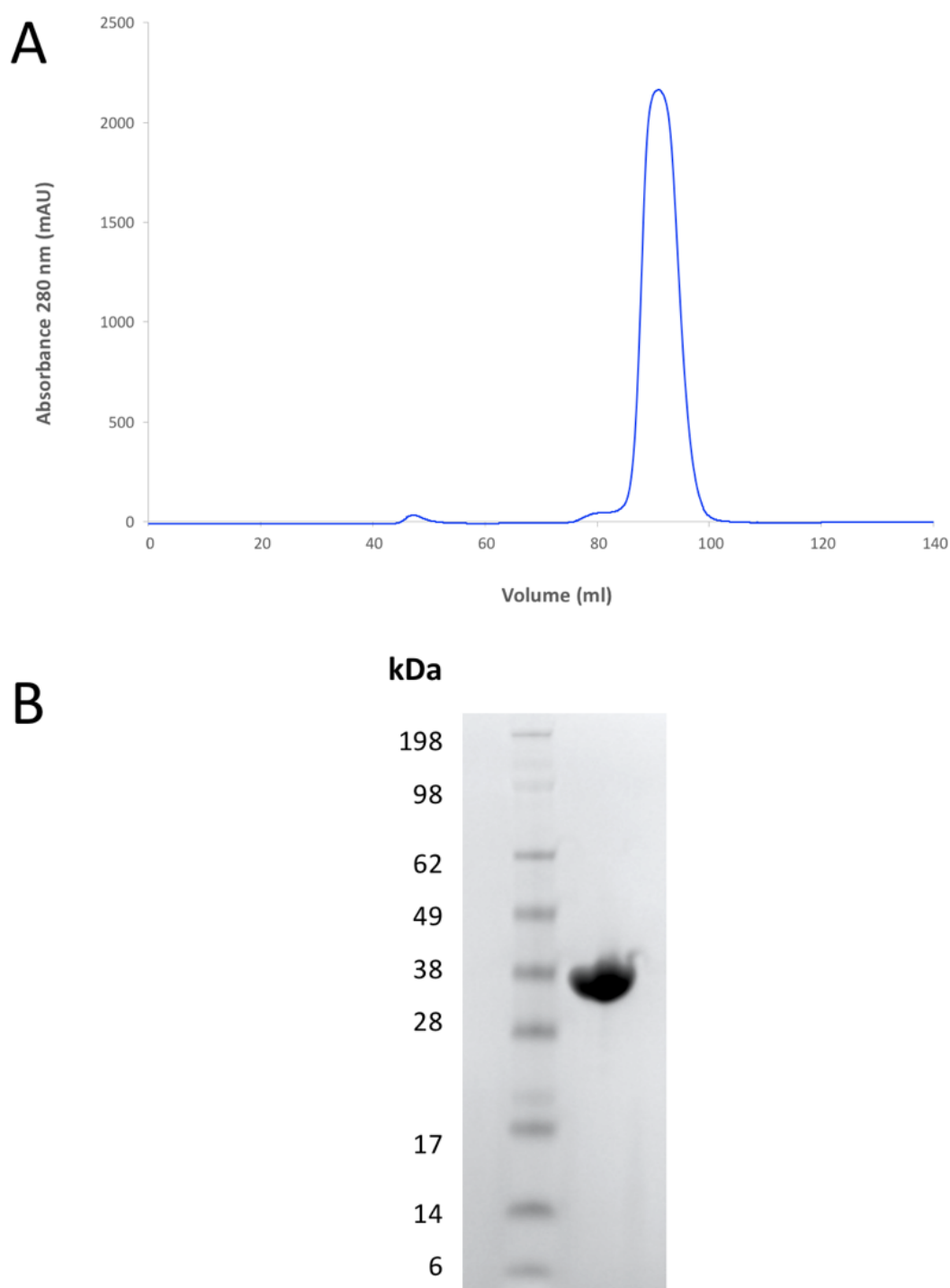


Figure 2.3.1.1. Purification of TauA. (A) SEC profile for TauA using the Superdex 200 prep grade HiLoad 16/600 column shows that TauA elutes at ~90 mL. According to manufacturer's data file, the elution volume of TauA is consistent with a monomeric species. (B) SDS-PAGE shows a single band at ~32 kDa corresponding to the expected MW of TauA.

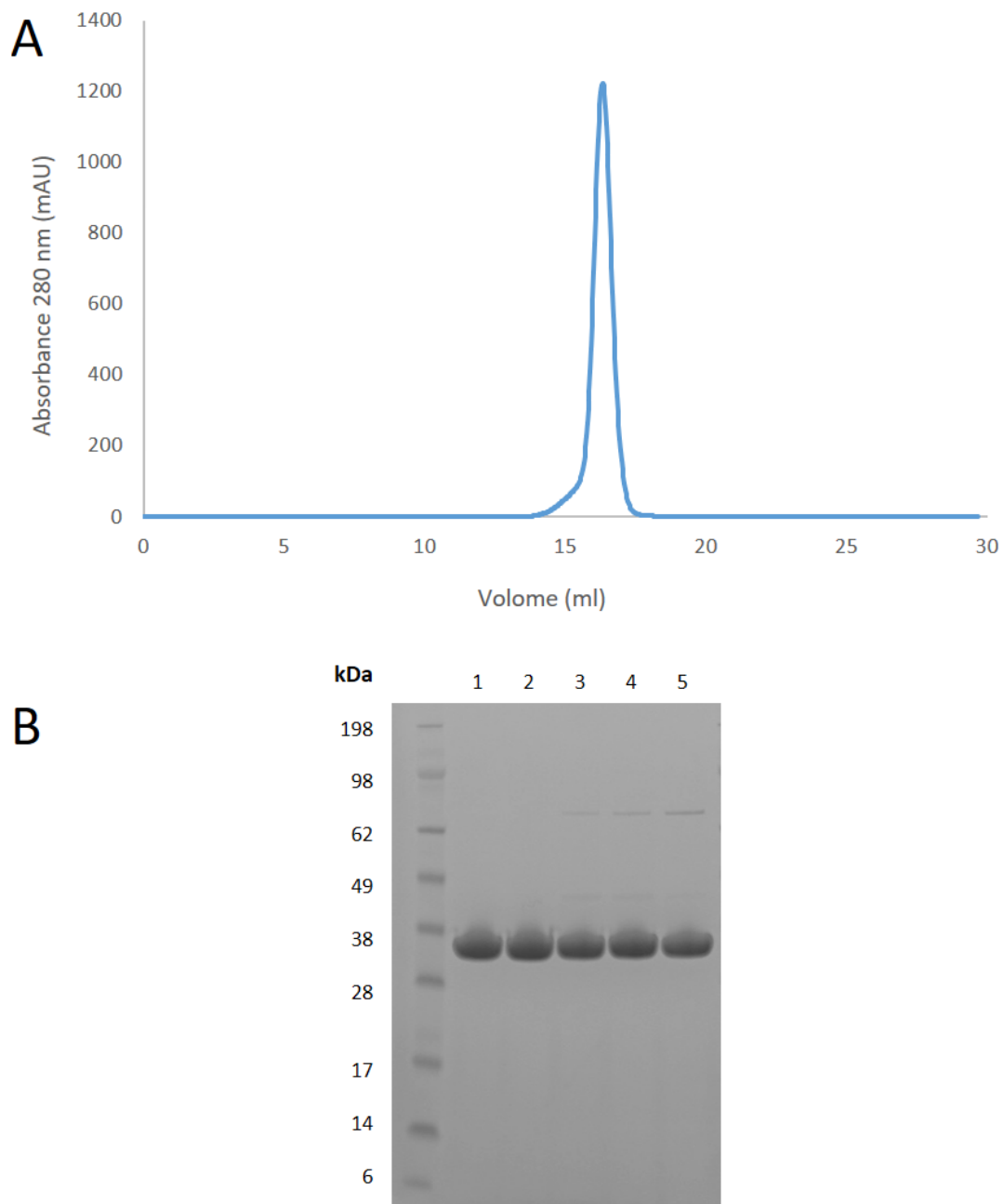


Figure 2.3.1.2. Purification of TauA mutants. (A) SEC profile for the Gln30Ala mutant using the Superdex 200 10/300 GL column. Almost identical monodisperse peaks were observed for other TauA mutants (Gln30Ala, Thr132Ala and Asp205Ala). Such a peak corresponds to ~30 kDa in MW, corresponding to the expected MW of TauA mutants. (B) SDS-PAGE of the purified wild-type and mutant TauA proteins. Lane 1: wild-type TauA, lane 2: Gln30Ala mutant, lane 3: Glu106Ala mutant, lane 4: Thr132Ala mutant, lane 5: Asp205Ala mutant.

2.3.2. TauA crystallisation

Apo TauA did not yield any crystals. This might be owing to the high flexibility of TauA which is unfavourable for crystallisation. TauA-aurine crystallised in the H3 condition of the PACT-premier screening plate, which contained 0.2 M sodium Iodide, 0.1 M Bis-Tris propane (pH 8.5) and 20% (w/v) PEG 3350. These crystals appeared as stacked thin planes after 4 °C incubation for five hours and grew to maximum size in 24 hours. Upon optimisation, single crystals appeared in 0.25 M sodium iodide, 0.1 M Bis-Tris propane (pH 8.5) and 24% (w/v) PEG 3350 at 20 °C (Figure 2.3.2.1). The TauA-aurine crystals diffracted X-rays to up to 1.31 Å. TauA in complex with 2-AEP, ACES and MES were crystallised in the same condition. These crystals appeared after 2 days and grew to maximum size after a week.

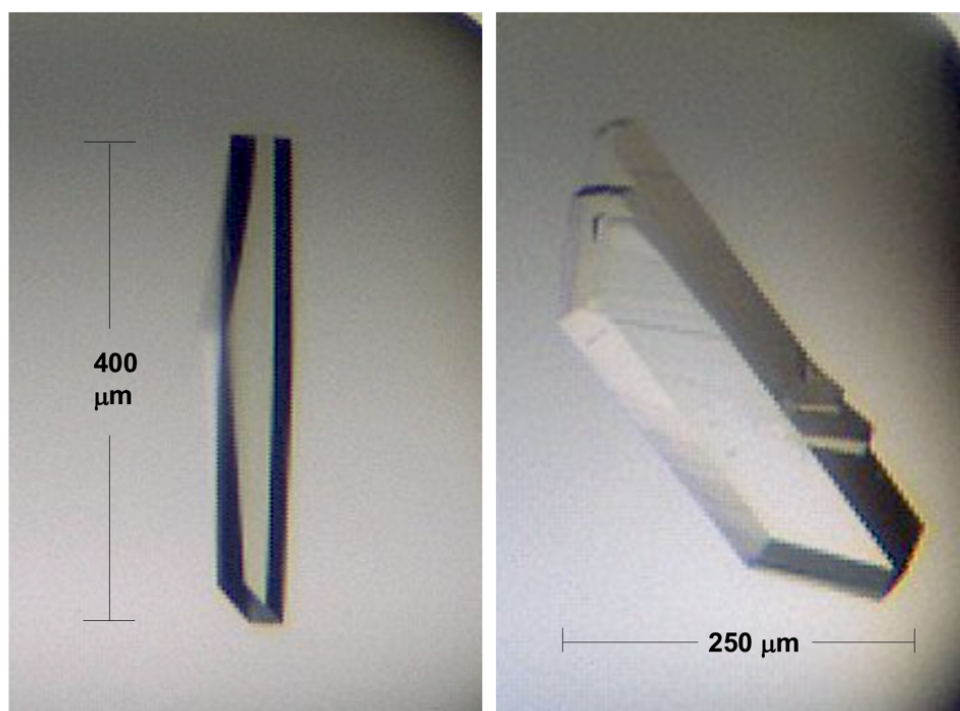


Figure 2.3.2.1. TauA-aurine crystals. Crystallisation condition: 0.25 M sodium iodide, 0.1 M Bis-Tris propane (pH 8.5) and 24% (w/v) PEG 3350 at 20 °C.

2.3.3. Structure of TauA in complex with taurine

Prior to the iodide-SAD experiment, molecular replacement was performed using apo SsuA from *E. coli* (PDB ID: 2X26, (Beale *et al.*, 2010)) and MES/MOPS/HEPES-bound SsuA from *X. axonopodis pv. citri* (PDB ID: 3KSJ, 3KSX and 3E4R, (de Araújo *et al.*, 2013)) as search models. However, no structure solution was found by molecular replacement. The lack of solution may be explained by three possible reasons. As discussed in section 1.0.5.6, SBPs share low sequence homology (often < 20%) and vary in size (25-70 kDa). Additionally, SBPs are flexible proteins that alternate between open and closed conformation. Finally, SBPs are formed of two domains each composed of a central β -sheet flanked by a number of α -helices. Polypeptide chains in β -sheet are rarely fully extended. In partially extended β -sheets, structural twists are formed, and the energetically preferred dihedral angles diverge greatly. Therefore, finding a suitable search model matching the overall conformation of taurine-bound TauA proved challenging.

For iodide-SAD, the dataset of native TauA-aurine crystal was collected at 2.7552 Å wavelength. Iodide ions form a wide variety of iodide ion-protein interactions with numerous hydrophobic or positively charged residues of TauA (Figure 2.3.3.1). For example, iodide ion 1 is coordinated by the positively charged Lys123 at protein surface. Iodide ion 2 is stabilised by three hydrophobic residues, Ile249, Val263 and Pro264, whereas iodide ion 3 forms interactions with backbone amides of Ser131, Thr132, Thr133, His134 and Tyr135. The space group of TauA-aurine crystal was determined to be $P2_12_12_1$ with two copies of TauA in the asymmetric unit (Figure 2.3.3.1). The native TauA-aurine dataset was collected at 0.9763 Å wavelength and refined at 1.31 Å resolution (Figure 2.3.3.2). Residues 23-319 were built in the electron density (Figure 2.3.3.1). Taurine was located at the substrate-binding site of both TauA

molecules in the asymmetric unit (Figure 2.3.3.3). Data collection and refinement statistics are summarised in Table 2.3.3.1.

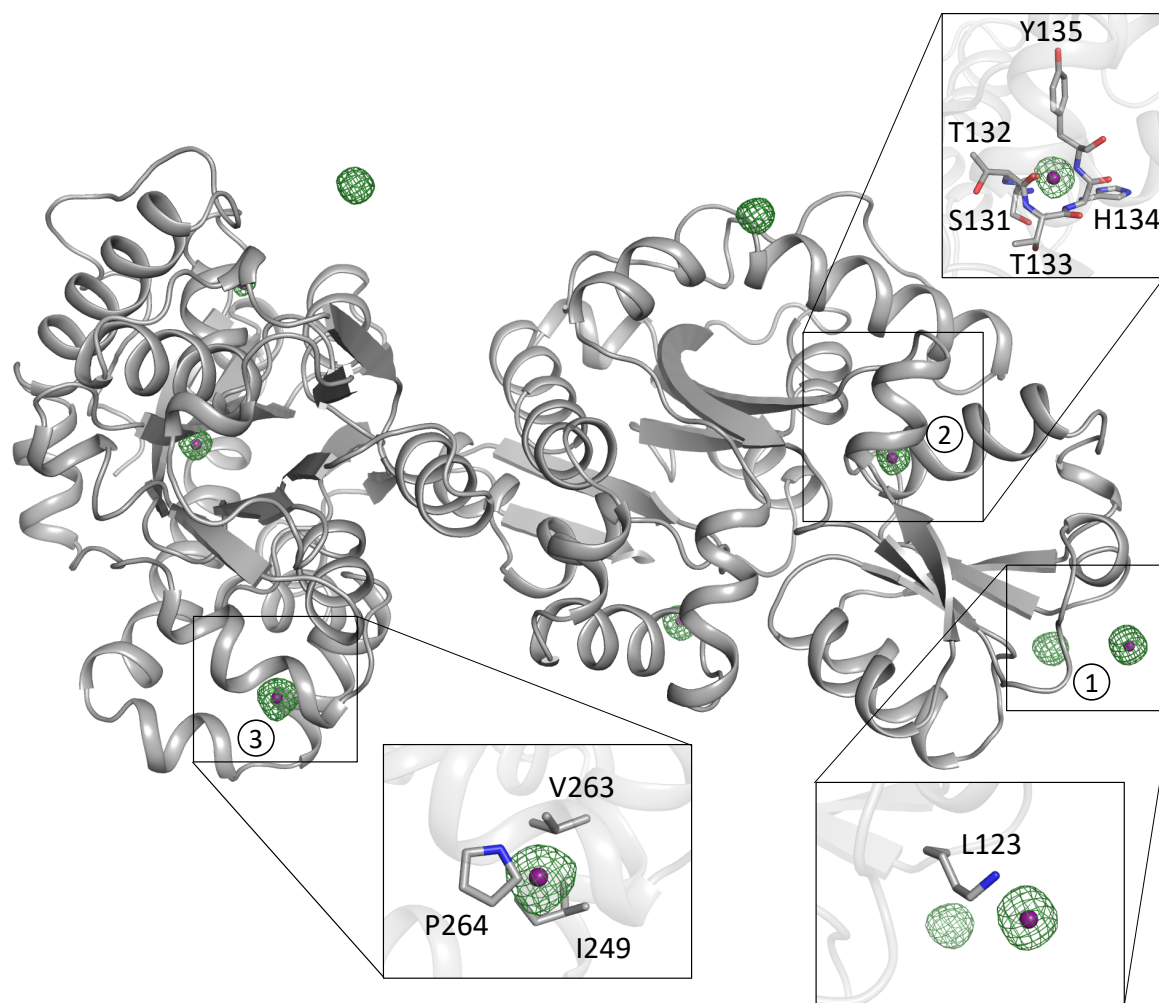


Figure 2.3.3.1. TauA anomalous difference maps. Anomalous difference maps of iodide ions contoured at 5 σ (green mesh). The two molecules of TauA in the asymmetric unit are shown in ribbon representation. The iodide ions are shown as purple spheres. The circled numbers indicate the different iodide ions. The insets show details of the iodide ion-protein interactions.

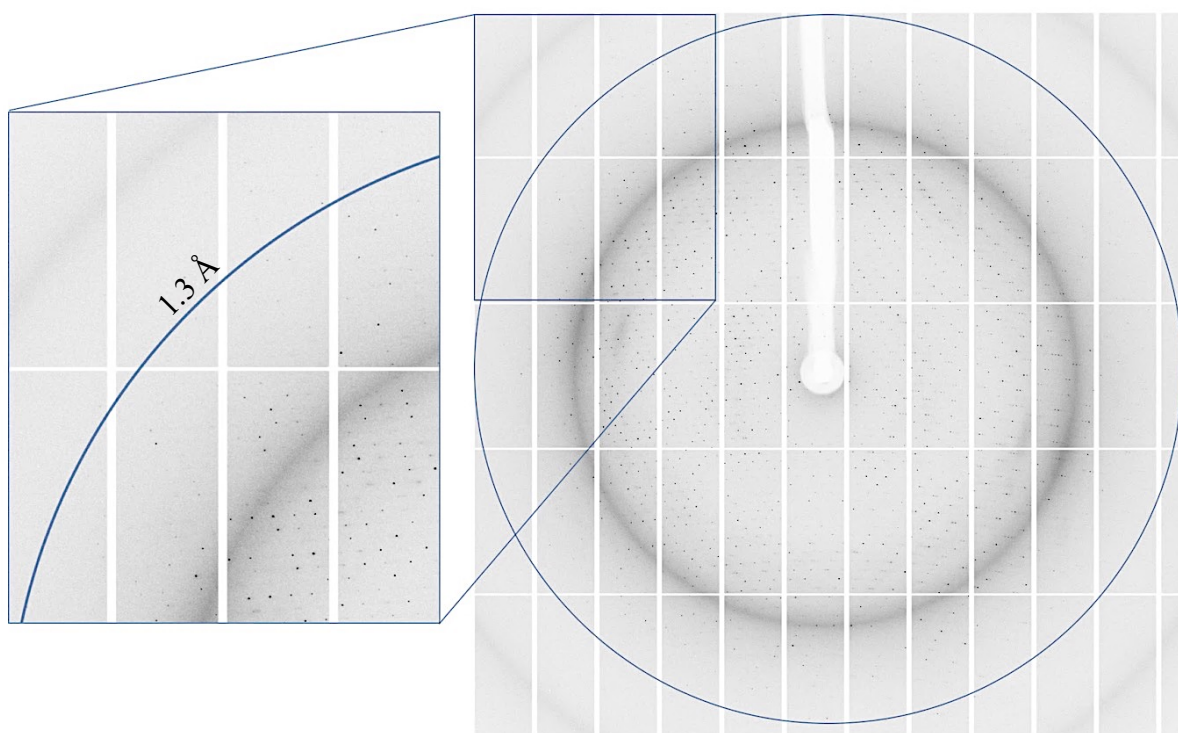


Figure 2.3.3.2. X-ray diffraction pattern for TauA-aurine crystal at 1.31 Å. The blue circle shows the diffraction limit at 1.3 Å and the inset shows the diffraction spots at high resolution.

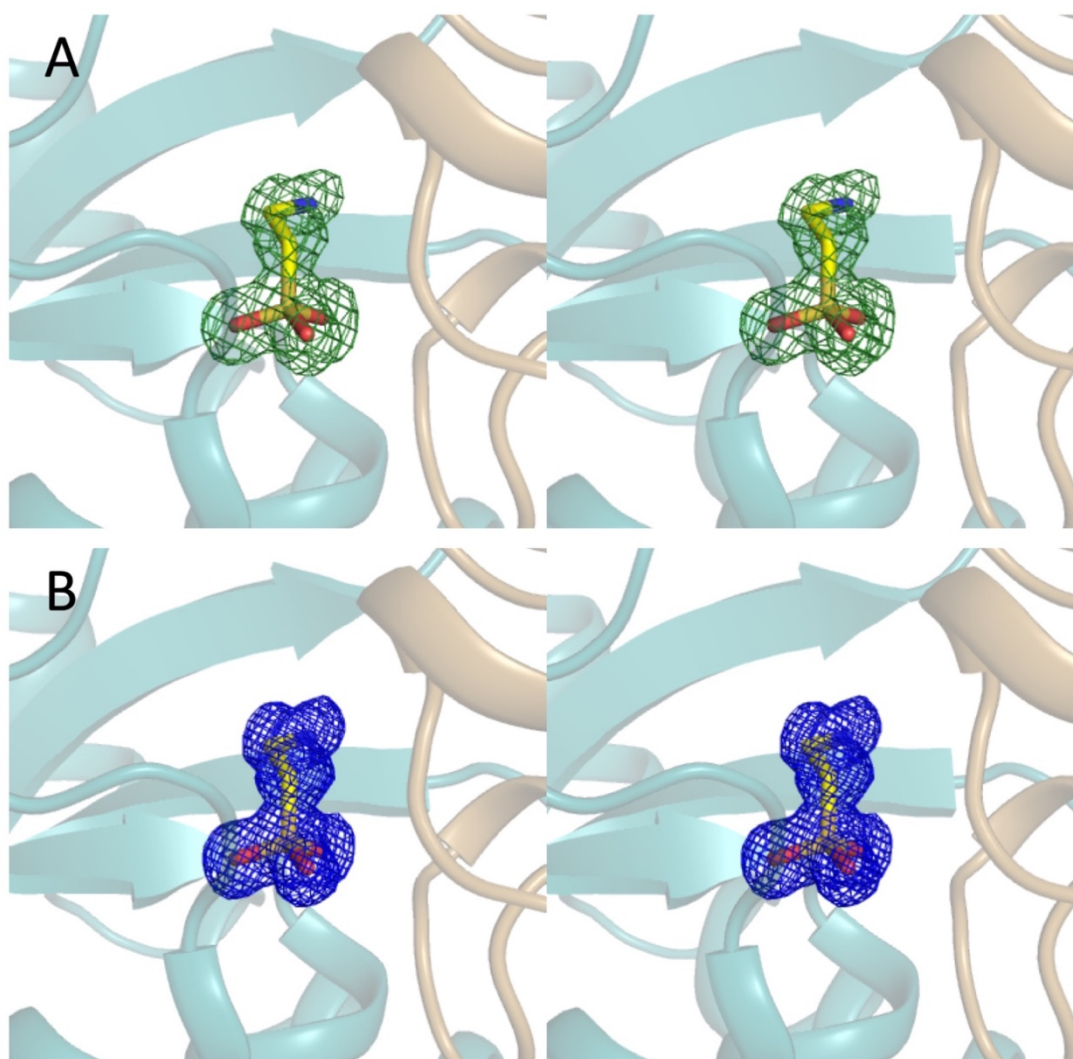


Figure 2.3.3.3. Stereo view of electron density maps for taurine. (A) The Fo-Fc omit maps of taurine contoured at 3σ (carve= $1.6 \text{ \AA}$) after omitting taurine from refinement (green mesh). (B) The 2Fo-Fc electron density maps of taurine contoured at 1σ (carve= $1.6 \text{ \AA}$) after refinement (blue mesh). Taurine is shown in stick format with carbon, sulfur, oxygen and nitrogen atoms coloured in yellow, gold, red and blue, respectively.

	TauA- taurine	TauA- taurine	TauA- 2-AEP	TauA- ACES	TauA- MES
Data collection					
Beamline	DLS-I23	DLS-I03	DLS-I24	DLS-I24	DLS-I24
Wavelength (Å)	2.7552	0.9763	0.9686	0.9686	0.9686
Resolution (Å)	163.66-1.77 (1.81-1.77)	67.53-1.31 (1.33-1.31)	65.68-1.62 (1.65-1.62)	35.9-1.5 (1.53-1.5)	66.59-1.50 (1.53-1.5)
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Cell dimensions					
a,b,c (Å)	44.98, 73.81, 163.66	44.96, 74.09, 164.12	45.1, 71.64, 164.55	45.16, 71.8, 164.78	45.16, 72.9, 163.65
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
NO. of total reflections	1133432 (17855)	417657 (20677)	433767 (19645)	330516 (14581)	407819 (18715)
NO. of unique reflections	46181(1840)	130542 (6414)	68655 (3375)	85811 (4180)	84688 (4118)
Completeness (%)	85.1 (59.6)	97.5 (96.9)	99.8 (99.9)	98.9 (97.8)	96.7 (95.3)
Redundancy	24.5 (9.7)	3.2 (3.2)	6.3 (5.8)	3.9 (3.5)	4.8 (4.5)
R _{meas}	0.113 (0.543)	0.068 (0.527)	0.109 (0.949)	0.084 (0.945)	0.132 (1.895)
Mean (I/sig(I))	20.4 (3.0)	11 (2.5)	10.9 (1.9)	11.9 (1.7)	7.4 (0.8)
Mn(I) half-set correlation CC 1/2	0.995 (0.889)	1 (0.8)	1 (0.6)	1 (0.5)	1 (0.3)
Anomalous completeness (%)	81.3 (58.5)				
Anomalous multiplicity (%)	12.3 (5.0)				
NO. of iodide sites	6	10	8	12	4
Refinement					
R _{work} /R _{free} (%)		13.1/17.6	13.3/20.0	14.6/21.8	15.7/22.0
Average B-factors (Å ²)		14.7	19.9	19.2	20.8
NO. of non-hydrogen atoms		5312	4934	5018	5018
RMS deviations					
Bond lengths (Å)		0.02	0.01	0.01	0.01
Bond angles (°)		1.97	1.79	1.81	1.82
Ramachandran					
Favoured (%)		96.58	96.6	96.45	96.1
Allowed (%)		3.08	3.06	3.05	3.4
Outliers (%)		0.34	0.34	0.51	0.51

Table 2.3.3.1. Data collection and refinement statistics for TauA crystals. The R-free value was calculated by excluding 5% of data from refinement. Values in parentheses are for the high-resolution shells.

TauA is composed of two α/β -domains connected by a two-loop hinge region (103-106 and 200-203) with the substrate-binding site located at the interface of the two domains (Figure 2.3.3.4). Domain I consists of residues 23-104 and 202-319, forming a central five-stranded β -sheet flanked by 11 α -helices. Domain II encompasses residues 105-201 and is composed of a five-stranded β -sheet core and four flanking α -helices. The central β -sheet has a topology of $\beta_2\beta_1\beta_3\beta_n\beta_4$, where $n=10$, representing the strand after the first crossover from domain II to domain I. TauA is a class II SBP belonging to cluster D (Berntsson *et al.*, 2010).

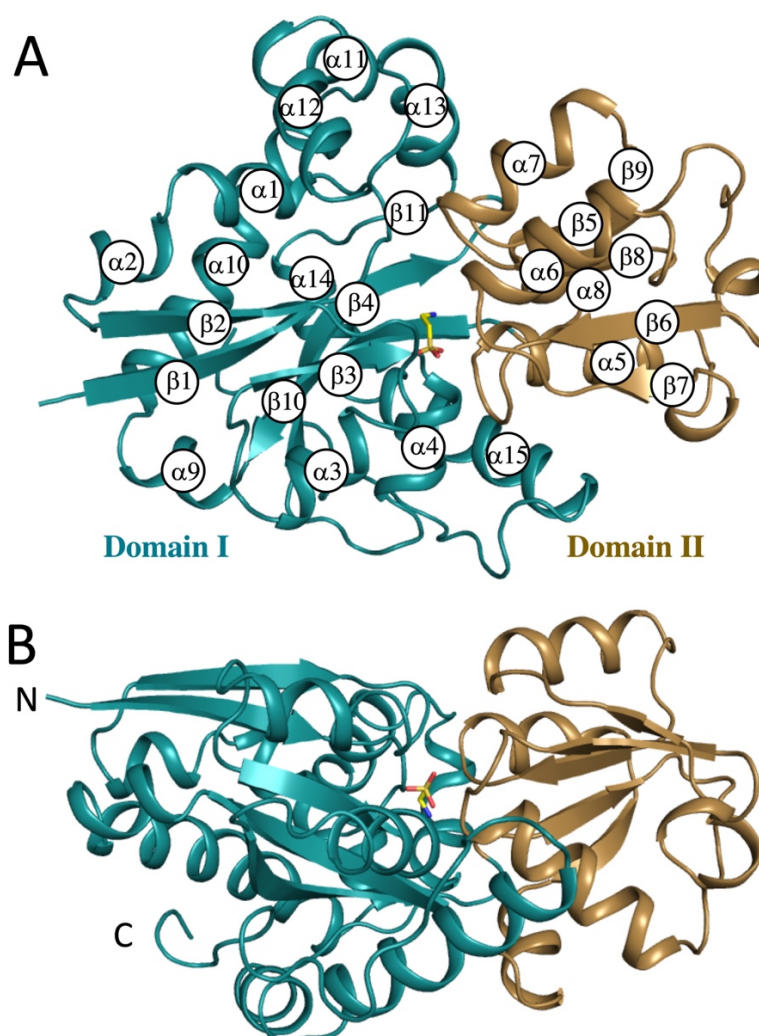


Figure 2.3.3.4. Structure of taurine-bound TauA. (A) Front-view and (B) side-view. Domain I and domain II are coloured in cyan and sand, respectively. The taurine molecule is shown in stick format with carbon in yellow, sulfur in gold, oxygen in red and nitrogen in blue. Secondary structure elements are indicated: α -helices (α) and β -strands (β).

The structure of taurine-bound TauA from *E. coli* was subjected to the DALI server (Holm & Sander, 1995). The Dali server is a network service for comparing the query protein structures against those in the Protein Data Bank (PDB). It aligns the 3D coordinates of different structures by comparing their intramolecular distances. Hits found in the DALI search are ranked in order of statistical significance. Similarities with a Z-score lower than 2 are spurious. The DALI search reveals that the closest structural homologue of taurine-bound TauA is the ligand-bound SsuA from *X. axonopodis pv. citri* (PDB ID: 3KSJ, 3KSX and 3E4R; Z-score $\approx$ 32%, RMSD=2 Å over $\sim$ 280 C α atoms) and apo SsuA from *E. coli* (PDB ID: 2X26; Z-score=28.1%, RMSD=3.1 Å over 277 C α atoms) (Figure 2.3.3.5) (de Araújo *et al.*, 2013; Beale *et al.*, 2010). RMSD (root-mean-square deviation) is the measure of the average distance between the C α atoms of superimposed proteins. SsuA is the alkanesulfonate-binding protein of the ABC importer SsuABC, involved in sulfur assimilation from a broad range of alkanesulfonates except for taurine, such as MES, HEPES and MOPS. *E. coli* TauA shares 22% (34%) and 23% (41%) sequence identity (similarity) with SsuA from *X. axonopodis pv. citri* and *E. coli*, respectively (Figure 2.3.3.6). Structural comparison shows that the taurine-bound TauA from *E. coli* and MES-bound SsuA from *X. axonopodis pv. citri* both exist in a closed conformation, regardless of the type of bound ligands, which is distinct from the apo open-state SsuA from *E. coli* (Figure 2.3.3.5). The distinct structures of ligand-bound and apo SBPs have validated the “Venus fly-trap” mechanism, where ligand binding stabilises the closed conformation (Figure 1.0.5.6) (Berntsson *et al.*, 2010).

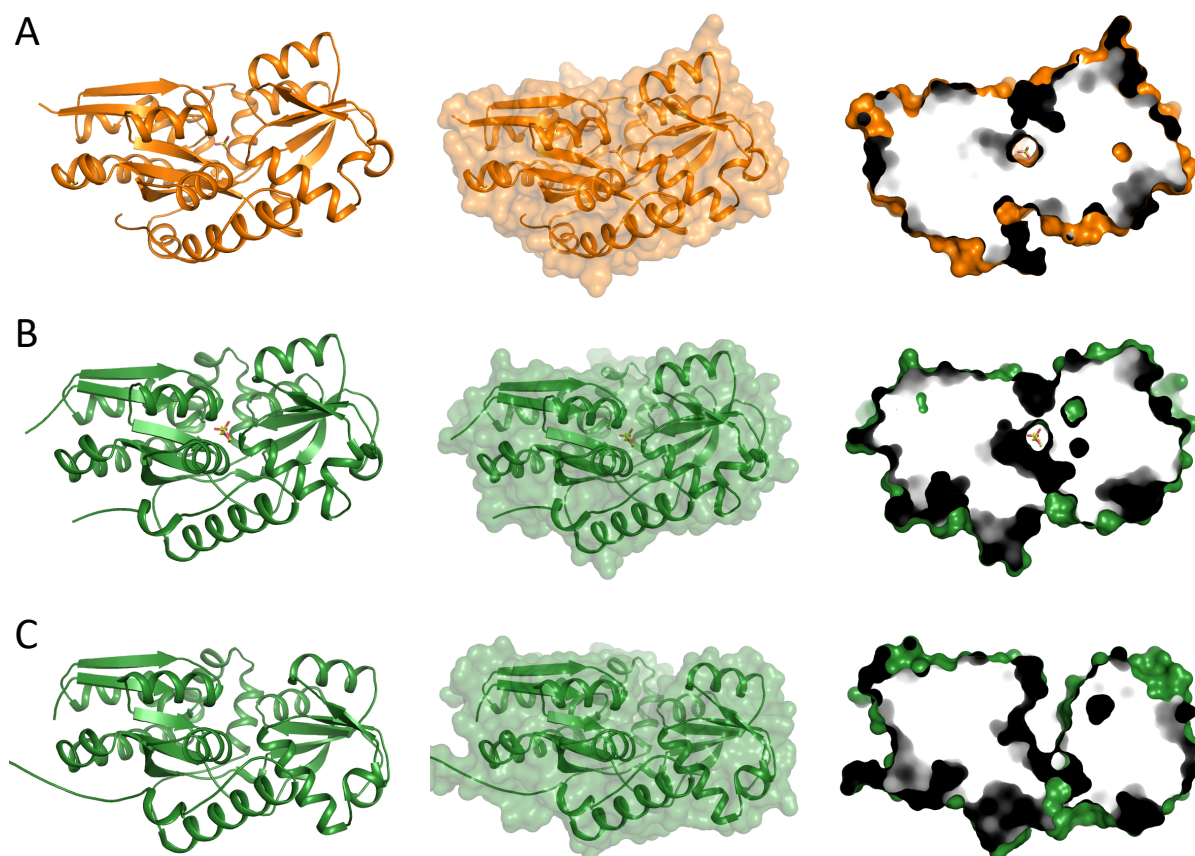


Figure 2.3.3.5. Structural comparison of TauA and SsuA. (A) Taurine-bound TauA from *E. coli* (orange). (B) MES-bound SsuA from *X. axonopodis* *pv. citri* (green) (PDB ID: 3KSJ, (de Araújo *et al.*, 2013)). (C) Apo SsuA from *E. coli* (green) (PDB ID: 2X26, (Beale *et al.*, 2010)). Left, middle and right panels show the ribbon, surface and slice surface representations, respectively. Carbon, sulfur, oxygen and nitrogen atoms in taurine and MES are coloured in yellow, gold, red and blue, respectively. The C-terminal disordered loop of apo SsuA from *E. coli* has been truncated for visualisation.

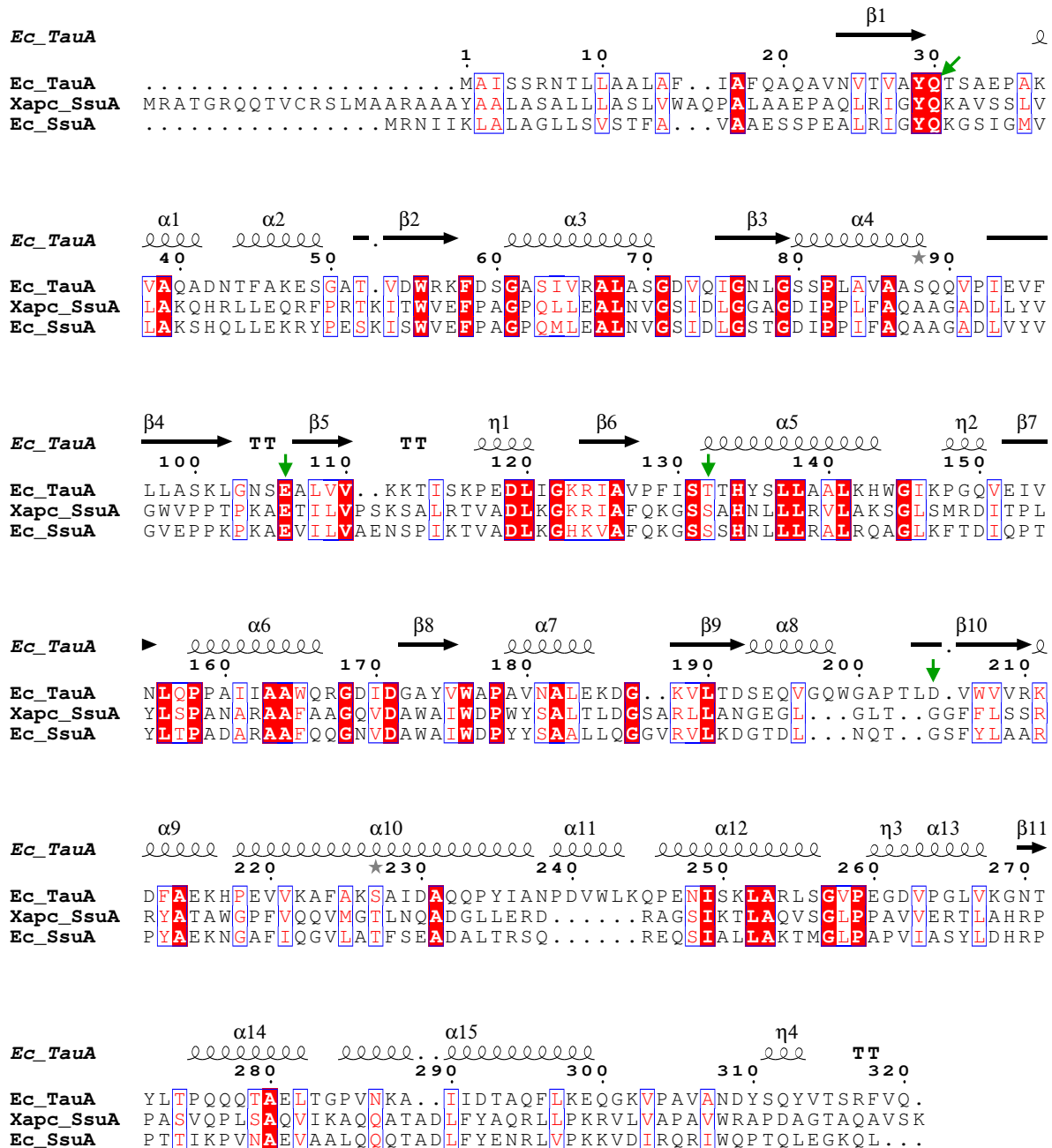


Figure 2.3.3.6. Sequence alignment of TauA and SsuA. TauA from *E. coli* (Ec_TauA), SsuA from *X. axonopodis* pv. *citri* (Xapc_SsuA) and SsuA from *E. coli* (Ec_SsuA). The conserved and similar residues are shown in red and blue boxes, respectively. Residue numbers and secondary structure elements of *E. coli* TauA are indicated on the top of the protein sequences. Residues of TauA that were mutated by site-directed mutagenesis are indicated by green arrows. Sequence alignment was performed by MultAlin (Corpet, 1988) and the figure was prepared using ESPript 3 (Robert & Gouet, 2014).

2.3.4. Characterisation of the TauA binding site

The crystal structure shows that TauA is composed of two α/β -domains with the taurine-binding site located at the domain interface (Figure 2.3.3.4). The taurine-binding pocket is composed of residues Gln30, Gly61, Gly79, Asp205 from domain I and Glu106, Ser131, Thr132 from domain II (Figure 2.3.4.1). Taurine stabilisation is mainly conferred through hydrogen bonds and van der Waals contact with residues from TauA. Taurine consists of a sulfonate group ($-\text{SO}_3$) and an ethylamine group ($-\text{H}_2\text{CH}_2\text{NH}_2$). The three sulfonate oxygens (designated O_1 , O_2 and O_3) form hydrogen bonds with the backbone amide groups of Gly61, Gly79 and Thr132. The sulfonate oxygens O_1 and O_3 also form hydrogen bonds with the side chains of Gln30 and Thr132, respectively. The sulfonate oxygen O_2 is coordinated by a water molecule through a hydrogen bond. The amino group of taurine ($-\text{NH}_2$) is hydrogen bonded to the carbonyl oxygens of Glu106, Thr132, Asp205 and a water molecule. The sulfonate oxygens and ethylamine carbons form van der Waals contact with residues Asp205, Ser131, Glu30, Gly61 and Gly79.

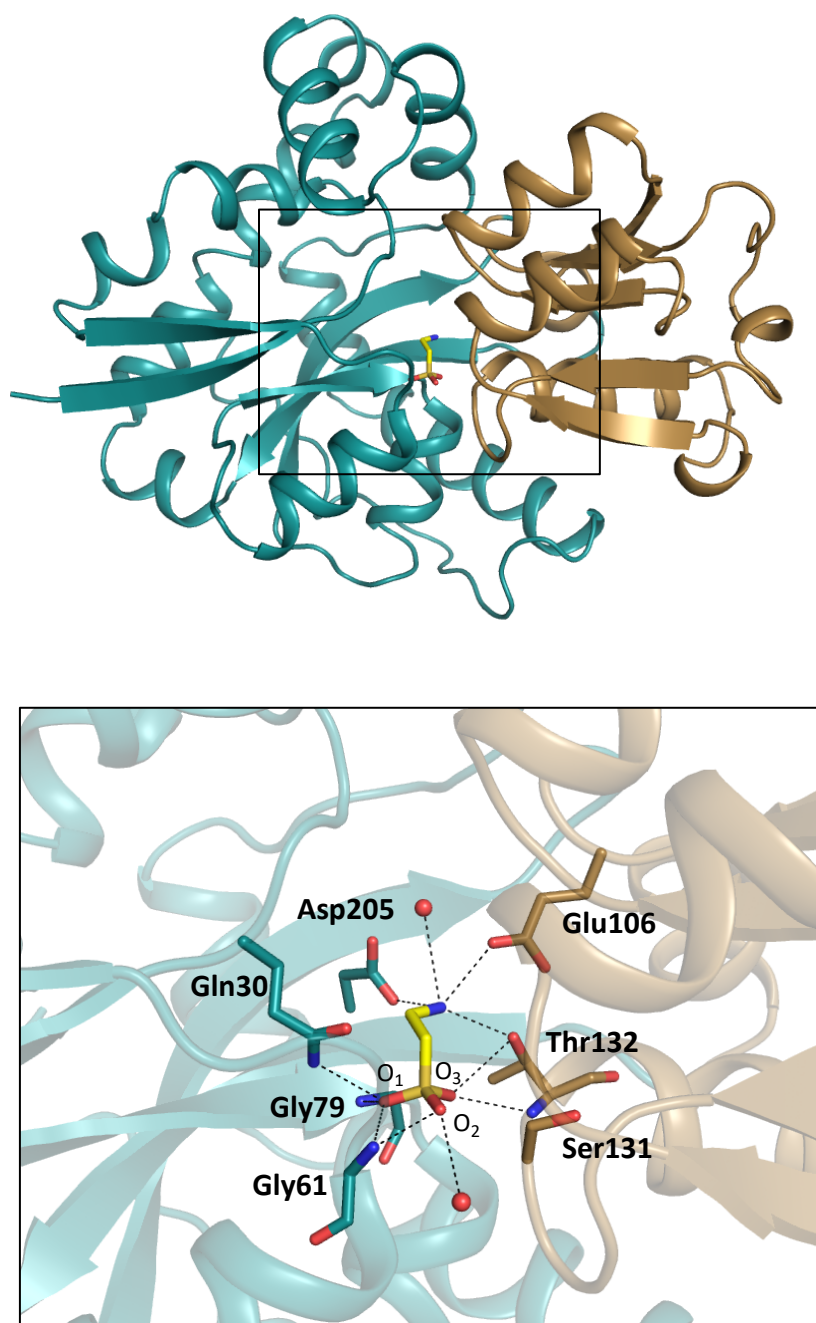


Figure 2.3.4.1. The taurine-binding site of TauA. Ribbon representation of the structure of taurine-bound TauA from *E. coli* (domains I and II are coloured in cyan and sand, respectively). The inset shows the detailed taurine-binding site and hydrogen bond interactions between TauA and taurine. Taurine and taurine-binding residues of TauA are shown in stick format. Carbon atoms of taurine, domain I and domain II residues of TauA are coloured in yellow, cyan and sand, respectively. Sulfur, oxygen and nitrogen atoms of taurine are coloured in gold, red and blue, respectively. Water molecules are shown as red spheres. Hydrogen bonds between TauA and taurine are presented as black dotted lines.

In order to validate roles of specific residues in taurine binding, alanine mutants at the binding site of TauA were designed and purified. Four residues, Gln30, Glu106, Thr132 and Asp205 of TauA, were substituted individually to alanine by site-directed mutagenesis. Binding affinities of taurine to wild-type TauA and mutants were determined by ITC. Wild-type TauA binds to taurine with a K_d of 1.61 nM (Table 2.3.4.1 and Figure 2.3.4.2), in agreement with the previously reported K_d value of 2.5 nM (Javaux *et al.*, 2007). The large number of interactions between TauA and taurine probably explains the tight binding of taurine (Figure 2.3.4.1). The binding affinity dropped to $\sim 3 \mu\text{M}$ upon loss of the hydrogen bond between Gln30 from TauA and the sulfonate group of taurine (Table 2.3.4.1 and Figure 2.3.4.2). Two hydrogen bonds were lost upon the Thr132 mutation, which resulted in a reduction of binding affinity to $\sim 21 \mu\text{M}$. Loss of the hydrogen bond between Asp205 and the amino group of taurine reduced the binding affinity to $\sim 23 \mu\text{M}$. The ITC data suggest the importance these three residues in the taurine binding. More significantly, complete loss of taurine binding was observed upon mutation of Glu106, suggesting that Glu106 is a key residue in taurine binding. Glu106 is located in domain II of TauA (Figure 2.3.4.1). It is likely that the E106A mutation has abolished the hydrogen bond between Glu106 and the amino group of taurine, thereby preventing the “Venus fly-trap” closure of TauA (Figure 1.0.5.6).

	TauA	Q30A	E106A	T132A	D205A
K_d (nM)	1.61	3480	-	20620	22880
stoichiometry	0.77 ($\pm 0.38\%$)	0.83 ($\pm 6\%$)	-	0.80 *	0.80 *
ΔH (kcal/mol)	-14.58 ($\pm 0.08\%$)	-4.27 ($\pm 8\%$)	-	-6.87 ($\pm 8\%$)	-8.32 ($\pm 2\%$)
ΔS (cal/mol/ $^{\circ}\text{C}$)	-8.69	10.60	-	-1.61	-6.65

Table 2.3.4.1. Thermodynamic parameters for taurine binding to TauA and mutants. NB: no binding. *: stoichiometry fixed to 0.80 during data fitting.

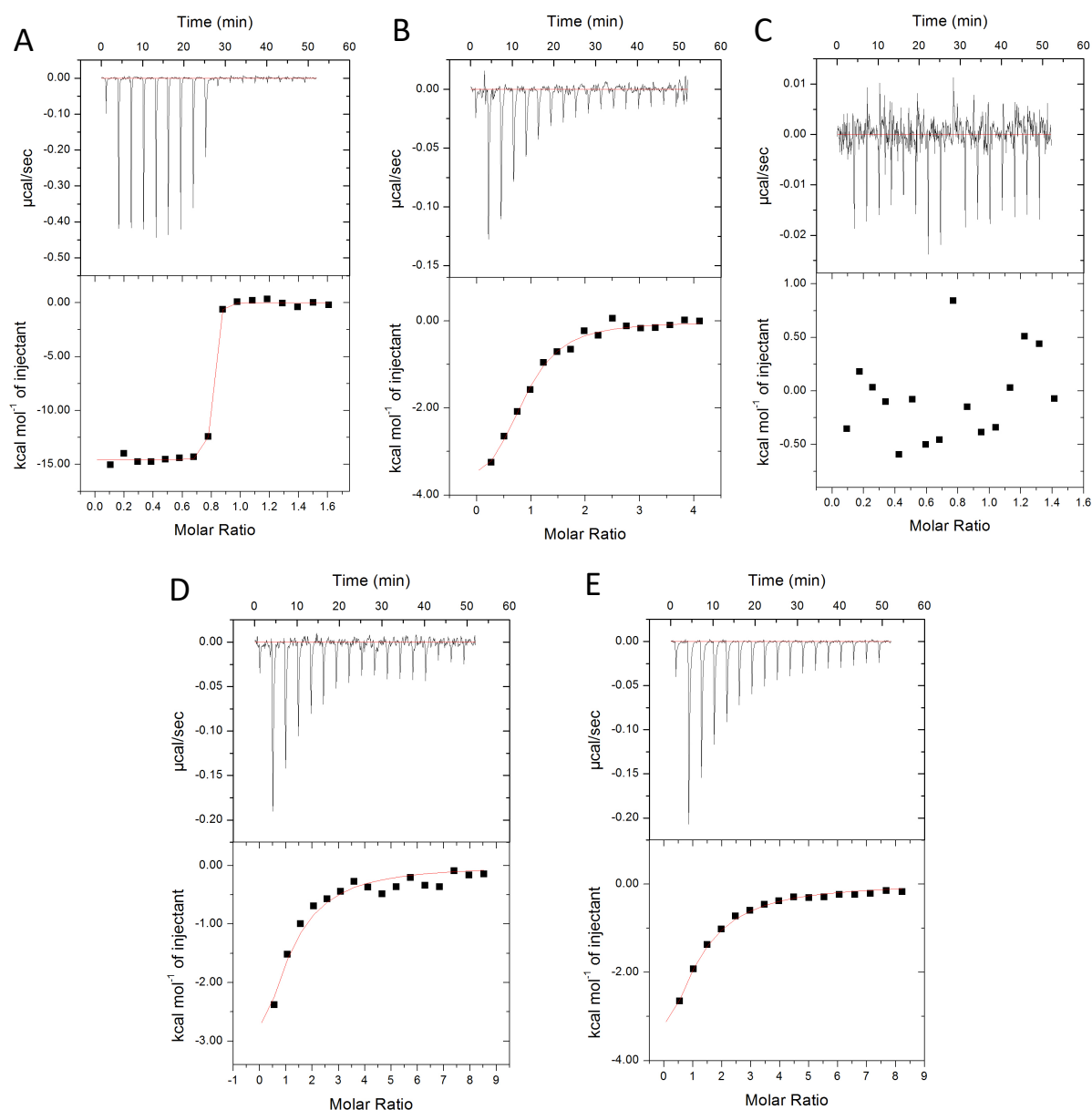


Figure 2.3.4.2. Calorimetric titration for taurine binding to TauA and mutants. Binding of taurine to wild-type TauA (A) and TauA mutants: Gln30Ala (B), Glu106Ala (C), Thr132Ala (D) and Asp206Ala (E). Each peak (top panel) represents an injection of 2.4 μL of taurine into 200 μL of wild-type TauA or TauA mutant. The bottom panel shows the integrated heat obtained from the raw data, after subtracting the heat associated with dilution of the ligand. Binding of taurine to TauA is exothermic with a binding affinity of 1.61 nM. All the TauA mutants show a decrease in taurine-binding affinity. Specifically, the Glu106Ala mutant totally loses the binding affinity for taurine.

2.3.5. TauA substrate specificity

Previous studies have shown that SsuA mediates uptake of various alkanesulfonates, whereas TauA shows higher specificity to taurine (Ploeg *et al.*, 1996; Eichhorn *et al.*, 1999). Binding of TauA for structural analogues of taurine (Figure 2.3.5.1) was investigated by ITC in order to understand the substrate specificity of TauA (Table 2.3.5.1 and Figure 2.3.5.2). 2-AEP was selected as a taurine analogue since it contains a phosphonate group instead of a sulfonate group (Figure 2.3.5.1B). Although 2-AEP and taurine share the same structural configuration, binding affinity between TauA and 2-AEP was measured to be 649351 nM, with a ~403324-fold reduction compared to binding of TauA for taurine. This suggests that although the phosphonate group has the same tetrahedral structure as the sulfonate group, it cannot be coordinated in the taurine-binding site and the ethylamine group alone is not sufficient to stabilise 2-AEP in the binding site of TauA. The loss of binding is likely due to the different bond length of P-O from the phosphonate group to S-O from the sulfonate group, which results in selectivity of TauA for taurine over 2-AEP. We also investigated whether TauA binds larger analogues of taurine like ACES (Figure 2.3.5.1C) and MES (Figure 2.3.5.1D). ACES and MES contain an acetamido and a morpholino group, respectively, after the ethylamine group. ITC data showed an affinity of 7463 nM for ACES binding to TauA, which was a ~4635-fold decrease compared to taurine. MES binds to TauA with a binding affinity of 2849 nM, which reduced by ~1770-fold. These data suggest that the substrate-binding site of TauA can accommodate the larger compounds ACES and MES. The significant decrease of binding affinity between TauA and taurine analogues has confirmed the high selectivity of TauA to taurine and suggested that the sulfonate group and ethylamine group are both important for taurine recognition. Enthalpy and entropy define the level of energy and disorder of a system, respectively. The ITC data suggest that reactions between TauA and taurine/ACES are

exothermic ($\Delta H < 0$), whereas binding of 2-AEP and MES to TauA are endothermic ($\Delta H > 0$). The amount of heat released upon taurine binding is ~42-fold higher than recorded for ACES binding, in agreement with the much higher affinity of TauA for taurine than ACES (~4635-fold). The amount of heat absorbed for binding of 2-AEP is ~3.5-fold higher than that of MES, suggesting that binding of MES is more favourable than 2-AEP. Additionally, the system becomes less disordered upon binding of taurine ($\Delta S < 0$), whereas binding of 2-AEP, ACES and MES increases the level of disorder ($\Delta S > 0$). This suggests that taurine binding is a more favourable reaction compared to binding of 2-AEP, ACES and MES.

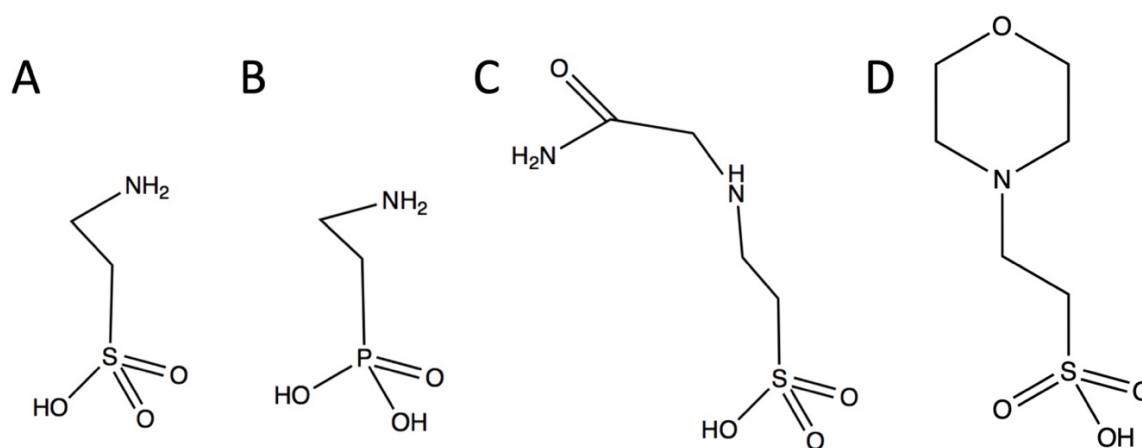


Figure 2.3.5.1. Taurine and its analogues. (A) taurine, (B) 2-AEP, (C) ACES, (D) MES.

	K_d (nM)	stoichiometry	ΔH (cal/mol)	ΔS (cal/mol/°C)
taurine	1.61	0.77 ($\pm 0.38\%$)	-14580 ($\pm 0.08\%$)	-8.7
2-AEP	649351	0.80 *	3952 ($\pm 0.03\%$)	27.8
ACES	7463	0.79 ($\pm 2.4\%$)	-345 ($\pm 3.3\%$)	22.3
MES	2849	0.87 ($\pm 0.5\%$)	1134 ($\pm 0.8\%$)	29.2

Table 2.3.5.1. Thermodynamic parameters for taurine and analogues binding to TauA. *: stoichiometry fixed to 0.8 during data fitting.

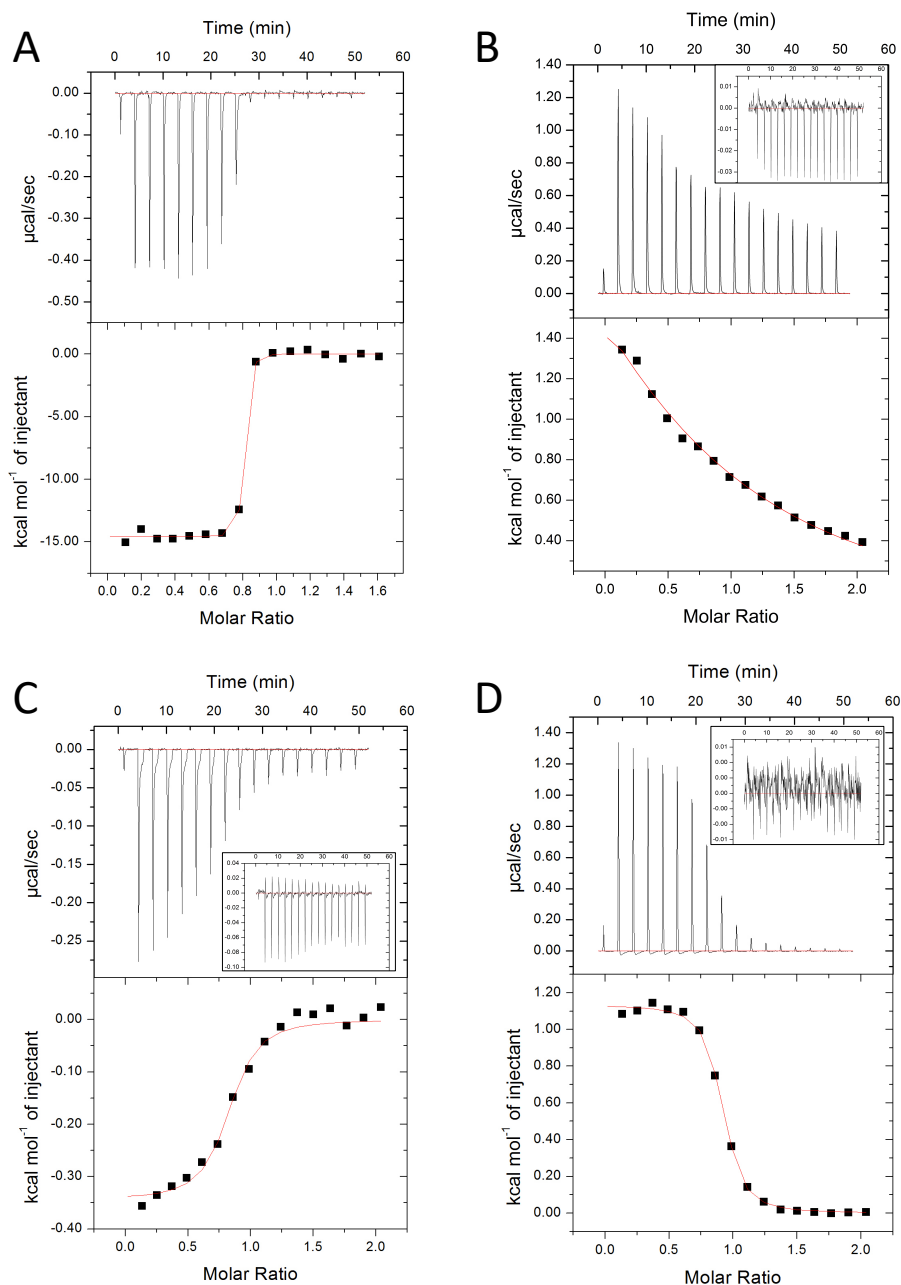


Figure 2.3.5.2. Calorimetric titration for the binding of different ligands to TauA. Each peak (top panel) represents an injection of 2.4 μL of taurine (A), 2-AEP (B), ACES (C) and MES (D) into 200 μL of TauA. The bottom panel shows the integrated heat obtained from the raw data, after subtracting the heat associated with dilution of the ligand (shown in insets). Results indicate that TauA has a high binding affinity for taurine ($K_d=1.61$ nM). 2-AEP, ACES and MES bind to TauA with affinities of ~ 649 μM , ~ 7 μM and ~ 3 μM , respectively. Binding of taurine and ACES to TauA is exothermic whilst binding of 2-AEP and MES is endothermic.

In order to better understand the structural basis of ligand binding, TauA was crystallised with the three taurine analogues and X-ray data were collected. The structures of TauA in complex with 2-AEP, ACES and MES were refined at a resolution of 1.62 Å, 1.5 Å and 1.5 Å, respectively (Table 2.3.3.1). All crystals belonged to the space group $P2_12_12_1$ with two copies of TauA in the asymmetric unit. The electron density maps unambiguously define the orientation of 2-AEP, ACES and MES in the binding site of TauA (Figure 2.3.5.3).

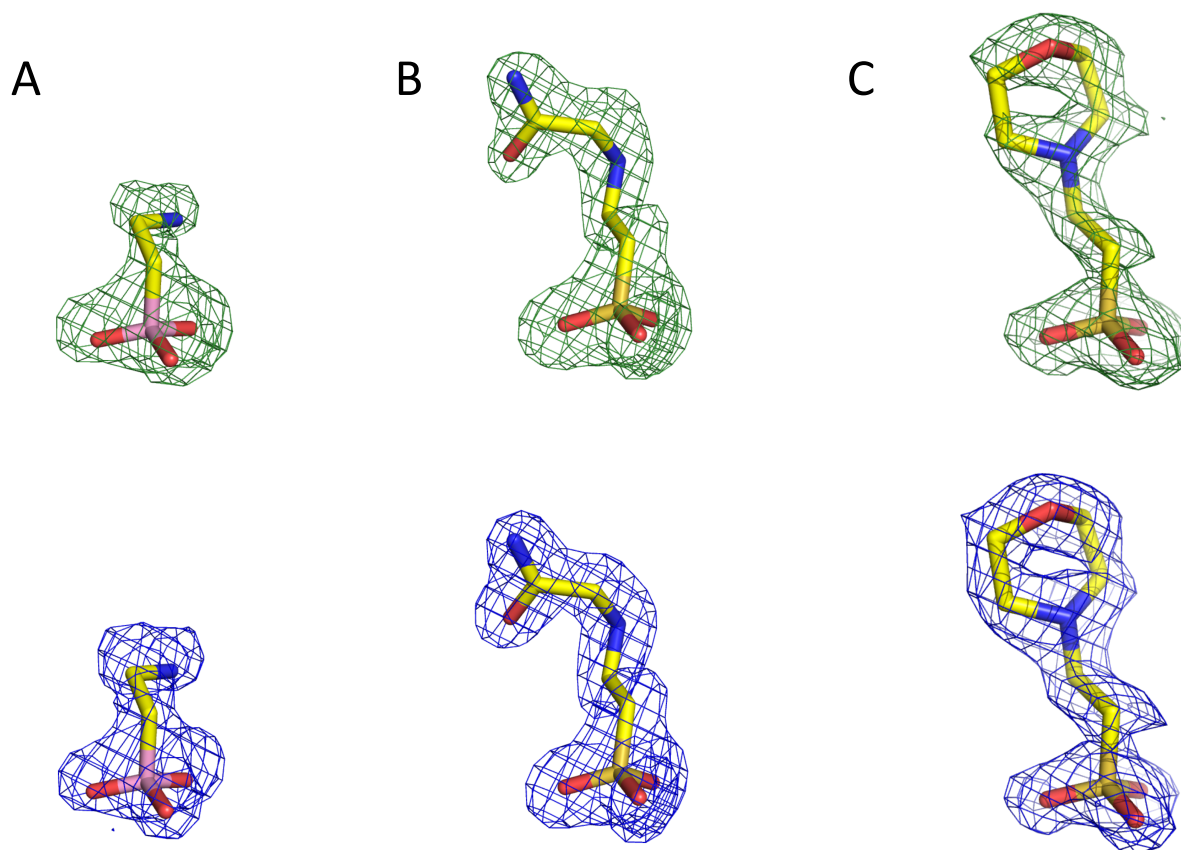


Figure 2.3.5.3. Electron density maps for 2-AEP, ACES and MES. (A) 2-AEP, (B) ACES and (C) MES. Top panel: the Fo-Fc omit maps of the ligands contoured at 3σ (carve=1.6 Å) after omission of the ligand from refinement (green mesh). Bottom panel: the 2Fo-Fc electron density maps of the ligands contoured at 1σ (carve=1.6 Å) after refinement (blue mesh). The ligand is shown in stick format with carbon, sulfur, phosphorus, oxygen and nitrogen atoms coloured in yellow, gold, pink, red and blue, respectively.

The overall structures of TauA in complex with taurine and taurine analogues appear identical, regardless of chemical structures of the bound ligands. This observation has confirmed that taurine and its analogues can induce conformational changes in TauA, allowing them to bind to the substrate binding site of TauA. Coordination of 2-AEP, ACES and MES in the binding pocket of TauA involves a similar set of residues as those for taurine (Figure 2.3.5.4). Coordination of the sulfonate/phosphonate group always involves a water molecule. Binding of 2-AEP involves in an additional hydrogen bond between the phosphonate group of 2-AEP and the carbonyl oxygen of Ser131 (Figure 2.3.5.4B). This hydrogen bond is absent in the TauA-taurine structure and is formed owing to the re-orientation of Ser131 side chain towards 2-AEP. The local conformational changes at Ser131 may allow TauA to distinguish the phosphonate group of 2-AEP for ligand binding. The acetamido group of ACES forms additional hydrogen bonds with the main chain of Ser32, the side chain of Asn77 and a water molecule (Figure 2.3.5.4C). In the TauA-MES complex, the morpholino group of MES is hydrogen bonded to Asp205 and is further stabilised by Glu106 through van der Waals contact (Figure 2.3.5.4D). The water molecule observed in the taurine, 2-AEP and ACES complex is displaced by the morpholino group of MES. Additionally, local conformation changes are observed in the TauA-MES structure, where Asp205 rotates 90° and Glu106 moves away from the ligand.

As mentioned earlier, binding of taurine and ACES is exothermic, whilst binding of 2-AEP and MES is endothermic (Table 2.3.5.1 and Figure 2.3.5.2). In the presence of 2-AEP and MES, Ser132, Glu106 and Asp205 adopt different rotamers to coordinate either the phosphonate group of 2-AEP or the voluminous morpholino group of MES (Figure 2.3.5.4). We propose that energy addition is probably necessary for local conformational changes in the 2-AEP/MES-binding site of TauA in order to overcome the energy barriers, resulting in endothermic binding reactions.

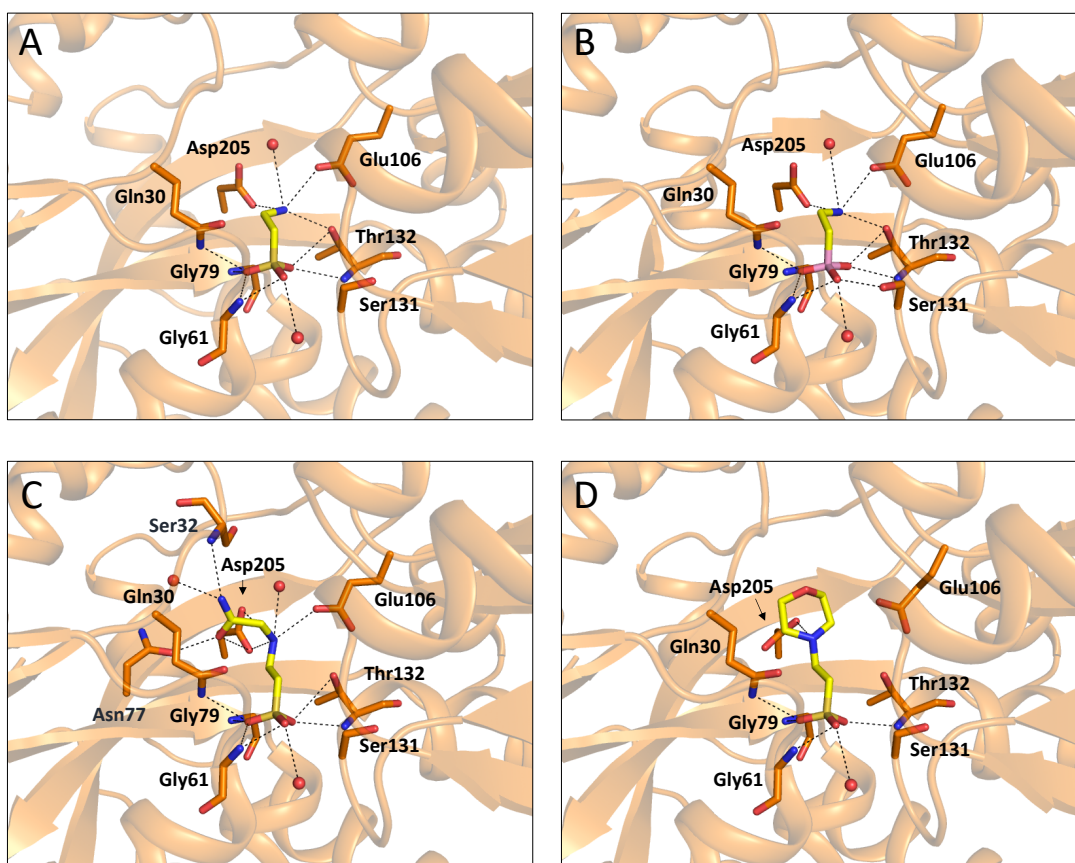


Figure 2.3.5.4. TauA in complex with taurine and its analogues. Hydrogen bonds between TauA and Taurine (A), 2-AEP (B), ACES (C) and MES (D) are presented as black dotted lines. In the TauA-2-AEP structure, Ser131 re-orientates to coordinate the phosphonate group of 2-AEP. Binding of ACES involves Ser32, Asn77 that are hydrogen bonded to the acetamido group of ACES. The morpholino group of MES is hydrogen bonded to Asp205 and is further stabilised by Glu106 through van der Waals contact. Ligands and ligand-binding residues of TauA are shown in stick format (yellow and orange, respectively). Sulfur, phosphorus, oxygen and nitrogen atoms are coloured in gold, pink, red and blue, respectively. Water molecules are shown as red spheres.

Although additional interactions were formed in the presence of 2-AEP, ACES and MES, their binding affinities have decreased drastically compared to taurine (Table 2.3.5.1 and Figure 2.3.5.2). Furthermore, water distributions in the binding site of TauA are different upon binding of different ligands (Figure 2.3.5.4). The acetamido group of ACES is coordinated by an additional water molecule. The morpholino group of MES has replaced the water molecule that

is present in the taurine, 2-AEP and ACES complexes. We therefore propose that desolvation of the binding site of TauA might contribute to the lower binding affinities of TauA for 2-AEP, ACES and MES over taurine, although they display more hydrogen bonds. In collaboration with Dr. Alfonso de Simone (Imperial College London), molecular dynamics (MD) simulations were performed based on structures of TauA in complex with the four ligands (Table 2.3.5.2). Apo TauA was created by separating taurine from the TauA-aurine structure. Water molecules were allowed to freely move and occupy the binding site of TauA. For each complex, an average potential energy per water molecule and an overall energy difference to the apo form were calculated. The simulation data show that three water molecules have been replaced in the presence of taurine, while 10 water molecules are substituted upon binding of 2-AEP, ACES and MES. Upon binding of different ligands, the average potential energy per water remains similar. However, the overall energy difference between the TauA-aurine complex and apo TauA is ~four times lower compared to TauA in complex with 2-AEP, ACES and MES. We conclude that desolvation of the binding site of TauA is detrimental to ligand binding, which is reflected by the lower binding affinity of TauA for 2-AEP, ACES and MES over taurine.

System	Average number of waters	Average potential energy per water (kJ/mol)	Overall difference with apo form (kJ/mol)
Apo protein	20.2	-93.4	0
Taurine	17.5	-94.4	+234.7
2-AEP	10.5	-87.7	+965.83
ACES	10.8	-86.1	+956.8
MES	10.2	-88.2	+987.0
Bulk Water		-92.2	

Table 2.3.5.2. MD simulation of water distributions in the binding site of TauA.

2.3.6. Comparison with SsuA

Structural comparison of the substrate-binding sites of TauA and SsuA reveals conserved structural elements in ligand binding and the structural basis for substrate selectivity of TauA and SsuA. Comparison between taurine-bound TauA, MES-bound TauA and MES-bound SsuA will be discussed in detail here.

In all three structures, the sulfonate group of taurine and MES contacts the proteins via a cluster of conserved residues including glutamine, glycine, serine and threonine(TauA)/serine(SsuA) (Figure 2.3.6.1). In the TauA-aurine structure, the amino group of taurine is stabilised by the carbonyl oxygens of Glu106, Thr132 and Asp205 from TauA and a water molecule via hydrogen bonds (Figure 2.3.6.1A), whereas the morpholino group of MES mainly contacts the proteins via van der Waals forces (Figure 2.3.6.1B and C). Variation in binding pocket volumes might dictate the substrate selectivity between TauA and SsuA. Asp205 constricts the binding pocket in the TauA-aurine complex (Figure 2.3.6.1A), whereas its equivalent residue in SsuA is Gly187, which abolishes a hydrogen bond (Figure 2.3.6.1C). This observation might explain why the Ssu system cannot recognise and transport taurine (Eichhorn *et al.*, 2000). Glu106 also defines the binding pocket of TauA (Figure 2.3.6.1A), whist its equivalent residue Glu88 in SsuA moves 1.8 Å away from the morpholino group of MES (measured at the carboxyl oxygen) (Figure 2.3.6.1C). Movement of Glu88 results in a 4.3-Å distance between the carbonyl oxygen of the glutamate and the amino group of taurine, which abolishes a hydrogen bond. This observation is in agreement with the mutagenesis and ITC data showing the complete loss of taurine binding upon alanine substitution of Glu106 of TauA (Table 2.3.4.1 and Figure 2.3.4.2). Glu106 in TauA-aurine would abolish the van der Waals contact with the morpholino group of MES in the SsuA-MES complex. The ligand-binding site of TauA also undergoes local conformational changes to coordinate the MES molecule. In the MES complex, Asp205 rotates

90° and Glu106 moves away from the ligand (0.9 Å measured at the carboxyl oxygen) (Figure 2.3.6.1B). Additionally, the MES molecule in the TauA-MES structure tilts ~10°, thus maintaining the van der Waals contact with Glu106 (Figure 2.3.6.1, B and C).

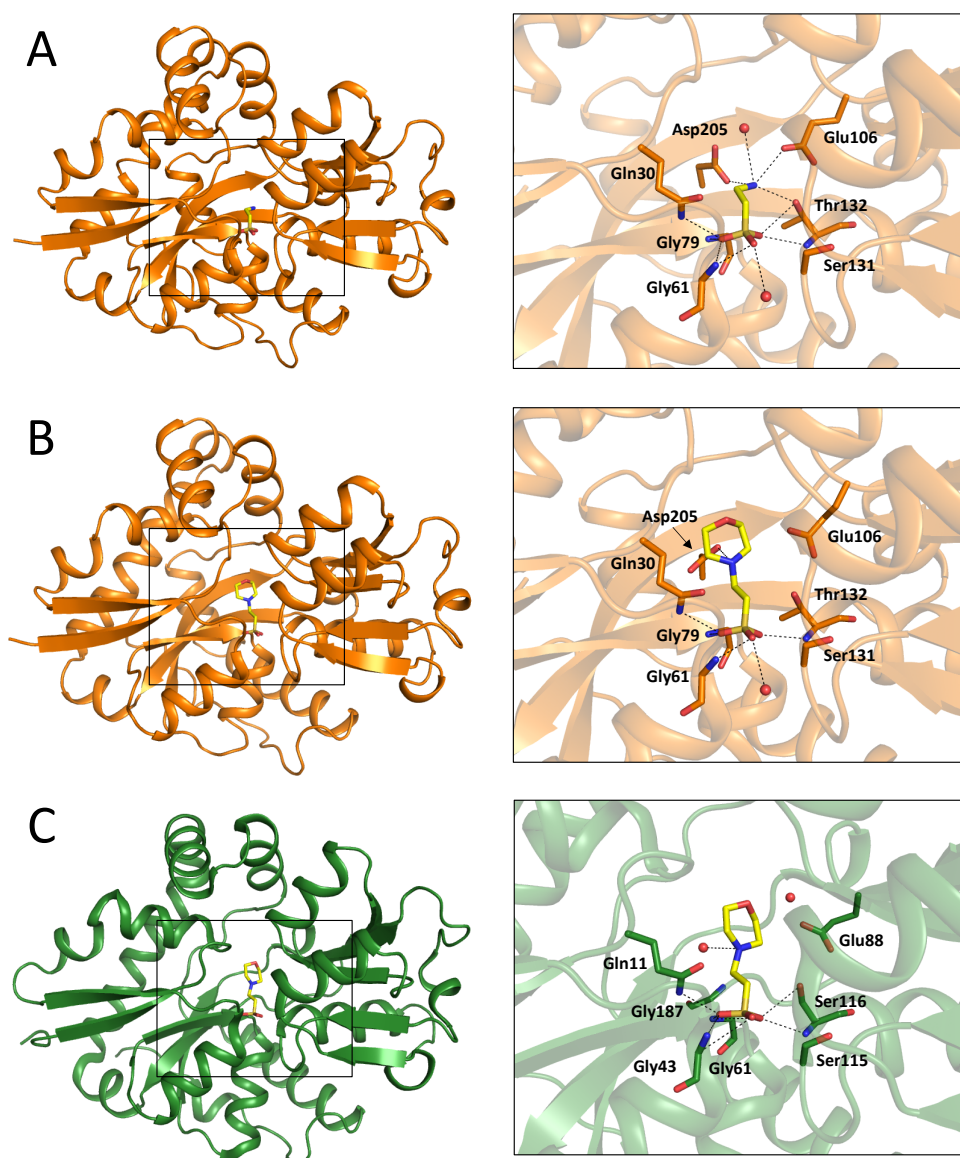


Figure 2.3.6.1. Structural comparison between TauA and SsuA. (A) TauA-*taurine*, (B) TauA-MES and (C) SsuA-MES (PDB: 3KSJ, de Araújo *et al.*, 2013). The insets show the detailed substrate-binding sites and hydrogen bonds (black dotted lines). Due to residue movements, SsuA has a larger substrate-binding pocket. In the presence of MES, Asp205 and Glu106 move away from the ligands, resulting in an expanded binding pocket of TauA. The MES molecule in TauA-MES tilts ~10° compared to that of SsuA-MES. Ligand and residues of TauA are shown in stick format. Carbon atoms of ligands and those of TauA and SsuA are coloured in yellow, orange and green, respectively. Sulfur, oxygen and nitrogen atoms are coloured in gold, red and blue, respectively. Water molecules are shown as red spheres.

2.3.7. Purification of TauBC

In order to better characterise the full-length ABC importer TauABC, the TauBC transporter was cloned and purified. TauBC was overexpressed and purified using the His₆-TEV-TauBC construct (Figure 2.2.8.1). A prerequisite for successful TauBC purification using this construct is that the soluble nucleotide-binding protein TauB and the inner-membrane protein TauC form a full ABC transporter.

TauBC was isolated from total *E. coli* membranes and solubilised in DDM. The SEC profile of TauBC run in a buffer containing DDM, showed a sharp monodisperse peak eluted at 12.8 mL (Figure 2.3.7.1A), which indicates the efficient recovery of a single TauBC species. There was no aggregation present on the gel filtration profile. The theoretical MWs of TauB and TauC are 28.297 and 29.812 kDa, respectively, making an ABC transporter of 58.1 kDa in MW. By comparison, the ABC exporter McjD from *E. coli* (65.425 kDa) eluted from the Superdex 200 10/300 GL column at ~12 ml as a homodimer (Choudhury *et al.*, 2014), indicating that the TauBC protein has also been purified as a dimer. A considerable amount of pure TauBC protein was obtained (15 mg/mL at 100 μ L). Two large bands were observed by Coomassie-stained SDS-PAGE (Figure 2.3.7.1B). The higher and lower bands corresponded to purified TauB and TauC. The protein band for TauC was lower than its theoretical MW owing to the masking effect of SDS on the hydrophobic region of the protein. SDS is an anionic surfactant that contains both hydrophobic and hydrophilic groups (Rath *et al.*, 2009). In SDS-PAGE, SDS disrupts the tertiary structure of proteins and coats the proteins with a uniform negative charge. Due to the presence of the hydrophobic domains, membrane proteins have higher loading capacity for SDS compared to globular soluble polypeptides, thereby carrying more negative charge. Consequently, membrane proteins generally run faster in SDS-PAGE.

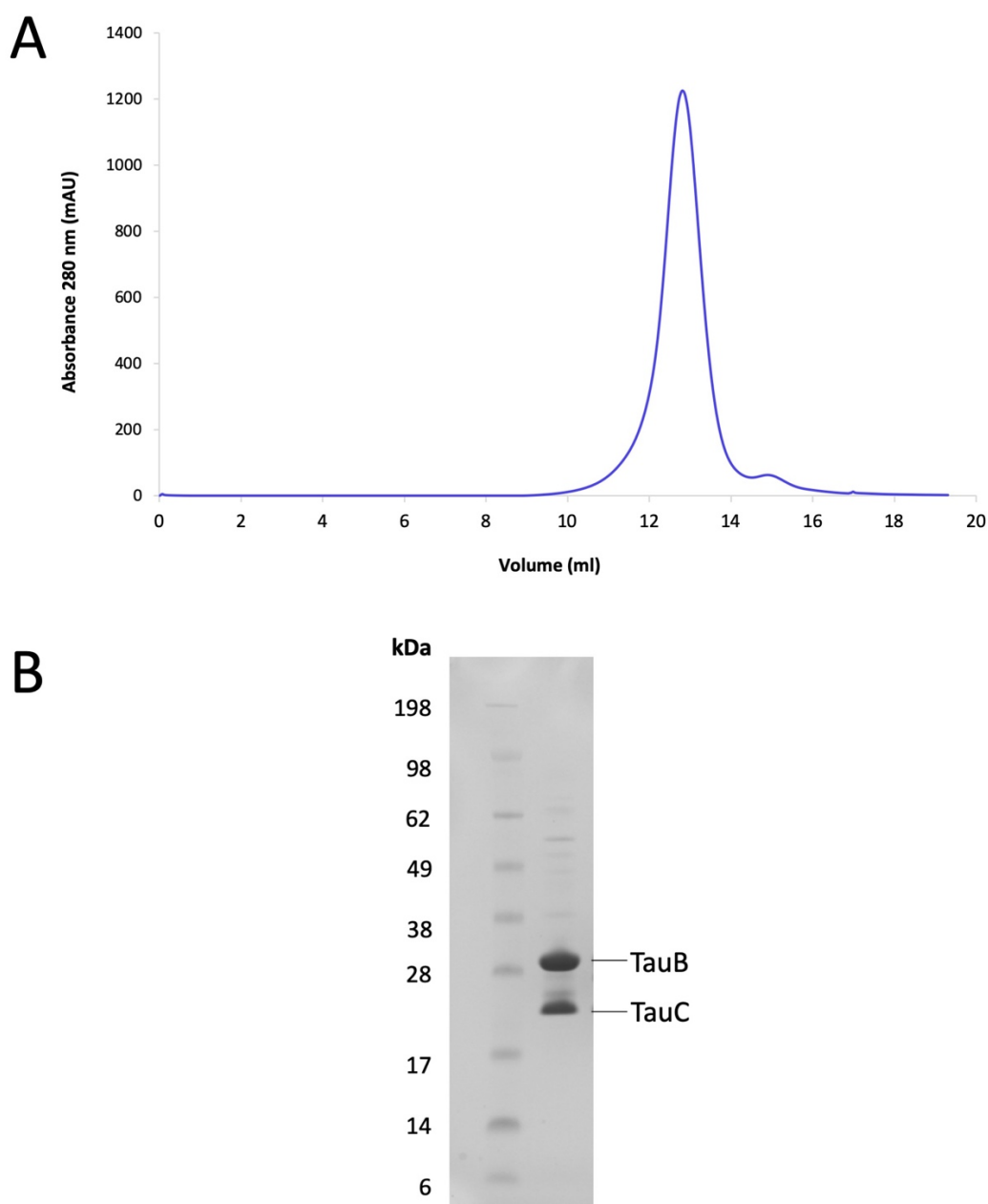


Figure 2.3.7.1. Purification of TauBC. (A) SEC profile for TauBC using the Superdex 200 10/300 GL column. TauBC eluted at 12.8 mL from the column, indicating that TauBC has been purified as a dimer. (B) SDS-PAGE shows two bands at ~28 kDa and ~24 kDa corresponding to TauB and TauC, respectively. The band for TauC is lower than its theoretical MW owing to the masking effect of SDS on the hydrophobic region.

2.3.8. Functional characterisation of TauBC

2.3.8.1. ATPase activity in detergent

Initially, the ATPase activity of TauBC was measured in detergent-protein micelles (Figure 2.3.8.1A). The purified TauBC protein was used for ATPase assays directly after concentration. In the absence of TauA/taurine, ATP was hydrolysed at a rate of 56.7 nmoles/min/mg (basal activity) by TauBC (1 μ M) (Figure 2.3.8.1B). The basal ATPase activity of TauBC shows that the purified TauBC protein was active in detergent, indicating that functional TauBC has been produced using DDM as a detergent. However, addition of TauA (10 μ M) and taurine (100 μ M) did not stimulate the ATPase activity of TauBC (1 μ M) (V_{max} =55.6 nmoles/min/mg). The detergent micelle is highly different from the native membrane environment. The TauBC protein was therefore reconstituted into proteoliposomes in order to provide a more physiologically relevant environment for ATPase assays.

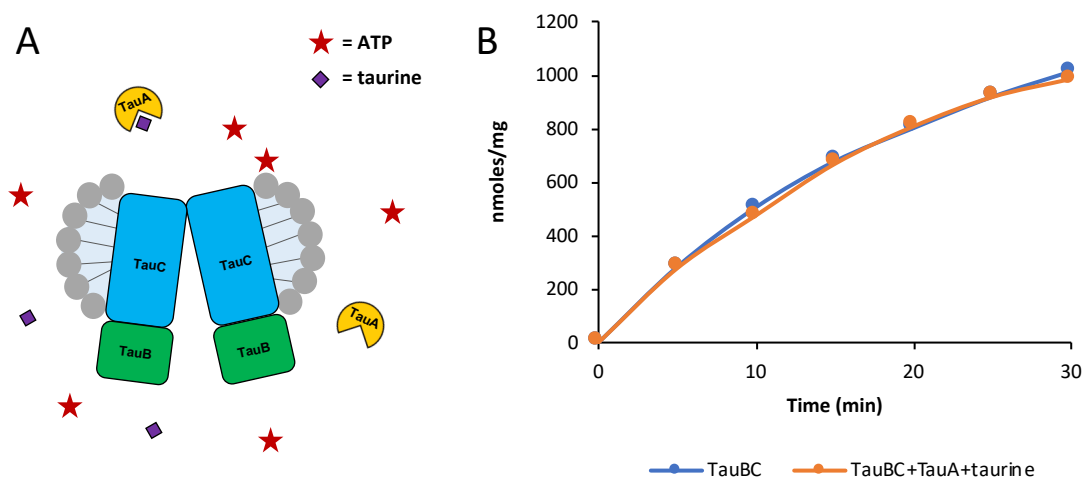


Figure 2.3.8.1. TauBC ATPase activity in detergent. (A) Schematic representation of TauBC in detergent micelles. (B) Time dependent ATPase activity of TauBC with (orange) and without (blue) TauA and taurine. TauBC shows a high basal ATPase activity, while with no stimulation of ATP hydrolysis was observed in the presence of TauA and taurine. The data correspond to the measurement of one experiment (n=1).

2.3.8.2. ATPase activity in proteoliposomes

To mimic a lipid environment, TauBC was reconstituted into proteoliposomes using a polar lipid extract from *E.coli*. Reconstitution into liposomes leads to two different orientations of TauBC in the lipid bilayer (Figure 2.3.8.2A). The basal activity of TauBC (1 μ M) displayed a V_{max} of 147.4 nmoles/min/mg in proteoliposomes (Figure 2.3.8.2B), which is 2.5-fold higher than that of DDM-solubilised TauBC (Figure 2.3.8.1B). Additionally, the ATPase assay only measures the ATPase activity of TauBC whose NBDs are outside the lipid bilayer, where ATP and all other components of the EnzChek kit are located. Therefore, unless all the TauBC molecules orient with their NBDs outside, the measured V_{max} represents only partial ATPase activity. For the ATPase stimulation test, TauA and taurine were added during proteoliposome reconstitution. There was no obvious stimulation of ATPase activity in the presence of TauA (10 μ M) and taurine (100 μ M) (V_{max} =142.2 nmoles/min/mg).

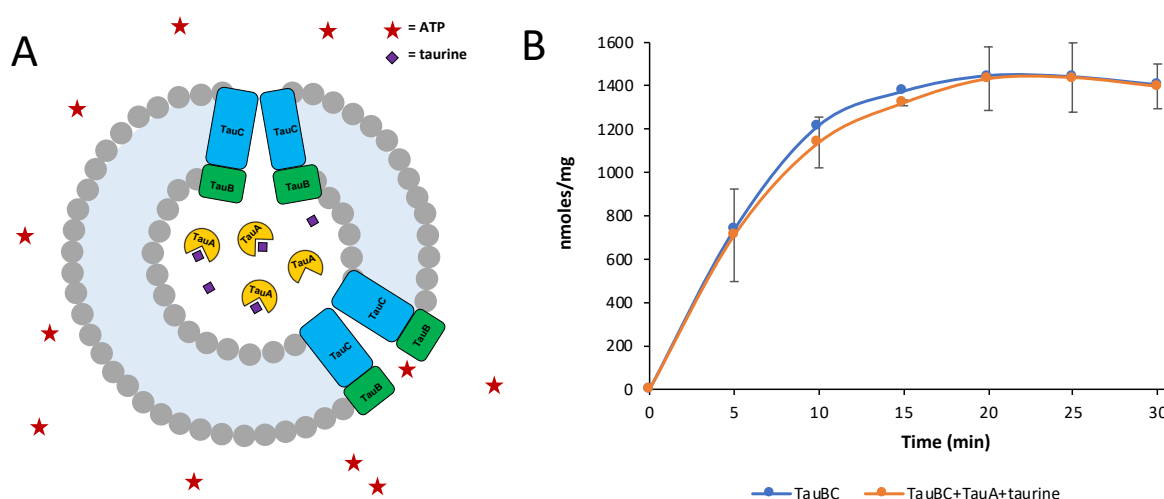


Figure 2.3.8.2. TauBC ATPase activity in proteoliposomes. (A) Schematic representation of TauBC in proteoliposomes. (B) Time dependent ATPase activity of TauBC with (orange) and without (blue) TauA and taurine. The ATPase activity of TauBC in proteoliposomes is 2.5-fold higher than that of DDM-solubilised TauBC. Addition of TauA and taurine did not stimulate the ATP hydrolysis of TauBC in proteoliposomes. The data correspond to mean values of two independent experiments from two batches of TauBC and TauA (n=2). Error bars indicate the $\pm$ standard deviation of recorded values.

2.4. Discussion

2.4.1. Cavity volume may determine substrate specificity

Structural comparison suggests that variation in the volumes of binding pockets might dictate the substrate selectivity between TauA and SsuA (Figure 2.3.6.1). Glu106 and Asp205 confine the binding cavity of TauA, whereas their equivalent residues either move away from the cavity (Glu88) or contain no side chain (Gly187). Due to the absence of these two confining residues, SsuA is able to coordinate larger alkanesulfonates such as MES, HEPES and MOPS (Eichhorn *et al.*, 2000). POCASA is an online server that predicts substrate-binding sites of query proteins by detecting pockets and cavities with a rolling sphere (Yu *et al.*, 2009). Interestingly, a POCASA search shows that the central cavity of TauA ($147 \text{ \AA}^3$) is 1.7-fold larger than that of SsuA, which is $255 \text{ \AA}^3$ in volume (Figure 2.4.1.1). The increased size of the SsuA-binding pocket aligns with the preference of SsuA for larger substrates.

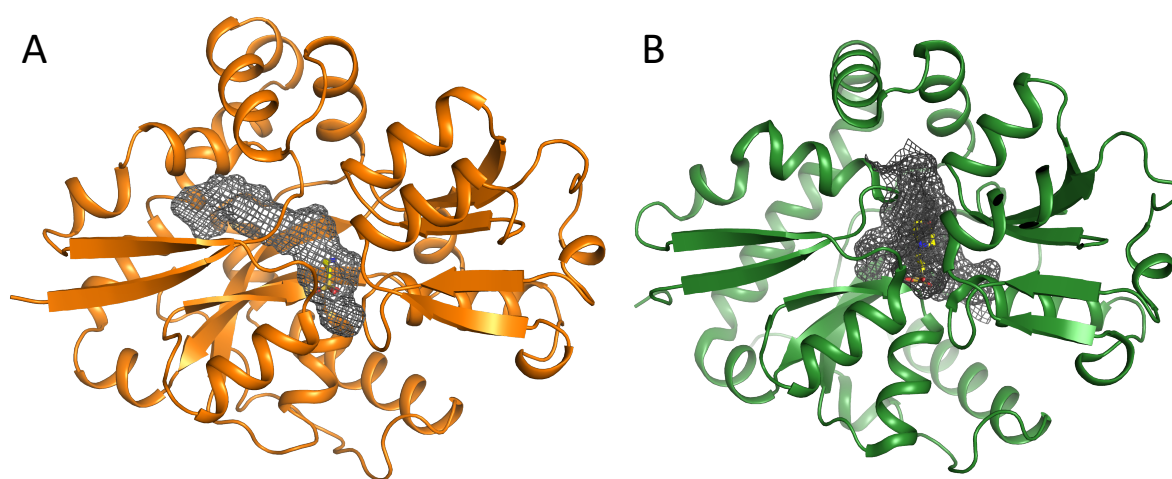


Figure 2.4.1.1. Cavity volumes of TauA and SsuA. (A) The TauA-aurine structure. (B) The SsuA-MES structure. The highlighted substrate-binding cavity volumes (grey mesh) are obtained with POCASA (Yu *et al.*, 2009).

Structural modulation of binding pocket volumes for ligand binding has been previously reported for other SBPs. The *Thermococcus litoralis* maltose-binding protein MBP and *Thermus thermophilus* glucose-binding protein GBP share 18% (60%) sequence identity (similarity). Superimposition of the MBP and GBP structures results in an RMSD of 2.6 Å over Cα atoms (Figure 2.4.1.2A) (Diez *et al.*, 2001; Cuneo *et al.*, 2006). MBP contains a large binding pocket that can accommodate disaccharides (*e.g.* maltose and trehalose). Glucose bound to GBP overlaps with one of the trehalose sugar rings in the MBP structure, whereas the site for the second trehalose sugar ring is occupied by an α-helix and two loops (Figure 2.4.1.2B). Therefore, GBP cannot coordinate di-saccharides. TauA does not contain secondary structure elements to confine the binding pocket. Instead, the smaller binding pocket of TauA compared to SsuA is owing to the presence of two residues, Glu106 and Asp205. Therefore, the binding pocket of TauA is more flexible. Upon local conformational changes, the binding pocket of TauA expands and thus TauA can accommodate larger taurine analogues.

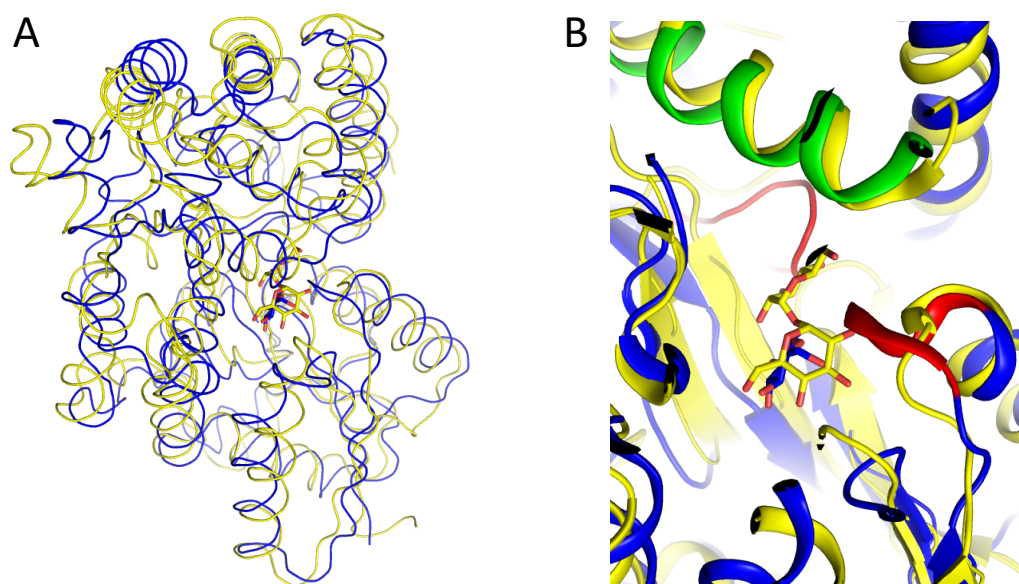


Figure 2.4.1.2. Comparison between MBP and GBP. (A) Superimposition of MBP from *T. litoralis* (yellow; PDB ID: 1EU8, (Diez *et al.*, 2001)) and GBP from *T. thermophilus* (blue; PDB ID: 2B3B, (Cuneo *et al.*, 2006)). The trehalose and glucose molecules are shown in stick format (yellow and blue respectively). Oxygen atoms are coloured in red. (B) Close-up view of the trehalose- and glucose-binding sites. An α-helix (green, residues 166-181) and two loops (red, residues 40-44 and 344-351) constrict the substrate-binding site of GBP.

2.4.2. Desolvation of the binding site confers promiscuity

Upon binding of different ligands, water distributions in the binding site of TauA vary greatly, indicating that water molecules might play a role in substrate binding. MD simulations show that desolvation of the binding site is another determinant in ligand binding of TauA. Upon binding of different ligands, different numbers of water molecules in the binding site of TauA are replaced by the ligands (Table 2.3.5.2). Desolvation of the binding site also results in a varied binding affinity (Table 2.3.5.1 and Figure 2.3.5.2). There are other SBPs that bind multiple substrates, like OppA from *Salmonella typhimurium* that binds peptides of two to five amino acid residues (Tame *et al.*, 1994; Tame *et al.*, 1996). In OppA from *S. typhimurium*, protein-ligand interactions are formed almost exclusively with the peptide backbone (Figure 2.4.2.1B) (Tame *et al.*, 1994). Interactions between ligand side chains and OppA do not contribute to ligand binding. Instead, OppA has a large hydrated cavity that enables accommodation of a wide range of chemical groups (Figure 2.4.2.1A). The enclosed water molecules serve as adaptable hydrogen bond donors or acceptors upon binding of different ligands (Tame *et al.*, 1994). The number of solvent molecules is adjusted according to the size of the ligand, where water molecules occupy the binding cavity not taken up by peptide side chains. Ligand binding of *S. typhimurium* OppA is therefore independent from the exact amino acid sequence of the peptide, owing to the presence of the hydrated binding cavity and its desolvation. Our work has shown that desolvation of the binding site is a key determinant in TauA binding for different ligands. However, the binding site of TauA is less hydrated than OppA, and taurine and its analogues are much smaller than oligopeptides. Therefore, binding affinities of TauA for different ligands vary more significantly (up to ~400000-fold) compared to OppA binding different peptides (~36-fold) (Tame *et al.*, 1996).

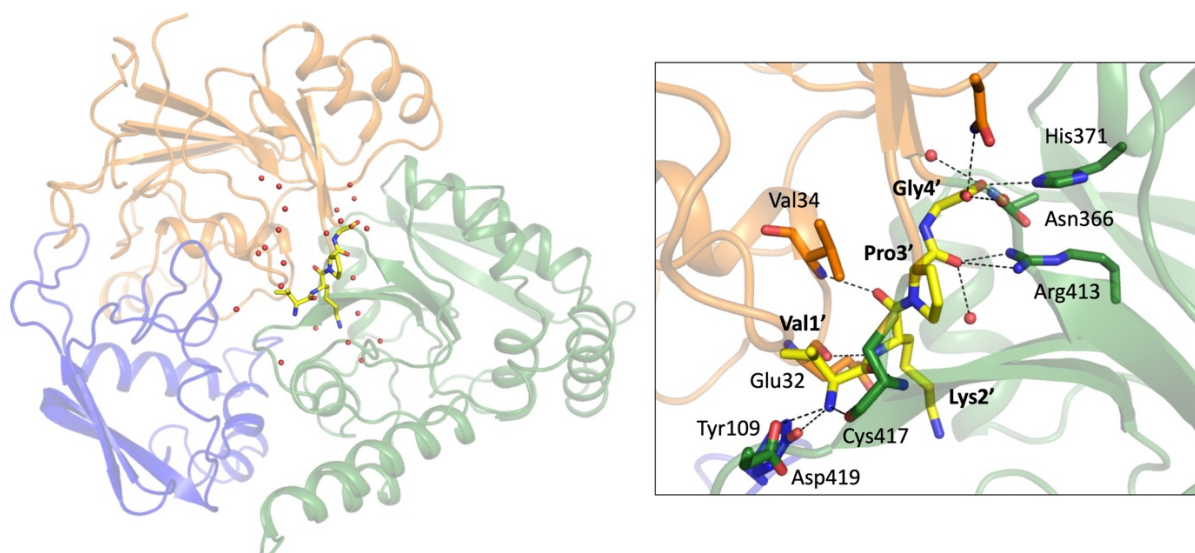


Figure 2.4.2.1. Structure of tetrapeptide-bound OppA from *S. typhimurium*. Domains Ia, Ib and II are coloured in orange, blue and green, respectively (PDB ID: 1OLA, (Tame *et al.*, 1994)). The inset shows the hydrogen bond interactions between OppA and the tetrapeptide (black dotted lines). The tetrapeptide and the peptide-binding residues are shown in stick format. Carbon atoms of the tetrapeptide, domain Ia, Ib and II residues of OppA are coloured in yellow, orange, blue and green, respectively. Sulfur, oxygen and nitrogen atoms are coloured in gold, red and blue, respectively. Water molecules are shown as red spheres.

2.4.3. Is TauABC a type I or type II ABC importer?

Our ATPase data indicate that purified TauBC is active in proteoliposomes ($V_{\max}=142.2$ nmoles/min/mg), while we did not observe stimulated ATPase by TauA and taurine. Based on our data, TauABC is likely to be a type II importer, which demonstrate high rate of basal ATPase activity and low stimulation by cognate SBPs (Lewinson & Livnat-Levanon, 2017). However, it is also possible that the current method of ATPase assay does not favour measurements of ATPase stimulation. The low nM binding affinity of TauA for taurine also suggests that TauABC might be a type II ABC importer. We used BLAST (Basic Local Alignment Search Tool) to search the sequence of TauC, the more heterogeneous component of the TauBC transporter, against proteins in PDB. All the BLAST results are type I ABC importers, including the D-methionine importer MetNI (Kadaba *et al.*, 2008), the amino acid importer ArtQN (Yu *et al.*, 2015), the molybdate/tungstate importer ModBC (Gerber *et al.*, 2008; Hollenstein *et al.*, 2007), the maltose importer MalFGK₂ (Khare *et al.*, 2009) and the acidic polysaccharide alginate importer AlgM₁M₂SS (Maruyama *et al.*, 2015) (sorted by alignment scores). This, however, shows that TauABC is a putative type I importer. Additionally, substrates of type I ABC importers are usually small, hydrophilic and are available at high concentrations (Locher, 2016). Taurine is small, highly hydrophilic and widely distributed in natural environment, which also suggests that TauABC is probably a type I importer. To date, the data we have do not support a clear classification of TauABC. Further work will be required to determine structures of TauBC in different transport states. Biophysical techniques, such as single-molecule Förster resonance energy transfer (smFRET) will also be used to probe conformational changes in TauBC during transport cycle. SmFRET measures distances at the 1-10 nm scale in single molecules, hereby retrieving conformational states and dynamics of ABC transporters (Gouridis *et al.*, 2015).

2.4.4. Conclusion and outlook

In this work, I have presented the structure of taurine-bound TauA and demonstrated the chemical basis for substrate discrimination. The TauA-aurine complex is the first example of a taurine-specific SBP and provides mechanistic insights into how bacteria tolerate low nutrient conditions. The nM binding affinity of TauA for taurine is consistent with reports for other SBPs such as *E. coli* thiamine-binding protein TbpA (Soriano *et al.*, 2008). In addition, the ability of TauA to bind more weakly to taurine analogues suggests that TauA has selectivity towards its physiological substrate taurine. Our data indicate that the acquired local conformational changes and desolvation of the substrate-binding site plays an important role in mediating this selectivity. Interestingly, desolvation of the substrate-binding site have previously been observed for OppA from *S. typhimurium* (Tame *et al.*, 1994). As such, we propose that this mechanism is likely to apply to other nutrient-specific SBPs. In recent years, it has become clear that SBPs display conformational dynamics that can be visualised by smFRET (Kim *et al.*, 2013). To compliment the structural insights gained from this study, future work will examine the taurine-dependent conformational changes on TauA using smFRET. Specifically, I am interested in defining the lifetimes of the open and closed states, and how taurine can alter the equilibrium between these states.

Following taurine recognition by TauA, docking of TauA at the ABC importer TauBC mediates delivery of taurine in the target cell. To our surprise, TauBC displayed a high basal ATPase activity in agreement with reports for type II ABC importers (Borths *et al.*, 2005). However, our inability to observe SBP-induced changes in ATPase activity suggests that modifications may be required to our assay in proteoliposomes. One possibility would be to reconstitute proteoliposomes using the biobeads method. Biobeads gradually remove detergents, providing a more gentle way of making proteoliposomes compared to the rapid

dilution method we used. We can also measure ATPase activity differently, for example, using the malachite green ATPase assay, which use malachite green as the indicator for phosphate liberation. If stimulated ATPase activity could be measured, we can test the stimulation effect of the TauA mutants and taurine analogues. Future work also includes determining structures of the TauBC transporter and the full-length TauABC importer. Conformational changes on TauBC during the transport cycle will also be examined by smFRET.

Chapter 3

Structural studies of *E. coli* YejA

The YejA structure will be published in collaboration with another group. We will contribute our structure alongside their structures and functional data.

3.0. Introduction

3.0.1. Antimicrobial peptides

Antimicrobial peptides contribute to innate immunity response in all domains of life, including microbes, plants and mammals (Reddy *et al.*, 2004). Antimicrobial peptides secreted by prokaryotic and eukaryotic cells mainly differ in their target specificity and efficacy. Generally, antimicrobial peptides produced by bacteria are more potent than those from eukaryotic cells. These peptides typically demonstrate high potency and a narrow spectrum of targets, making them promising candidates for therapeutic intervention.

In natural environments, diverse species of microorganism coexist and form microbial communities where they compete constantly for space and resources to survive (Fredrickson & Stephanopoulos, 1981). The use of antimicrobial peptides is probably the most ancient and widespread microbial competition strategy employed by microorganisms. Bacteriocins are potent antibacterial peptides and proteins that are ribosomally synthesized by Gram-negative bacteria to compete with phylogenetically close species (Cotter *et al.*, 2013). Most bacteriocins demonstrate antibacterial activity at pM or nM concentrations. Bacteriocins consist of colicins (*e.g.* colicin Ia and colicin 3E) and microcins (*e.g.* MccJ25 and MccC7/C51) (Sharma *et al.*, 2013; Duquesne *et al.*, 2007). Colicins represent a type of high MW antimicrobial proteins (30-80 kDa), whilst microcins are peptides that possess lower MW (1-10 kDa), high hydrophobicity and high tolerance to unfavourable conditions such as proteases, extreme temperature and pH (Rebuffat, 2013). The producer strain often confers protection against antimicrobial peptides by a dedicated self-immunity system (Duquesne *et al.*, 2007).

3.0.2. Microcins

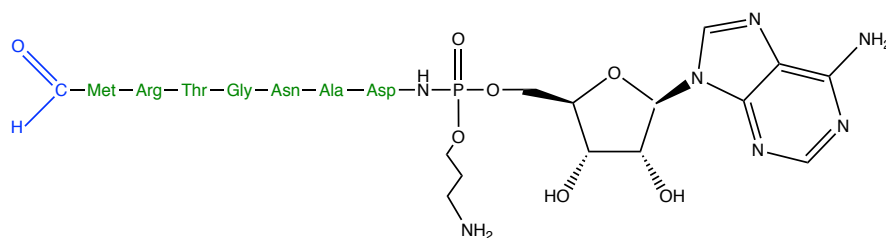
Microcins are antimicrobial peptides produced by Gram-negative bacteria, especially Enterobacteria (Baquero *et al.*, 1978; Duquesne *et al.*, 2007). Microcins are produced under unfavourable conditions such as stress and nutrient starvation (Connell *et al.*, 1987; Chiuchiolo *et al.*, 2001; Fomenko *et al.*, 2001). Mature microcins display low minimum inhibitory concentrations (MICs) (within a nM range) and a narrow spectrum of targets. To date, at least 16 microcins have been identified, among which only eight have been isolated and structurally characterised.

Microcins are divided into two classes based on their MWs, the presence of disulphide bonds and post-translational modifications (Table 3.0.2.1 and Figure 3.0.2.1) (Duquesne *et al.*, 2007). Class I consists of three plasmid-encoded microcins, MccB17, MccC7/C5 and MccJ25. Microcins of class I display MWs below 5 kDa and extensive backbone post-translational modifications. Class II contains microcins with higher MW (5-10 kDa) and is further divided into class IIa and class IIb. Class IIa includes three plasmid-encoded microcins (MccN, MccV and MccL) that are free from post-translational modifications and contain zero, one and two disulphide bond(s), respectively. Class IIb consists of four chromosome-encoded linear microcins that carry a C-terminal siderophore modification: MccE492, MccM, MccH47 and presumably MccI47. The complexity of microcin biosynthetic pathways allows high diversity in their structures and mechanisms of action. Such properties make microcins efficient weapons for bacterial infections and attractive substitution of conventional antibiotics.

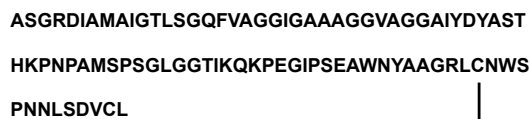
		Precursor	Leader peptide	Mature microcin	
Class	Microcin	Number of residues	Number of residues	Number of residues	Molecular mass (Da)
Class I	MccB17	69	26	43	3093
	MccC7/C51	7	0	7	1177
	MccJ25	58	37	21	2107
Class IIa	MccN	90	17	73	7457
	MccV	103	15	88	8733
	MccL	105	15	90	8884
Class IIb	MccE492	103	19	84	8717
	MccM	92	15	77	7283
	MccH47	75	15	60	4865
	MccI47	77	15	62	6276

Table 3.0.2.1. Structural characteristics of secreted microcins. Adapted from (Duquesne *et al.*, 2007).

Class I
MccC7/C51



Class IIa
MccV



Class IIb
MccE492

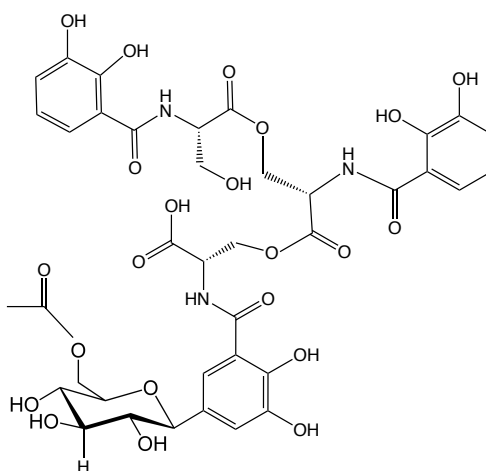


Figure 3.0.2.1. Representative examples of different microcin classes. Microcins of class I (*e.g.* MccC7/C51) have MWs below 5 kDa and extensive backbone post-translational modifications. Microcins of class II have MWs between 5-10 kDa. Class IIa microcins (*e.g.* MccV) lack post-translational modifications and contain various numbers of disulphide bond(s) (indicated by the black line). Class IIb contains chromosome-encoded linear microcins (*e.g.* MccE492) carrying C-terminal siderophore modifications.

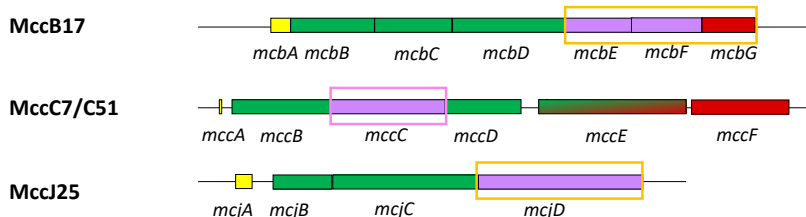
3.0.3. Microcin genetic systems

Microcins are encoded by plasmid or chromosomal gene clusters. Such gene systems contain a conserved organisation of open reading frames (Table 3.0.3.1 and Figure 3.0.3.1). The minimal structure of these gene clusters is composed of (1) a structural gene encoding the peptide precursor, (2) a self-immunity gene that generally sits adjacent to the former and encodes the self-immunity factor to protect the producing strain against its own antimicrobial substance, and (3) a gene encoding the efflux transport system required for the secretion of microcins (Duquesne *et al.*, 2007). Gene encoding post-translational modification enzymes are also found in the open reading frames, in which case microcin precursors undergo a maturation process, leading to the formation of mature microcins.

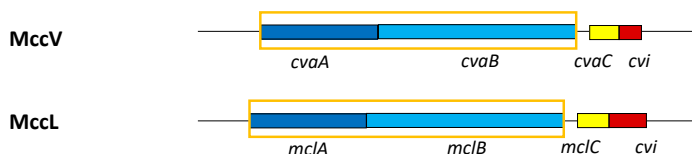
Microcin	Structural gene	Self-immunity gene(s)	Export gene(s)	Post-translational modification genes	Gene of unknown function
MccB17	<i>mcbA</i>	<i>mcbE, mcbF, mcbG</i>	<i>mcbE, mcbF</i>	<i>mcbB, mcbC, mcbD</i>	-
MccC7/C51	<i>mccA</i>	<i>mccC, mceE, mceF</i>	<i>mccC</i>	<i>mccB, mceD, mceE</i>	-
MccJ25	<i>mcjA</i>	<i>mcjD</i>	<i>mcjD</i>	<i>mcjB, mcjC</i>	-
MccV	<i>cvaC</i>	<i>cvi</i>	<i>cvaA, cvaB</i>	-	-
MccL	<i>mclC</i>	<i>mclI</i>	<i>mclA, mclB</i>	-	-
MccE492	<i>mceA</i>	<i>mceB</i>	<i>mceH, mceG, mceF</i>	<i>mceC, mceD, mceI, mceJ</i>	<i>mceE</i>

Table 3.0.3.1. Typical gene clusters characterising microcins of different classes. Adapted from (Duquesne *et al.*, 2007).

Class I microcins



Class IIa microcins



Class IIb microcins

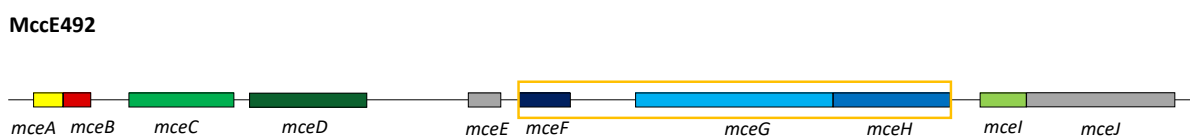


Figure 3.0.3.1. Genetic organisation of the microcin systems. Orange and pink boxes highlight the ABC and MFS transporters, respectively. Genes encoding microcin precursors, proteins for self-immunity, export and post-translational modification are coloured in yellow, red, blue and green, respectively. The purple box indicates genes involved in both self-immunity and export. Green and red gradation shows *mccE*, which is required for both post-translational modification and self-immunity. Genes with unknown function are coloured in grey. Adapted from (Duquesne *et al.*, 2007).

3.0.4. Microcin biosynthesis and export

Most microcins are produced as an inactive peptide precursor comprising a N-terminal leader peptide and a C-terminal core peptide that constitutes the final mature microcin (Rebuffat, 2013). Microcin C does not have a leader peptide and is the only microcin known to be exported without cleavage (Garcia-Bustos *et al.*, 1985). Leader peptides of class I microcins are variable in length (26 and 37 amino acids for MccB17 and MccJ25, respectively) and do not share a common sequence (Duquesne *et al.*, 2007). Class II microcins generally contain a shorter leader peptide of ~15 amino acids in length and a conserved sequence M-R-X-I/L-X₉-G-G/A (where X is any amino acid) (Duquesne *et al.*, 2007). Although the precise role has yet to be clarified, leader peptides are often linked to inactivation, protection against degradation, precursor maturation and microcin secretion.

Upon removal of leader peptides, mature microcins are exported to the extracellular medium. There is no standard export model for class I microcins (Garrido *et al.*, 1988; Gonzalez-Pastor *et al.*, 1995; Fomenko *et al.*, 2003; Solbiati *et al.*, 1999; Delgado *et al.*, 1999). Class I microcins use either ABC transporters (MccB17 and MccJ25) or secondary transporters (MccC7/C51), which are encoded by the microcin gene cluster, for their export. Secretion of class II microcins displays a canonical model. It involves (1) the ABC transporter that cleaves the leader peptides and exports the mature microcins across the inner membrane, (2) the OMF TolC that enables export of microcins to the extracellular medium by forming a pore in the outer membrane, and (3) the accessory protein that serves as a connector between (1) and (2) (Gilson *et al.*, 1990; Gaggero *et al.*, 1993; Lagos *et al.*, 2001; Pons *et al.*, 2004). Both the ABC transporter and its accessory protein are encoded by the microcin gene system, while the TolC protein is encoded by another gene cluster.

3.0.5. Microcin uptake and mechanism of action

Analysis of microcin-resistant cells has revealed the mechanism of microcin uptake (Figure 3.0.5.1). Several microcins hijack nutrient uptake receptors to enter bacterial cells (Duquesne *et al.*, 2007). Most microcins use iron siderophore receptors to penetrate the outer membrane, either of the hydroxamate-type (FhuA) in the case of MccJ25, or the catecholate-type (FepA, Cir, Fiu) for MccV, MccL, MccE492, MccM and MccH47. By contrast, MccB17 and MccC7/C51 rely on the outer membrane channel OmpF (Laviña *et al.*, 1986; Salomón & Farias, 1993; Patzer *et al.*, 2003; Destoumieux-Garzón *et al.*, 2006). The uptake of many microcins (MccJ25, MccE492, MccH47 and MccM) requires the TonB/ExbB/ExbD inner-membrane complex for energy transduction (Pugsley *et al.*, 1986; Salomón & Farias, 1995; Trujillo *et al.*, 2001; Patzer *et al.*, 2003). The inner-membrane protein SbmA is required for the import of MccJ25 and MccB17 (Laviña *et al.*, 1986; Salomón & Farias, 1995).

Class I microcins mainly target intracellular enzymes involved in DNA/RNA synthesis (Figure 3.0.5.1) (Duquesne *et al.*, 2007). MccJ25 binds to RNA polymerase secondary channel, thus preventing nucleoside triphosphates from trafficking to the active site of RNA polymerase (Adelman *et al.*, 2004). MccB17 is a DNA gyrase inhibitor shown to induce irreversible accumulation of DNA-DNA gyrase complexes, inducing the SOS response and blocking DNA replication (Vizán *et al.*, 1991). MccC7/C51 gains antimicrobial activity upon *in vivo* cleavage of its peptide region. The processed MccC7/C51 (MccC7/C51*) is an aspartyl tRNA synthetase inhibitor that blocks aminoacylated tRNA synthesis (Metlitskaya *et al.*, 2006). Class II microcins, however, target the cell phospholipid bilayers or inner-membrane proteins (Duquesne *et al.*, 2007). MccV and MccE492 disrupt inner-membrane permeabilization, while MccH47 targets the F₀F₁ ATP synthase (Yang & Konisky, 1984; Trujillo *et al.*, 2001; Bieler *et al.*, 2006).

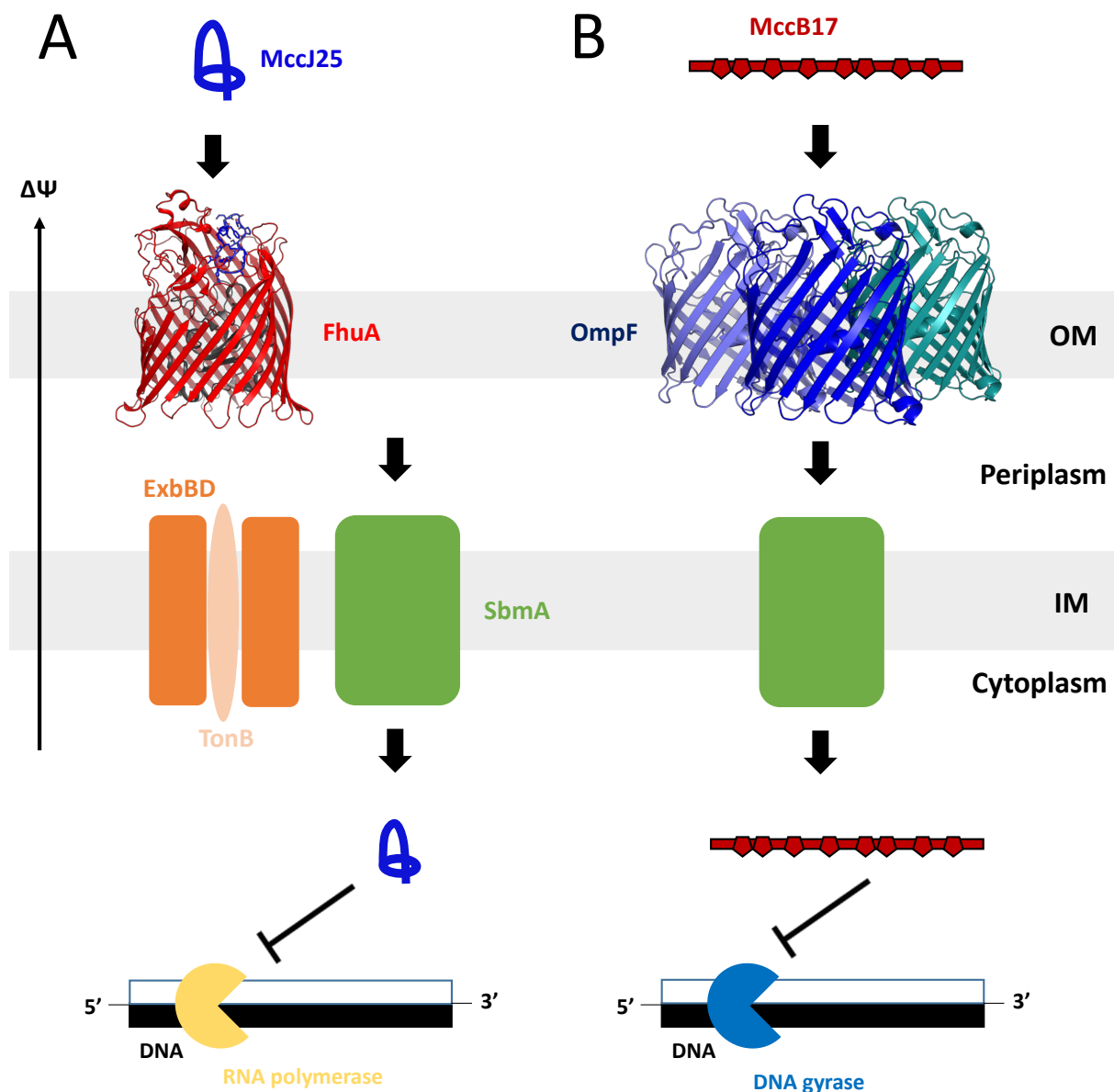


Figure 3.0.5.1. MccJ25 and MccB17 uptake and mechanism of action. (A) MccJ25 enters *E. coli* cells through the outer-membrane porin FhuA via a TonB/ExbB/ExbD-dependent mechanism and the inner-membrane permease SbmA. Intracellular MccJ25 inhibits the activity of RNA polymerase by obstructing the passage of incoming nucleoside triphosphates during protein synthesis. (B) MccB17 enters the periplasm and cytoplasm through the outer-membrane channel OmpF and the inner-membrane protein SbmA, respectively. MccB17 inhibits the DNA supercoiling and relaxation reactions of DNA gyrase, thus preventing DNA replication. Adapted from (Duquesne *et al.*, 2007).

3.0.6. MccJ25: mechanism for hijacking FhuA

MccJ25 is a hydrophobic peptide of 21 amino acids in length (Figure 3.0.6.1) (Solbiati *et al.*, 1996; Rosengren *et al.*, 2003). MccJ25 shows potent antibacterial activity against a narrow spectrum of Enterobacteria species, such as *E. coli* and certain strains of *Salmonella* (Duquesne *et al.*, 2007). MccJ25 exhibits a distinct lasso topology. The linkage between Gly1' and Glu8' forms a macrolactam ring (Gly1'-Glu8'), through which a C-terminal tail (Tyr9'-Gly21') is threaded and sterically locked by two bulky aromatic residues, Phe19' and Tyr20', that are located at each side of the ring (Rosengren *et al.*, 2003). MccJ25 inhibits the activity of RNA polymerase.

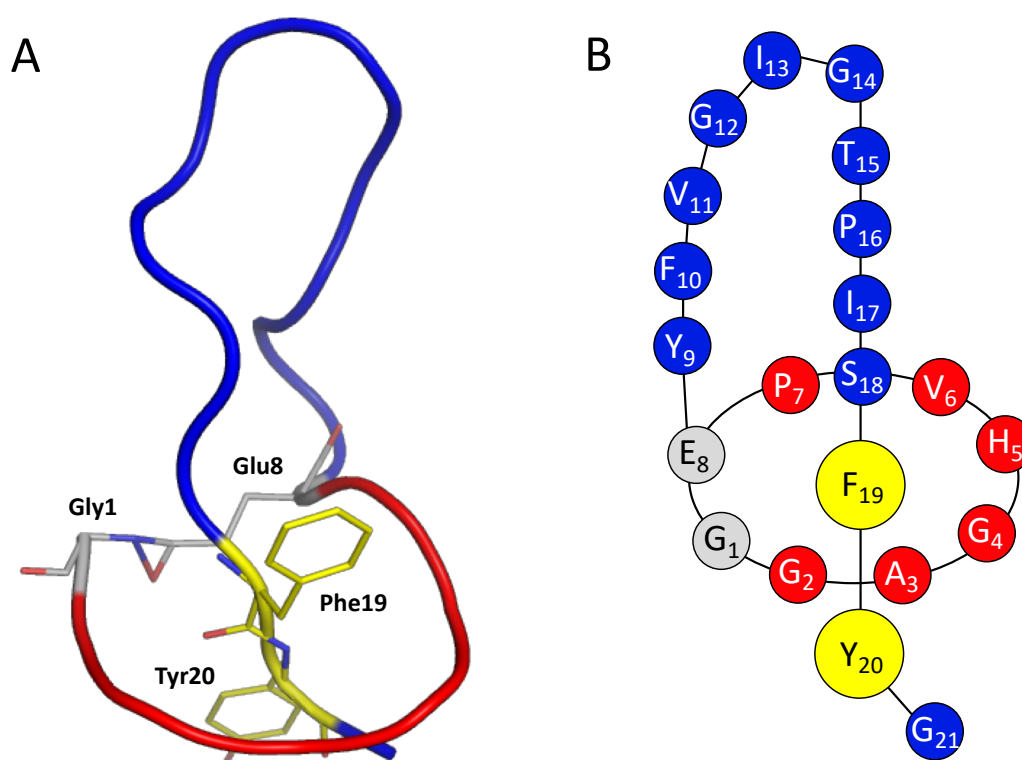


Figure 3.0.6.1. Structure of MccJ25. (A) Ribbon representation of MccJ25 (PDB ID: 1Q71, (Rosengren *et al.*, 2003)). (B) Graphic illustration of MccJ25 structure and sequence. Different colours indicate the individual components: N-terminal macrolactam ring (red), C-terminal tail (blue), lactam bound linkage between Gly1' and Glu8' (grey), two aromatic residues Phe 19' and Tyr20' (yellow).

FhuA is the receptor of the Fe (III)-chelated hydroxamate siderophore ferrichrome, which enables iron internalisation for bacterial survival and growth (Locher *et al.*, 1998). There is no structural similarity between ferrichrome and MccJ25. FhuA constitutes an N-terminal plug domain and a β -barrel domain composed of 22 antiparallel β -stands (Locher *et al.*, 1998). The plug domain is folded inside the β -barrel from the periplasmic side, which provides a large ligand-binding pocket opening to the extracellular medium. The TonB/ExbB/ExbD complex transduces the proton motive force across the inner membrane to FhuA for substrate transport.

MccJ25 hijacks FhuA by mimicking the interaction mode of ferrichrome to induce TonB/ExbB/ExbD-dependent transport (Locher *et al.*, 1998; Mathavan *et al.*, 2014). The structure of FhuA-MccJ25 reveals that MccJ25 sits vertically inside the β -barrel and fully occupies the substrate translocation pathway of FhuA. Ferrichrome forms extensive hydrogen bonds with the plug domain of FhuA, whereas only two MccJ25 residues (Ala3' and His5') form three hydrogen bonds with the plug domain, suggesting that minimal conserved interactions are sufficient for MccJ25 to hijack FhuA. The structure shows that His5' inserts more deeply inside the plug. Competition assays with phage T5 and native mass spectrometry (MS) suggest that His5' is essential for MccJ25 recognition (Mathavan *et al.*, 2014). The loop of MccJ25 (residues Tyr9'-Ser18') is also important for FhuA recognition. Once MccJ25 binds to the extracellular binding pocket of FhuA, the plug domain undergoes major conformational changes, enabling MccJ25 transport (Locher *et al.*, 1998; Mathavan *et al.*, 2014).

Microcins have evolved a mechanism to enter bacterial cells via naturally occurring nutrient uptake systems (Duquesne *et al.*, 2007). Examples of microcin internalisation include MccJ25 hijacking FhuA and SbmA, and MccB17 hijacking OmpF and SbmA. Another such import machinery is the YejABEF ABC importer that is hijacked by MccC7/C51. The structural mechanism of how MccC/C51 hijacks YejABEF remains unknown.

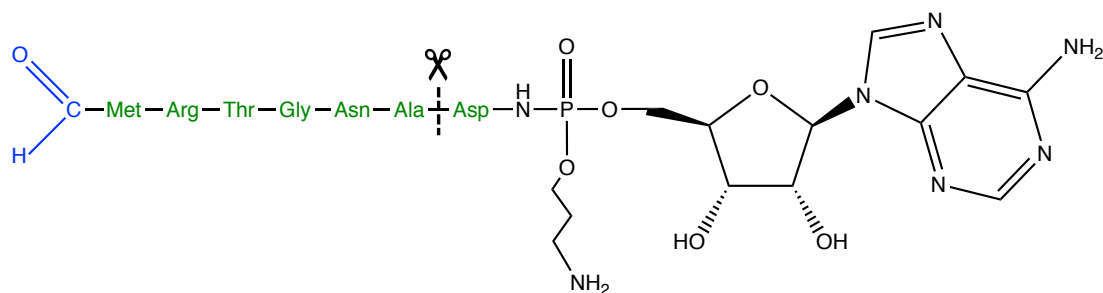
3.1. YejA

3.1.1. MccC7/C51

MccC7/C51 is the smallest microcin isolated to date (Duquesne *et al.*, 2007). It is produced by *E. coli* strains (*e.g.* MC4100 or TG1) harbouring either a 43-kb single-copy pMccC7 plasmid (Garcia-Bustos *et al.*, 1984) or a 38-kb low-copy number pMccC51 plasmid (Kurepina *et al.*, 1993). The two gene clusters share 98-100% sequence identity (Fomenko *et al.*, 2003). MccC7 and MccC51 display an identical structure (Figure 3.1.1.1A).

Full-length MccC7/C51 is an N-formylated heptapeptide with a modified adenosine monophosphate (AMP) covalently attached to its C-terminal aspartate through an N-acyl phosphoramidate bond (Guijarro *et al.*, 1995). MccC7/C51 is the only microcin that is post-translationally modified by the addition of a nucleotide. Intact MccC7/C51 does not show any inhibitory effect. In the cytoplasm, MccC7/C51 are cleaved forming MccC7/C51* (Figure 3.1.1.1B), an analogue of aspartyl-adenylate (Kazakov *et al.*, 2008).

A. MccC7/C51



B. MccC7/C51*

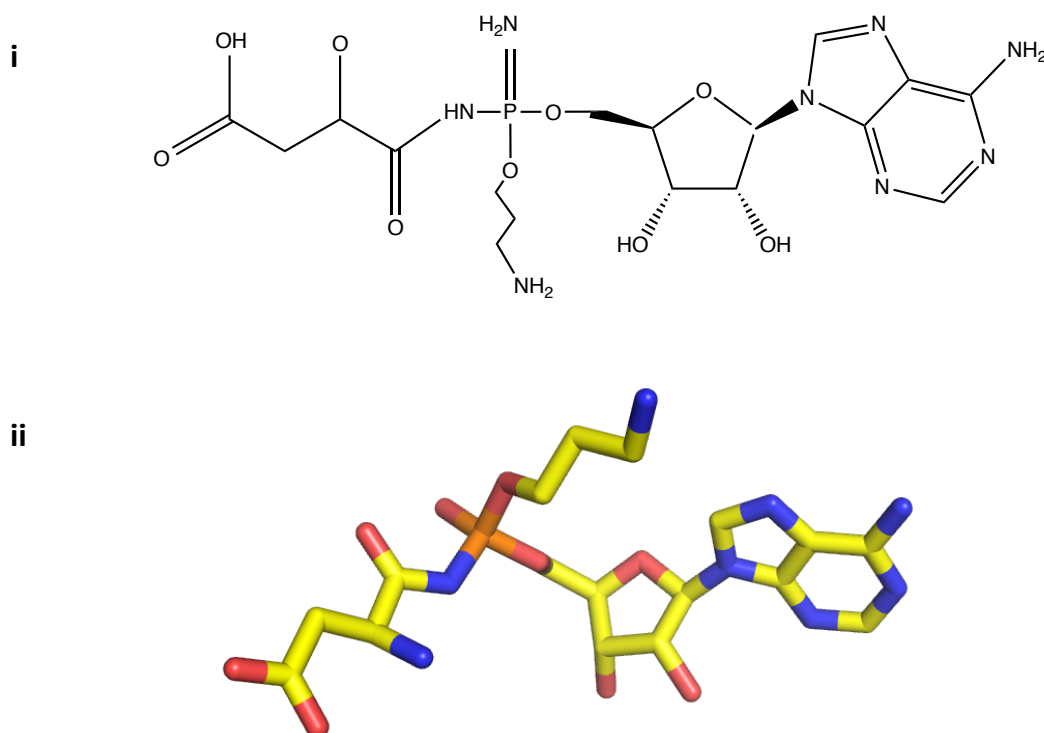


Figure 3.1.1.1. Structures of MccC7/C51 and processed MccC7/C51*. MccC7/C51 (A) is an N-formylated heptapeptide with a modified AMP covalently attached to its C-terminal aspartate. Upon entering susceptible bacterial cells, the peptide is intracellularly cleaved between Ala6' and Asp7', resulting in (B) the processed MccC7/C51*, a translation inhibitor. The scissors indicate the cleavage site of MccC7/C51. The graphic illustration (i) and crystal structure (stick format) (ii) of MccC7/C51*. Carbon, oxygen, nitrogen and phosphorus atoms are coloured in yellow, red, blue and orange, respectively (Agarwal *et al.*, 2012).

MccC7/C51 is produced by *E. coli* strains harbouring plasmids with the *mccABCDEF* gene cluster (Table 3.0.3.1 and Figure 3.0.3.1). The *mccABCDE* gene cluster is transcribed from a single promoter, while the downstream *mccF* gene is transcribed in the opposite direction (Gonzalez-Pastor *et al.*, 1995; Fomenko *et al.*, 2003). Transcription of the *mccABCDE* promoter is activated upon nutrient deprivation (*i.e.* stationary stage). It has been proposed that the *mccF* gene is expressed even when cells are in the logarithmic phase.

The 21-bp structural gene, *mccA*, encodes the 7-amino-acid peptide precursor MccA (Met'-Arg'-Thr'-Gly'-Asn'-Ala'-Asn') (Figure 3.1.1.2) (González-Pastor *et al.*, 1994). The genes *mccB*, *mccD* and *mccE* encode proteins responsible for post-translational modification of MccA, leading to formation of the mature MccC7/C51 (Figure 3.1.1.2) (Gonzalez-Pastor *et al.*, 1995; Fomenko *et al.*, 2003). MccB contains a nucleotide-binding domain (NBD) and a metal-binding motif, which resembles E1 ubiquitin-activating enzymes that catalyse the covalent attachment of AMP to proteins (Taylor *et al.*, 1998; Leimkühler *et al.*, 2001). MccB attaches the modified AMP to MccA *in vitro* at the expense of two ATP molecules per turnover (Roush *et al.*, 2008). MccD and MccE are both required for attachment of the aminopropyl group to the MccA-AMP intermediate, a product of the MccB-catalysed reaction. Addition of the aminopropyl group increases the biological activity by ~10 fold (Metlitskaya *et al.*, 2009). MccD shares sequence homology to S-adenosylmethionine-dependent methyltransferases (Fomenko *et al.*, 2003). Disruption of MccD results in formation of MccC7/C51 without the aminopropyl group. MccE consists of a N-terminal domain that resembles pyridoxal 5'-phosphate-dependent amino acid decarboxylases and a C-terminal region that shows similarities to RimL, an enzyme that catalyses the acetylation of ribosomal proteins (Fomenko *et al.*, 2003; Vetting *et al.*, 2005). The N- and C-terminal domains of MccE are required for post-translational modification of MccC7/C51 and self-immunity of the producing strains, respectively. It is proposed that the N-terminal domain of MccE catalyses the formation of the

n-aminopropanol through a homoserine decarboxylation reaction (Gonzalez-Pastor *et al.*, 1995; Fomenko *et al.*, 2003). MccD then transfers the n-aminopropanol to AMP, forming mature MccC7/C51. In summary, three post-translational steps are required for the maturation of MccC7/C51: (1) conversion of Asn7' to an Asp amide, which is catalysed by MccB, (2) formation of an N-acyl phosphoramidate bond between the Asp amide group and the α -phosphate of AMP by MccB, and (3) aminopropylation of the MccC7/C51 intermediate, which is achieved by MccD and the N-terminal domain of MccE.

Bacteria that produce microcins must protect themselves against toxicity. MccC and the C-terminal domain of MccE are involved in self-immunity towards MccC7/C51 in *E. coli* (Figure 3.1.1.2). MccC is an MFS exporter. MccC confers self-immunity to the producer cells by exporting microcins, hereby reducing the accumulation of MccC7/C51 (Fomenko *et al.*, 2003). The outer-membrane component that extrudes MccC7/C51 to the extracellular medium remains unknown. The C-terminal domain of MccE is an acetylase (Novikova *et al.*, 2010). It detoxifies processed MccC7/C51* by acetylating the amino group of the amino acid adenylate (Agarwal *et al.*, 2011). The additional acetyl group interferes with the formation of hydrogen bonds and results in steric clashes between MccC7/C51* and its cellular target. *mccF* encodes the third self-immunity protein, MccF (Gonzalez-Pastor *et al.*, 1995; Fomenko *et al.*, 2003). MccF is a peptidase that hydrolyses the carboxamide bond that links aspartate with the nucleotide (Tikhonov *et al.*, 2010). Both MccC7/C51 and MccC7/C51* are substrates of MccF.

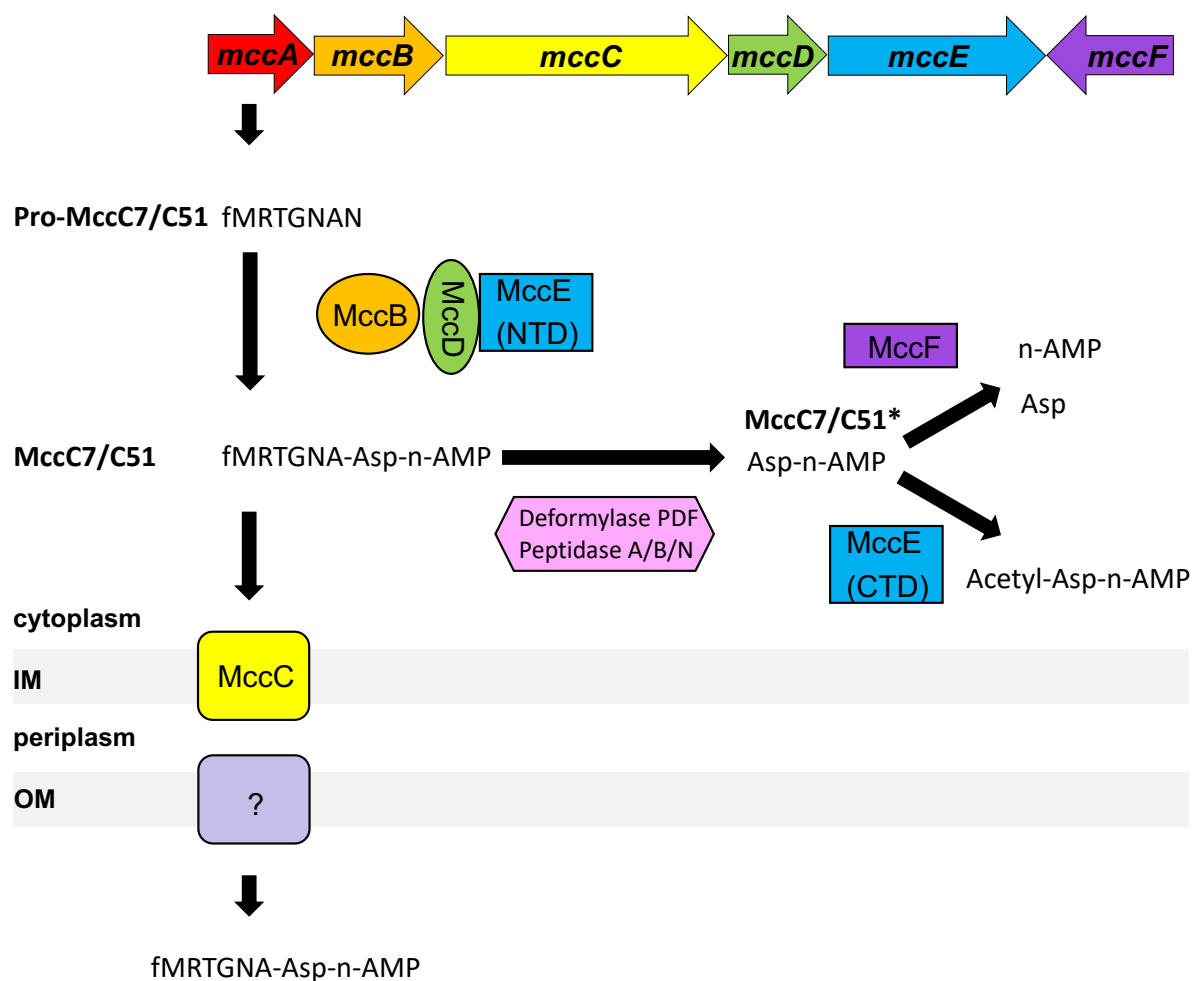


Figure 3.1.1.2. MccC7/C51 maturation and export. The peptide precursor MccA undergoes post-translational modifications by MccB, MccD and the N-terminal domain of MccE, resulting in formation of mature MccC7/C51. MccC, the C-terminal domain of MccE and MccF are involved in self-immunity against MccC7/C51 in *E. coli*. MccC is an MFS-type exporter that transports MccC7/C51 into the periplasm. The outer-membrane component that extrudes MccC7/C51 to the extracellular medium is unknown. The C-terminal domains of MccE and MccF protect the producer cells by inactivating the processed MccC7/C51*. Adapted from (Severinov *et al.*, 2011).

3.1.2. MccC7/C51 hijacks OmpF and YejABEF

MccC7/C51 hijacks two nutrient uptake systems to enter sensitive bacteria: the outer-membrane porin OmpF and the inner-membrane ABC importer YejABEF (Novikova *et al.*, 2007). Porins represent a group of hollow β -barrel proteins that form a pore in cellular membranes to enable passive diffusion of different types of molecule (Schirmer, 1998). Porins mostly exist in the outer membrane of Gram-negative bacteria. ABC transporters and bacterial ABC importers are discussed in detail in sections 1.0.5.

MccC7/C51 enters the outer membrane of sensitive bacteria through either the porin OmpF or other yet-identified outer-membrane transporters (Novikova *et al.*, 2007). The periplasmic MccC7/C51 subsequently penetrates the inner membrane via the ABC importer YejABEF (Novikova *et al.*, 2007). Since *yef* mutants are highly resistant to MccC7/C51, its maturation intermediates and chemical analogues, the YejABEF system has been shown to be the only inner-membrane transporter involved in MccC7/C51 uptake in *E. coli* (Novikova *et al.*, 2007). Once into the cytoplasm of sensitive cells, MccC7/C51 undergoes a series of cleavage reactions by the peptide deformylase PDF to deformylate the first methionine residue and one or more of the three none-specific peptidase PepA, B and N to remove the peptide tail. The resulting product, the processed MccC7/C51, is an analogue of aspartyl adenylate and thus is a translation inhibitor targeting on aspartyl-tRNA synthetase (Kazakov *et al.*, 2008). Intact MccC7/C51 shows no inhibitory effect on translation, while the processed MccC7/C51* does not affect cell growth of MccC7/C51-sensitive cells (Novikova *et al.*, 2007). Thus, MccC7/C51 is thought to be a Trojan horse-type antimicrobial compound (Reader *et al.*, 2005): the peptide chain enables active uptake of MccC7/C51 by YejABEF, and cleavage of the peptide allows MccC7/C51 to release the inhibitory aminoacyl-nucleotide moiety. A graphic illustration of MccC7/C51 uptake and the mechanism of action is presented in Figure 3.1.2.1.

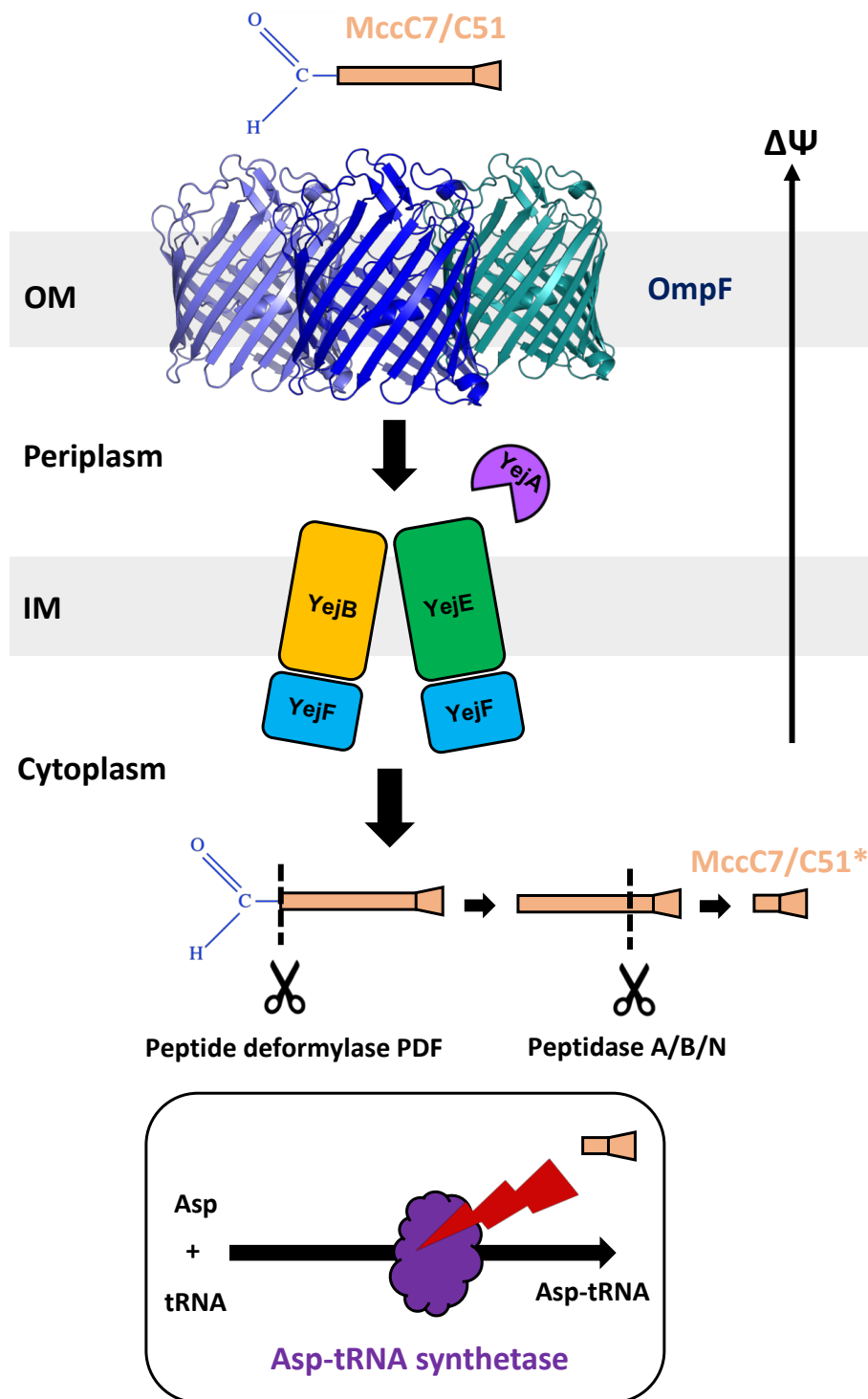


Figure 3.1.2.1. MccC7/C51: uptake and mechanism of action. MccC7/C51 penetrates the outer membrane through the porin OmpF and the inner membrane through the ABC importer YejABEF. Once in the cytoplasm, MccC7/C51 undergoes serial cleavages by PDF and peptidase A/B/N, leading to formation of the processed MccC7/C51*, a modified aspartyl-adenylate. MccC7/C51* is a translation inhibitor that targets aspartyl-tRNA synthetase. The scissors indicate cleavage of MccC7/C51.

3.1.3. YejABEF

YejABEF is a member of the ABC superfamily (Saurin *et al.*, 1999). Genes encoding YejA, YejB, YejE, YejF as well as an additional gene, *yejG*, are located in a single operon (Figure 3.1.3.1). The product of the *yejG* gene has no functional relationship to the YejABEF system (Novikova *et al.*, 2007). The physiological role of YejG is unknown. In *E. coli*, the *yejABEF* gene cluster is transcribed in one direction, while transcription of the downstream *yejG* gene is in the opposite direction (Novikova *et al.*, 2007). YejB and YejE are homologous integral transmembrane domains (TMDs) that form translocation pores in the inner membrane; YejF forms homodimeric cytoplasmic ATP-binding domains (NBDs) that drive substrate translocation via ATP binding and hydrolysis. YejA is the periplasmic substrate binding protein (SBP) (Antal *et al.*, 2005). The complete subunit composition of the ABC transporter is YejABEF₂. It is proposed that YejA captures the substrate from the periplasm and delivers it to the transmembrane proteins YejB and YejE. The substrate is subsequently translocated into the cytoplasm upon ATP binding and hydrolysis by the homodimeric YejF proteins. Deletion of each of the *yejABEF* genes leads to complete resistance to MccC7/C51, suggesting that all the four components are essential for MccC7/C51 uptake (Novikova *et al.*, 2007).

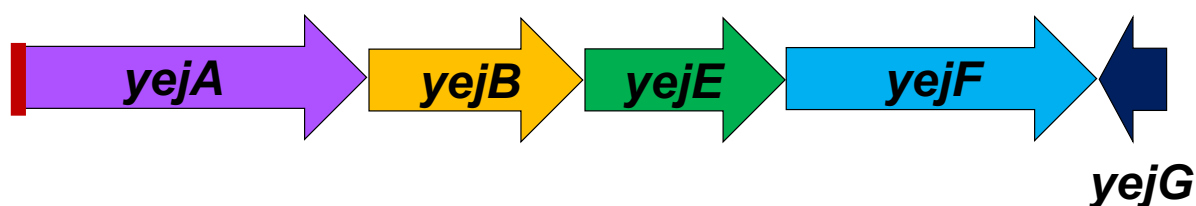


Figure 3.1.3.1. Organisation of the *yejABEFG* operon from *E. coli*. Arrows indicate the direction of gene transcription. *yejA*, *yejB/E* and *yejF* encodes the SBP, TMDs and NBDs of the ABC transporter YejABEF, respectively. The red box indicates the sequence encoding the signal peptide of YejA. Accession codes: P33913 (YejA), P0AFU0 (YejB), P33915 (YejE), P33916 (YejF), P0AD21 (YejG).

The mechanism of how microcins hijack certain transport proteins has been previously elucidated (*e.g.* MccJ25 hijacking FhuA), but not for ABC importers. It is unknown how MccC7/C51 hijacks the ABC importer YejABEF and what is the physiological substrate (*i.e.* MccC7/C51 unrelated) of the YejABEF system. Based on sequence similarity, YejABEF probably constitute an ABC importer that is involved in oligopeptide internalisation. However, there is currently no direct evidence demonstrating its physiological function and natural substrate. SBPs are responsible for substrate recognition and therefore dictate specificity to the cognate ABC importer. Substrate binding to YejA is an important initial step in its transport through YejABEF. Vondenhoff *et al.* (2011) have characterised the minimal peptide chain length required for YejABEF-mediated uptake of MccC7/C51 analogues. However, structure of YejA is not yet been determined. No structural information of how MccC7/C51 interacts with YejA is currently available.

3.1.4. Aim of research

The aim of this research is to illustrate structural details governing substrate recognition and binding of YejA, thereby providing information for structure-based design of novel antibiotics.

Specific objectives are as follows:

1. To determine structures of YejA and YejA in complex with MccC7/C51.
2. To determine the binding affinity between YejA and MccC7/C51 or its analogues.
3. To identify the physiological substrate(s) of YejABEF in light of the structures.

3.2. Materials and Methods

3.2.1. *yejA* cloning

The *yejA* gene is present in genomes of a range of different organisms including *E. coli* (1815 bp), *Salmonella enterica* (1823 bp), *Pseudomonas putida* (1836 bp), *S. dysenteriae* (1815 bp) and *Ketogulonicigenium vulgare* (1917 bp). *yejA* in *E. coli* encodes a protein of 604 amino acids in length (69,669 Da). The gene fragment encoding mature YejA of *E. coli* (residues 20-604, accession number: P33913) was amplified using *E. coli* chromosomal DNA as a template. Primers *yejAFw* and *yejARv* were designed and purchased from Invitrogen (Table 3.2.1.1). The amplified *yejA* gene was inserted into the pEHisTEV vector (Liu & Naismith, 2009) using the NcoI and XhoI restriction sites. This resulted in a construct encoding an N-terminal His₆-tagged YejA protein with a TEV protease cleavage site (Figure 3.2.1.1). Three amino acids remain at the N-terminus of the mature protein owing to the linker and the TEV protease site. The pEHisTEV-*yejA* was verified by DNA sequencing (Source BioScience).

Primer	Sequence
<i>yejAFw</i>	5'-CATGCCATGGTGAACGTCACCGTGGCG-3'
<i>yejARv</i>	5'-CCGCTCGAGTTATTGCACGAAGCGCGAG-3'

Table 3.2.1.1. Primer sequences for *yejA* gene.

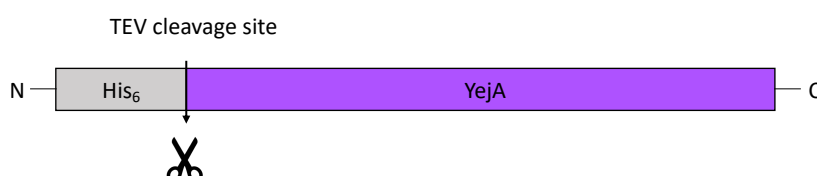


Figure 3.2.1.1. Schematic representation of the His₆-TEV-YejA construct.

3.2.2. YejA overexpression

The pEHisTEV-*yejA* construct was transformed into the BL21 (DE3) strain of *E. coli*, plated onto LB-agar plate supplemented with 50 µg/mL kanamycin and incubated overnight at 37 °C. A single colony was picked and used to inoculate 100 mL of LB media supplemented with 50 µg/mL kanamycin. The starter culture was grown at 37 °C with 220 rpm shaking overnight. 10 mL of the overnight starter culture was used to inoculate 1 L of LB media supplemented with 50 µg/mL kanamycin. The culture was incubated at 37 °C with shaking at 220 rpm. Cells grown to an OD₆₀₀ of ~0.6 were induced by addition of IPTG to a final concentration of 0.5 mM. Cell growth continued at 22 °C for 18 hours. Cells were harvested by centrifugation at 6,000 g for 10 minutes and stored at -80 °C until further use.

3.2.3. YejA purification

The whole cell pellet was resuspended in PBS buffer supplemented with 20 mM imidazole (pH 7.5), 3 mM MgCl₂, deoxyribonuclease (50 unit/mL), lysozyme (0.1 mg/mL) and PefaBloc (0.1 mg/mL). Cell-free lysate was prepared by sonication (10 cycles; 10-second sonication and 10-second interval per cycle) followed by centrifugation (200,000 g for 1 hour) and filtration (0.22 µm). Sonication was performed on ice and the following purification steps were performed at 4 °C. For the IMAC step, the supernatant was loaded on a His-Trap column pre-equilibrated with PBS containing 20 mM imidazole (pH 7.5). The column was washed with 20 CV of PBS containing a linear imidazole concentration from 20 mM to 250 mM (pH 7.5). The ÄKTAexpress (GE Healthcare) system was used to monitor the UV₂₈₀ absorbance of the flow through. The His₆-tagged YejA protein was eluted from PBS buffer containing 125 mM

imidazole. The purity of different fractions was assessed by SDS-PAGE. Fractions with high purity were collected and TEV protease was added with a protein-to-TEV ratio of 10:1 to cleave the His-tag. The sample was dialysed overnight in 2 L of buffer A (20 mM Tris (pH 7.5) and 150 mM NaCl).

The overnight cleaved sample was loaded on a His-Trap column pre-equilibrated with buffer A (reverse-IMAC). The flow through and 2 CV of buffer A wash were collected and concentrated to a final volume of 500 μ L, and injected on a Superdex 200 10/300 GL column, pre-equilibrated with buffer A. The ÄKTApurifier system was used to monitor the UV₂₈₀ absorbance of the fractions. Fractions containing YejA protein were collected and concentrated to 10 mg/mL using a 50 kDa MW cut-off concentrator (Thermo Fisher Scientific).

3.2.4. Isothermal titration calorimetry

ITC was performed to measure the binding of AMP to YejA (MicroCal iTC₂₀₀ microcalorimeter, Malvern Instruments). The ITC buffer consisted of 20 mM Tris (pH 7.5) and 150 mM NaCl. The protein and AMP were diluted to 20 μ M and 140 μ M, respectively, using the ITC buffer. The pH of protein and ligand was adjusted at 4 °C until identical. The measurement was carried out at 25 °C. 200 μ L of protein and 60 μ L of ligand were added into the sample cell and injection syringe, respectively. Seventeen 2.4 μ L injections were performed with a 0.5 μ L pre-injection, a 180-second interval and a 0.5 μ L/second injection speed. Control experiments were performed where ligand was injected into the ITC buffer, with the same experimental parameters. The heat of interaction was obtained by subtracting control measurements (heat of dilution) from the experimental measurements. The baseline was generated, and the

concentration-normalized binding isotherm was integrated using the MicroCal Origin software (version 7) (Origin, 2002). The first injection peak was removed from the isotherm.

3.2.5. YejA crystallisation

Crystallisation trials were carried out using the vapour diffusion sitting drop method on 96-well CrystalQuick plates (Grenier). Purified YejA at a concentration of 10 mg/mL (145 μ M) was screened against the JCSG-plus, PACT-premier and Index screening plates at 20 °C and 4 °C. The reservoir volume was 100 μ L and the sitting drop consisting of 100 nL of YejA and 100 nL of reservoir solution was dispensed by a Mosquito robot. YejA crystallised in the H3 condition of the JCSG-plus screening plate, which contained 0.1 M Bis-Tris (pH 5.5) and 25% (w/v) PEG 3350. Crystallisation conditions were optimised against different pH ($\pm$ 0.5), temperature (20 °C and 4 °C) as well as different protein (5 mg/mL, 10 mg/mL and 30 mg/mL) and PEG concentrations ($\pm$ 4% (w/v) with 2% (w/v) interval), on CrystalQuick plates. Microseeding was carried out using a contemporary method with the Seed Bead (Hampton Research). Briefly, following addition of 10 μ L of reservoir solution, the macrocrystals were crushed into microcrystals with the Crystal Crusher (Hampton Research) and transferred to the Seed Bead placed on ice. This step was repeated five times until all of the crushed crystals had been transferred to the Seed Bead. This resulted in $\sim$ 50 μ L of solution containing crushed crystals in the Seed Bead. The Seed Bead was vortexed on ice (three cycles; 30-second vortex plus 30-second interval per cycle). The sitting drop was made of 100 nL of YejA, 30 nL of the prepared microseed and 70 nL of reservoir solution.

3.2.6. Data collection and structure determination

Crystals were cryoprotected by transferring the crystals to the crystallisation buffer supplemented with 20% glycerol (v/v) and flash-cooled in liquid nitrogen for data collection. X-ray data were collected at a wavelength of 0.9686 Å on the I24 beamline at the Diamond Light Source. Data were indexed using Xia2 (Winter, 2010). YejA crystals belonging to the space group $P2_12_12_1$ diffracted X-rays up to 1.5 Å resolution. Initial phases were obtained by Phaser (McCoy *et al.*, 2007) molecular replacement using PDB 5ICQ (N-terminal domain) and 4ONY (C-terminal domain) as search models. An initial model was built by Buccaneer (Cowtan, 2006) autobuild. Iterative cycles of model building and refinement were completed using Refmac5 (Murshudov *et al.*, 2011) and COOT (Emsley & Cowtan, 2004). During refinement, strong electron density was observed at the interface of domain I and domain II, which enabled straightforward model building of the backbone of a polypeptide that consists of 11 amino acids (hendcapeptide). Water molecules were automatically added by ARP/wARP (Langer *et al.*, 2008) and further manually checked in COOT. The final YejA structure was refined at 1.5 Å with a R-factor of 15.6% and R-free of 21.3%. The final model has 97.45 % and 2.55% of the residues located in the favoured and allowed Ramachandran regions, respectively (Ramachandran *et al.*, 1963).

3.2.7. Sample preparation for MS

In order to separate the bound peptide from the YejA protein, MS specialists denatured YejA with urea (8M) and performed solid phase extraction using a C₁₈ ZipTip (Millipore) column. However, the following MS analysis showed that multiple YejA peptides were present in the YejA sample, which hindered identification of the bound peptide.

In order to separate abundant peptides from the crystallised protein-polypeptide complex, crystallisation desolvation was performed, followed by purification with a PD SpinTrap G-25 (GE Healthcare) column. YejA crystals were harvested, washed in protein buffer containing 20mM Tris (pH7.5) and 150 mM NaCl, and dissolved in 100 µL of protein buffer. To remove PEG molecules, the sample was loaded onto a PD SpinTrap G-25 column, pre-equilibrated with 300 µL of protein buffer. 100 µL of dissolved YejA crystals were applied in the middle of the packed bed of the PD SpinTrap G-25 column. The sample was eluted by centrifugation (800 g for 2 minutes). To remove all the sample, 70 µL of protein buffer was loaded onto the column, followed by 2 minutes of centrifugation at 800 g. This procedure was repeated once. In total three fractions containing 5 µg/mL, 12 µg/mL and 17 µg/mL of proteins were obtained. All three fractions were sent for MS analysis.

MS experiments by matrix-assisted laser desorption ionization (MALDI) MS, MALDI tandem mass spectrometry (MS/MS), Nano-scale liquid chromatography (nLC)-electrospray ionization (ESI) MS/MS were performed by the Mass Spectrometry and Proteomics Facility at the University of St Andrews. Crystal solution and PD SpinTrap purification were performed by Feng Qu.

3.3. Results

3.3.1. YejA purification

The sequence of YejA in *E. coli* possesses a 19-residue signal peptide. The signal peptide was removed to ensure expression of the recombinant protein in the cytoplasm. Following the IMAC and reverse-IMAC steps, SEC was performed using the Superdex 200 10/300 GL column. Mature YejA (67.521 kDa) was eluted at ~15 mL as a sharp monodisperse peak (Figure 3.3.1.1A), which indicated the efficient recovery of a single YejA species. By comparison, a protein standard of bovine serum albumin (BSA) (67 kDa) elutes at a volume of ~14 mL according to the column data specification, suggesting that mature YejA was purified as a monomer. The single band from SDS-PAGE indicates that pure YejA was obtained (Figure 3.3.1.1B). The protein band for YejA ran faster than its theoretical MW, which might be due to the higher SDS binding capacity of YejA.

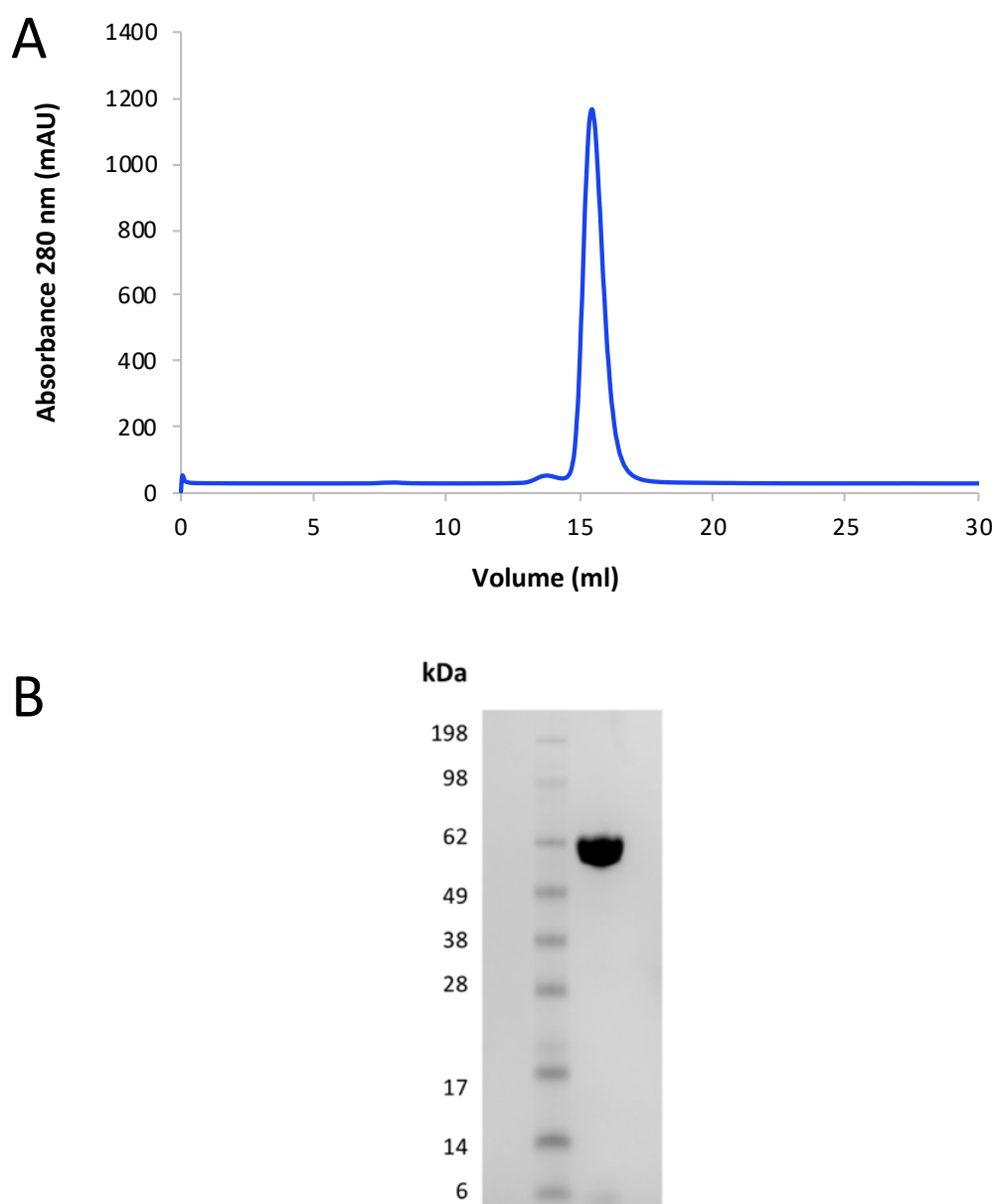


Figure 3.3.1.1. Purification of YejA. (A) SEC profile for YejA using the Superdex 200 10/300 GL column. Mature YejA was purified as a monomer as indicated by the sharp monodisperse peak. (B) SDS-PAGE reveals a single band at ~62 kDa corresponding to mature YejA.

3.3.2. Binding of AMP to YejA

MccC7/C51 is an N-formylated heptapeptide that contains a modified AMP covalently attached to its C-terminal aspartate through an N-acyl phosphoramidate bond (Figure 3.1.1.1A) (Guijarro *et al.*, 1995). Full-length MccC7/C51 hijacks the outer-membrane porin OmpF and the inner-membrane ABC importer YejABEF to enter the cytoplasm of sensitive bacteria (Novikova *et al.*, 2007). In the host cytoplasm, MccC7/C51 undergoes a series of cleavage processes and releases the processed MccC7/C51* (Figure 3.1.1.1B), a translation inhibitor (Kazakov *et al.*, 2008). The processed MccC7/C51* is not recognised by the YejABEF system (Novikova *et al.*, 2007).

Since mature MccC7/C51 was not available, the binding affinity between the purified YejA and AMP, the C-terminal nucleotide of MccC7/C51, was initially studied. Binding between YejA and AMP was not observed by ITC (Figure 3.3.2.1). It was firstly explained by the fact that YejA probably only requires the recognition of the peptide chain (Met'-Arg'-Thr'-Gly'-Asn'-Ala'-Asn') for MccC7/C51 uptake, whilst it does not recognise or bind AMP. Later, the high-resolution crystal structure revealed that a polypeptide occupied the active site of YejA (Figure 3.3.4.2), which might explain the lack of binding affinity shown by ITC (Figure 3.3.2.1). It is likely that the polypeptide binds to apo YejA with a higher affinity than AMP. Therefore, AMP could not outcompete the hendecapeptide that occupied the substrate-binding site of YejA during protein purification and crystallisation. The structure of YejA in complex with MccC7/C51 is needed to unravel the binding mode between YejA and MccC7/C51.

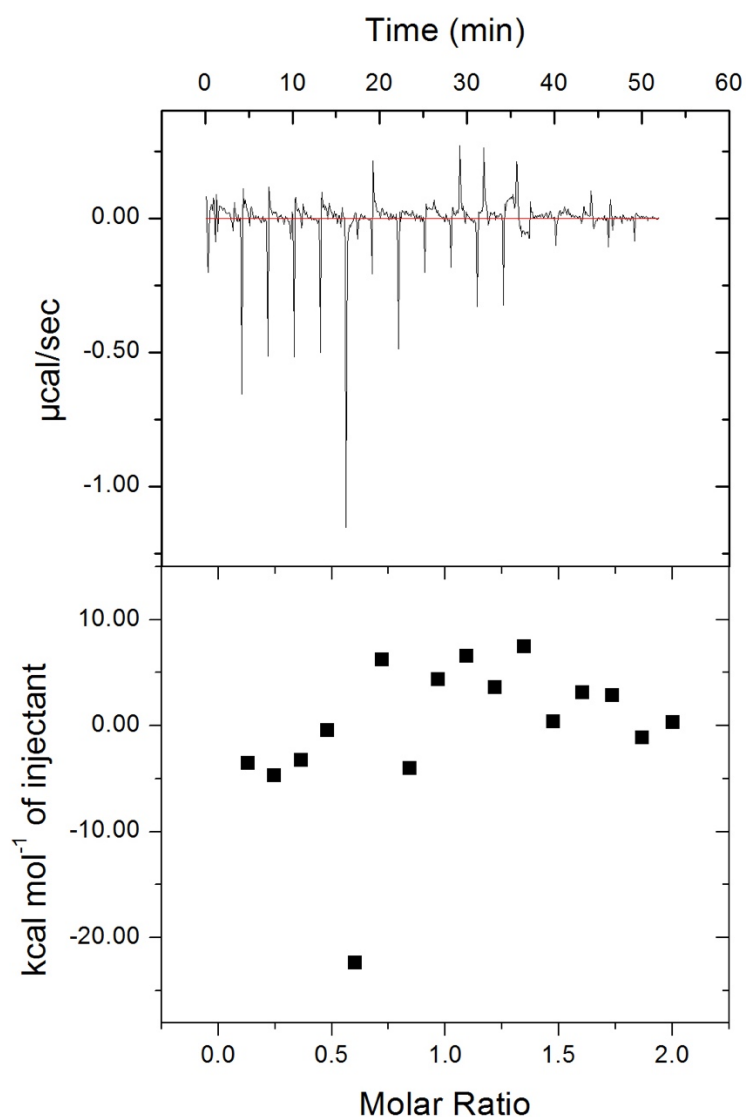


Figure 3.3.2.1. Calorimetric titration for the binding of AMP to YejA. Each peak (top panel) represents an injection of 2.4 μL of AMP (20 μM) into 200 μL of YejA at 25 $^{\circ}\text{C}$. The bottom panel shows the integrated data obtained from the raw data, after subtracting the heat associated with dilution of the ligand. Binding between YejA and AMP is not observed.

3.3.3. YejA crystallisation

YejA (10 mg/mL) crystallised in the H3 condition in the JCSG-plus screening plate, which contained 0.1 M Bis-Tris (pH 5.5) and 25% (w/v) PEG 3350. These crystals appeared as clusters of needles after 20 °C incubation for a week and diffracted X-rays to ~ 2.5 Å. Crystal conditions were further optimized in the 96-well sitting drop CrystalQuick plates. 5mg/ml of YejA in reservoir containing 0.1 M Bis-Tris (pH 5.5) and 23% (w/v) PEG 3350 produced more separated crystals. Microseeding was performed in parallel using the initial YejA crystals as seeds. Seeding accelerated the formation of YejA crystals. YejA crystals with seeding appeared after 4 °C overnight incubation (Figure 3.3.3.1). Crystallisation optimisation and microseeding improved the crystal quality to 1.5 Å (Figure 3.3.4.1 and Table 3.3.4.1).

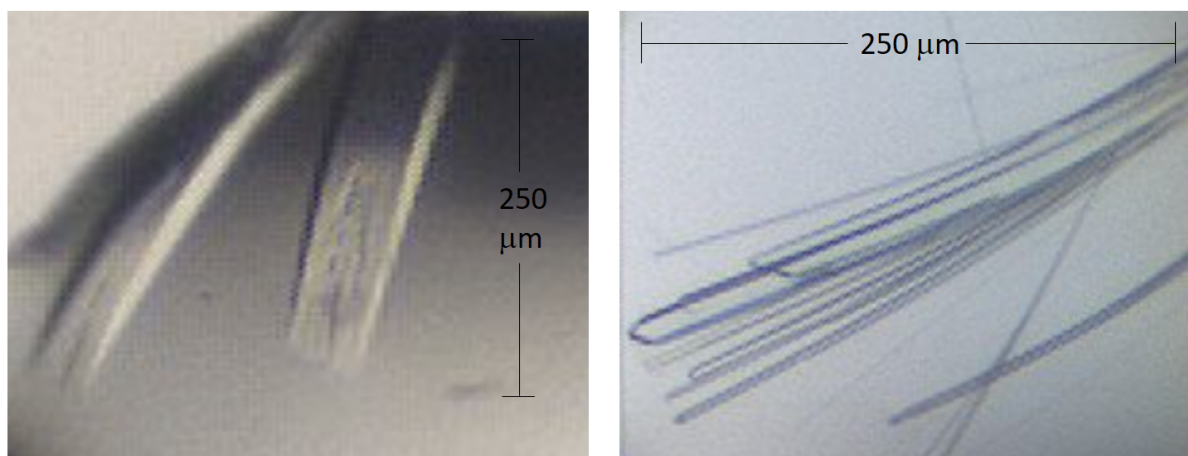


Figure 3.3.3.1. YejA crystals. Crystallisation condition: 0.1 M Bis-Tris (pH 5.5) and 23% (w/v) PEG 3350 with seeding at 4 °C.

3.3.4. Overall structure of YejA

The structure of YejA from *E. coli* was refined at a resolution of 1.5 Å (Figure 3.3.4.1). The space group was determined to be $P2_12_12_1$ with one copy of YejA in the asymmetric unit. Residues 21-601 were observed in electron density maps (Figure 3.3.4.2). SBPs usually exhibit high binding affinity to their ligands (nM to μ M range) (Scheepers *et al.*, 2016). Consequently, ligands are often co-purified and co-crystallised with their SBPs. During refinement, strong electron density was observed at the interface of domain I and domain II (Figure 3.3.4.3A), which could not be accounted for by the YejA protein. The electron density enabled straightforward model building of the backbone of a hendecapeptide (Figure 3.3.4.3B). Such co-crystallisation has also been described for AppA from *B. subtilis*, and OppA from *L. lactis* and *E. coli* (Berntsson *et al.*, 2009; Levnikov *et al.*, 2005b; Klepsch *et al.*, 2011). Data collection and refinement statistics are summarised in Table 3.3.4.1.

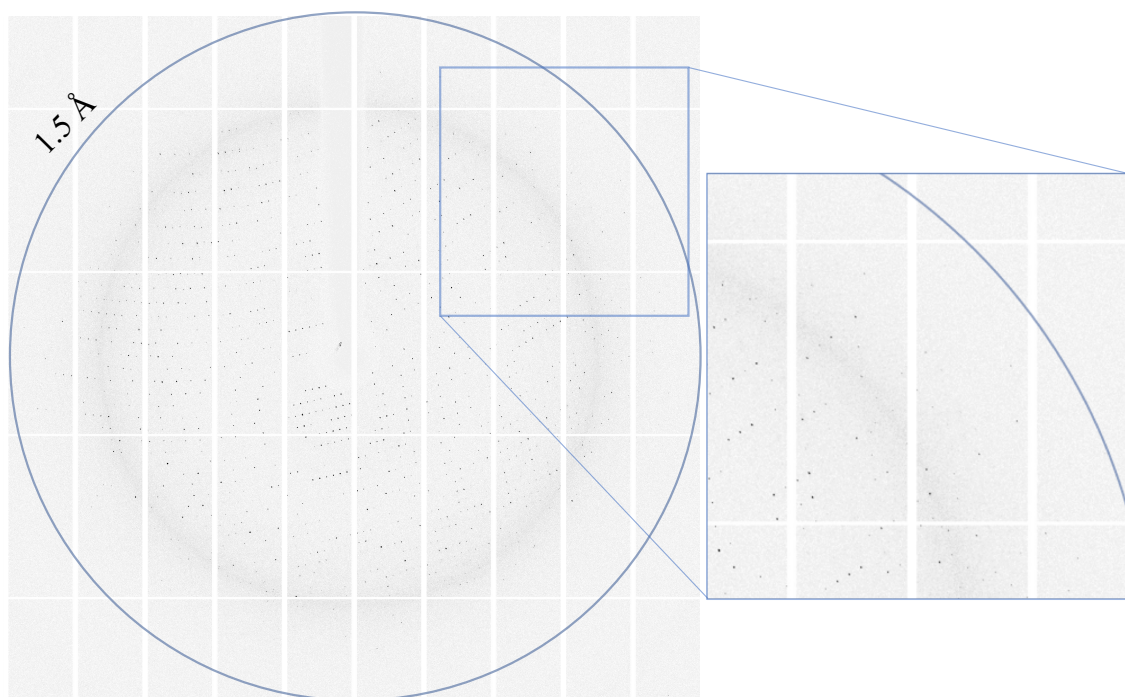


Figure 3.3.4.1. X-ray diffraction pattern for YejA crystal at 1.5 Å. The blue circle shows the diffraction limit at 1.5 Å and the inset indicates diffraction spots at high resolution.

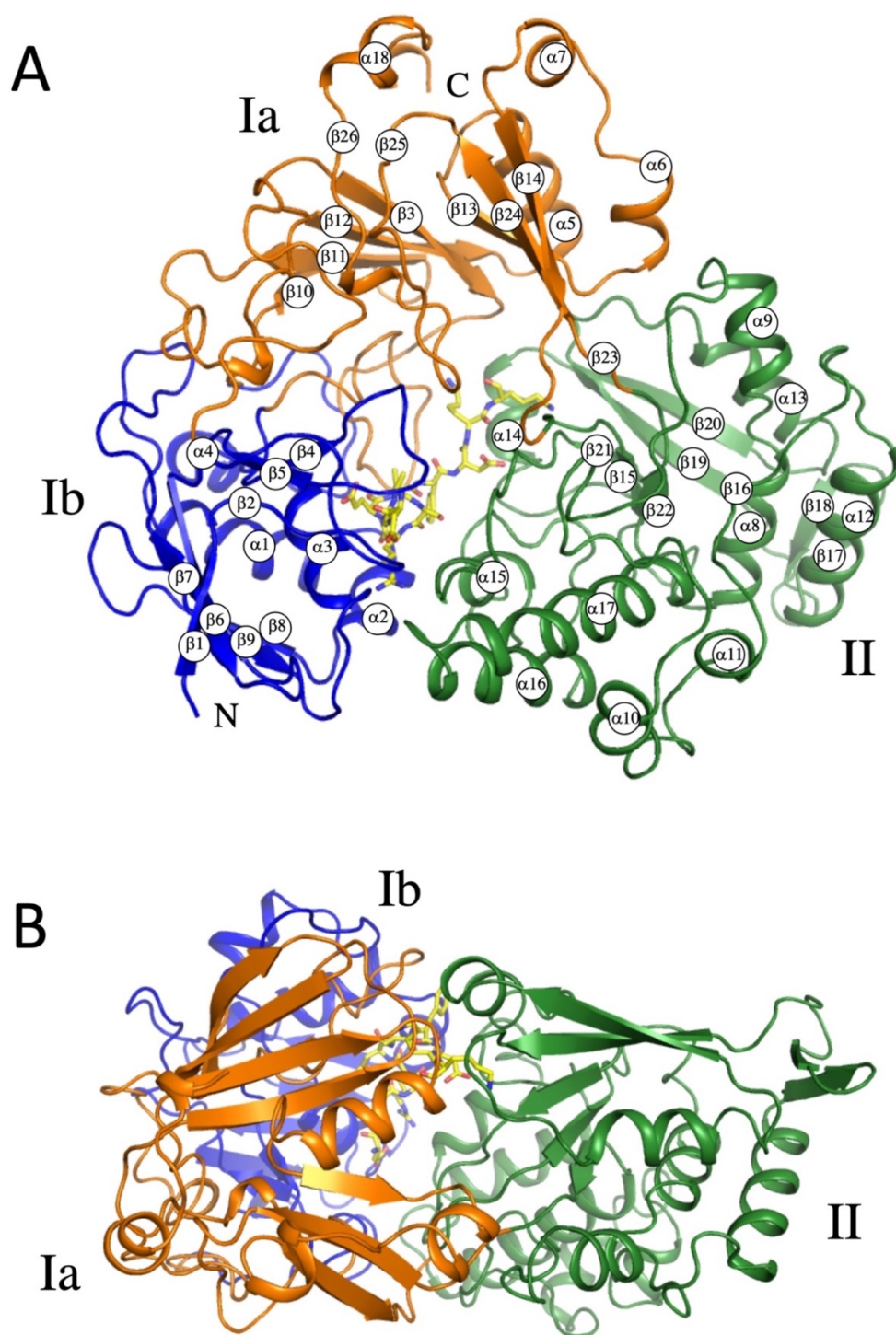


Figure 3.3.4.2. Structure of YejA in complex with an oligopeptide. Front-view (A) and side-view (B) of the YejA-peptide structure. Domain Ia, domain Ib and domain II are coloured in orange, blue and green, respectively. The bound peptide is shown in stick format, with carbon, oxygen and nitrogen atoms coloured in yellow, red and blue, respectively. Secondary structure elements are indicated: α -helices (α) and β -strands (β).

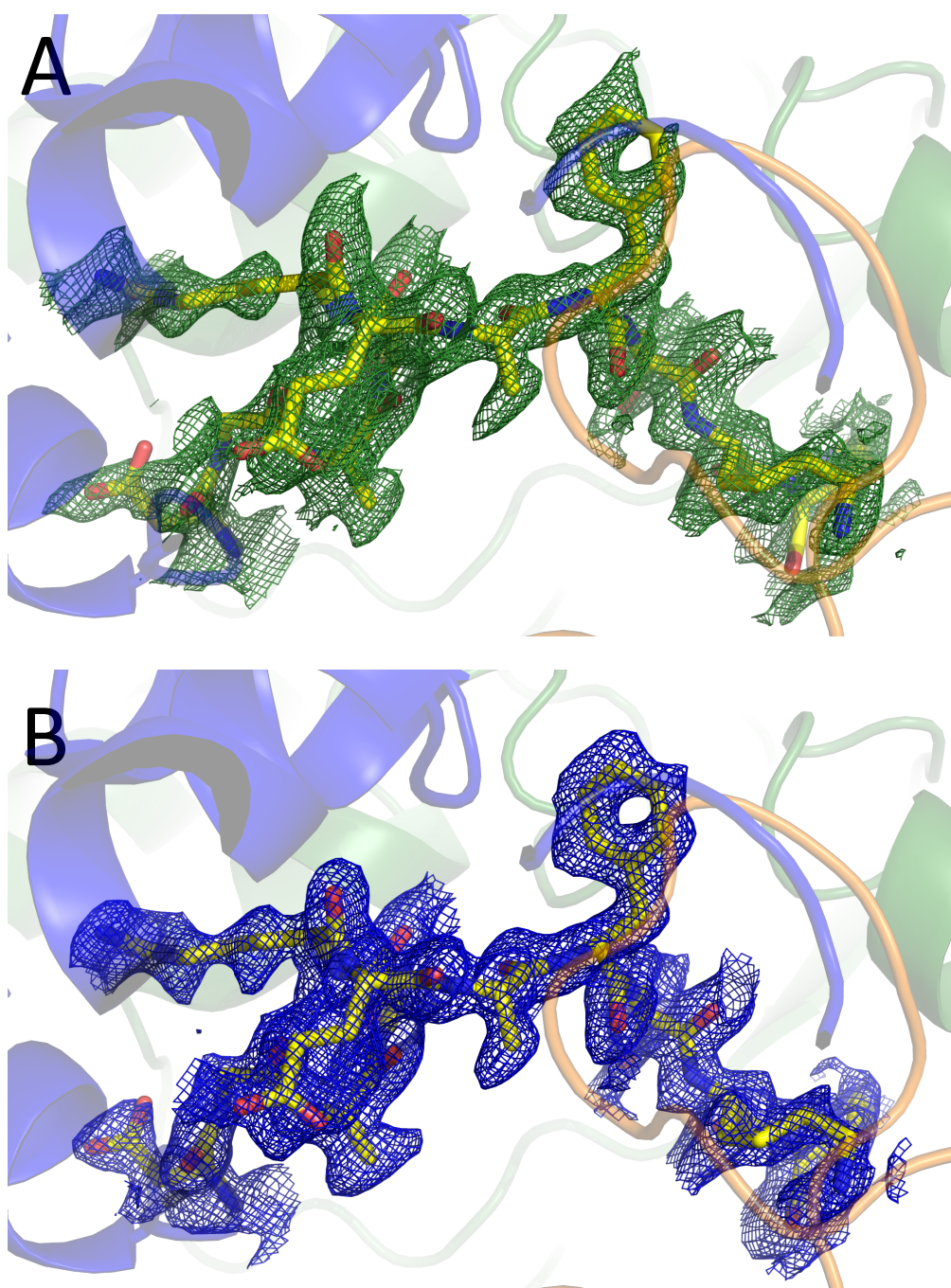


Figure 3.3.4.3. Electron density maps for the co-crystallised oligopeptide. (A) The Fo-Fc omit maps of the peptide contoured at 3σ (carve=3 Å) after omitting the peptide from refinement (green mesh). (B) The 2Fo-Fc electron density maps of the hendecapeptide contoured at 1σ (carve=3 Å) after refinement (blue mesh). The peptide is shown in stick format with carbon, oxygen and nitrogen atoms coloured in yellow, red and blue, respectively.

Data collection	
Beamline	DLS-I24
Wavelength (Å)	0.9686
Resolution (Å)	1.5 (62.19)
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Cell dimensions	
a,b,c (Å)	70.87, 77.39, 104.45
α, β, γ (°)	90, 90, 90
NO. of total reflections	417657 (20677)
NO. of unique reflections	130542 (6414)
Completeness (%)	97.5 (96.9)
Redundancy	3.2 (3.2)
R _{meas}	0.068 (0.527)
Mean (I/sig(I))	11 (2.5)
Mn(I) half-set correlation CC 1/2	1 (0.8)
Refinement	
R _{work} /R _{free} (%)	15.6/ 21.3
Average B-factors (Å ²)	25.739
NO. of non-hydrogen atoms	5412
RMS deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.83
Ramachandran	
Favoured (%)	97.45
Allowed (%)	2.55
Outliers (%)	0

Table 3.3.4.1. Data collection and refinement statistics for YejA crystals. The R-free value was calculated by excluding 5% of data from refinement. Values in parentheses are for the high-resolution shells.

YejA is composed of two domains, with domain I containing residues 21-304 and 558-601, and domain II consisting of residues 305-557 (Figure 3.3.4.2). The two domains are connected by a hinge region composed of two segments (residues 300-309 and 555-561). A substrate-binding site is buried in between the domain interface, characteristic of a closed-state SBP. Domain I can be divided into two subdomains, domain Ia and domain Ib. Domain Ia consists of residues 39-84, 209-304 and 558-601, forming a central β -sheet ($\beta_{10}\beta_{11}\beta_{12}\beta_3\beta_{13}\beta_{24}\beta_{14}$) flanked by a number of α -helices. Domain Ib encompasses residues 21-38 and 85-208 and is composed of a central five-stranded β -sheet ($\beta_8\beta_9\beta_7\beta_6\beta_1$), a β -hairpin ($\beta_3\beta_4$) and flanking α -helices. Domain II also contains a core five-stranded β -sheet ($\beta_{22}\beta_{15}\beta_{21}\beta_{19}\beta_{20}$), a β -hairpin ($\beta_{17}\beta_{18}$) and a number of flanking α -helices. Based on the secondary structure arrangement, YejA is classified as a type II SBP belonging to cluster C (Berntsson *et al.*, 2010). Compared to SBPs of other clusters, SBPs in cluster C are larger in size and possess an additional domain (*i.e.* domain Ib). Ligands of other cluster C SBPs are typically enclosed by domain Ia and domain II (Figure 3.3.4.4) (Levdikov *et al.*, 2005; Berntsson *et al.*, 2009), whereas in the YejA-peptide structure, the bound peptide is located at the interface of domains Ia, Ib and II (Figure 3.3.4.2).

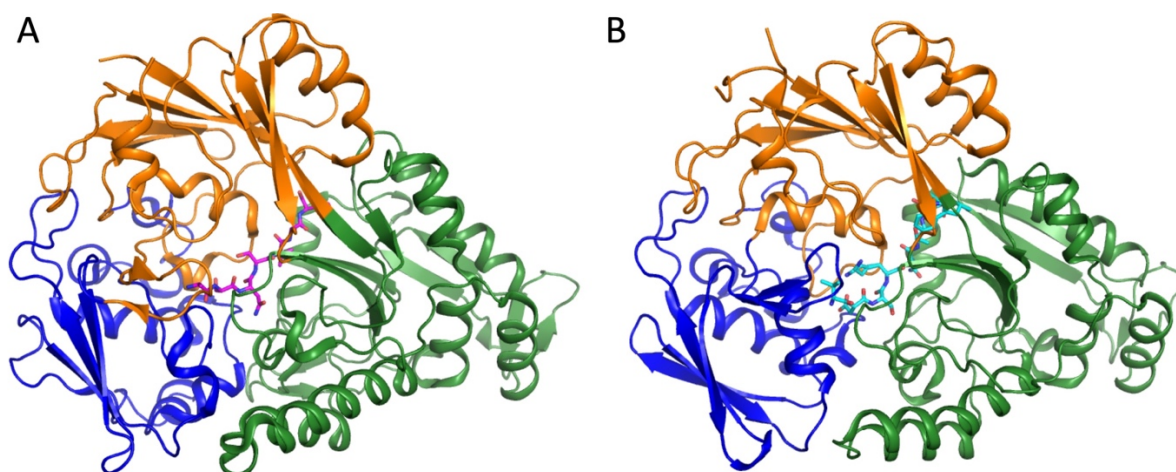


Figure 3.3.4.4. Top view of AppA and OppA. (A) AppA from *B. subtilis* (PDB ID: 1XOC, (Levdikov *et al.*, 2005)). (B) OppA from *L. lactis* (PDB ID: 3DRF, (Berntsson *et al.*, 2009)). In AppA and OppA, domain Ia and domain II are linked by a two-segment hinge region and form an interface that encloses the ligand. Domain Ia, domain Ib and domain II are coloured in orange, blue and green, respectively. The bound peptides inside the cavity of AppA and OppA are shown in magenta and cyan stick format, respectively. Oxygen and nitrogen atoms are coloured in red and blue, respectively.

A DALI search (Holm & Sander, 1995) revealed that the structure of *E. coli* YejA most closely matches the apo CuMbn-binding protein MbnE from *Methylocystis parvus* OBBP (Figure 3.3.4.5A) (PDB ID: 5ICQ, (Dassama *et al.*, 2016)) (Z-score=37, RMSD=3.8 Å over 552 Cα atoms). CuMbns are copper (Cu)-loaded methanobactins (Mbns), small molecules produced by methanotrophs for Cu acquisition (DiSpirito *et al.*, 1998). The oligopeptide-binding protein AppA from *B. subtilis* (Figure 3.3.4.5B) (PDB ID: 1XOC, (Levdikov *et al.*, 2005)) is the next highest-scoring match (Z-score=28.3, RMSD=2.7 Å over 478 Cα atoms). Although containing similar overall folds, the oligopeptide-binding protein OppA from *L. lactis* (PDB ID: 3DRF, (Berntsson *et al.*, 2009)) only gives a Z-score of 21.6 (Figure 3.3.4.5C). The differences in structural similarity could be anticipated from the sequences, with MbnE, AppA and OppA having 31.9% (49.1%), 22.1% (35.6%) and 19.1% (33.4%) sequence identity

to YejA, respectively (Figure 3.3.4.6). Structural comparison shows that the peptide-bound YejA, AppA and OppA are in a closed state, whereas the apo MbnE displays an open conformation (Figure 3.3.4.2 and Figure 3.3.4.5). The overall structures of ligand-bound oligopeptide-binding proteins appear identical, regardless of the sequences of the bound peptides, and are distinct from the apo protein. This observation has validated the “Venus fly-trap” mechanism for SBPs, where ligand binding stabilises the closed conformation (Figure 1.0.5.6) (Berntsson *et al.*, 2010).

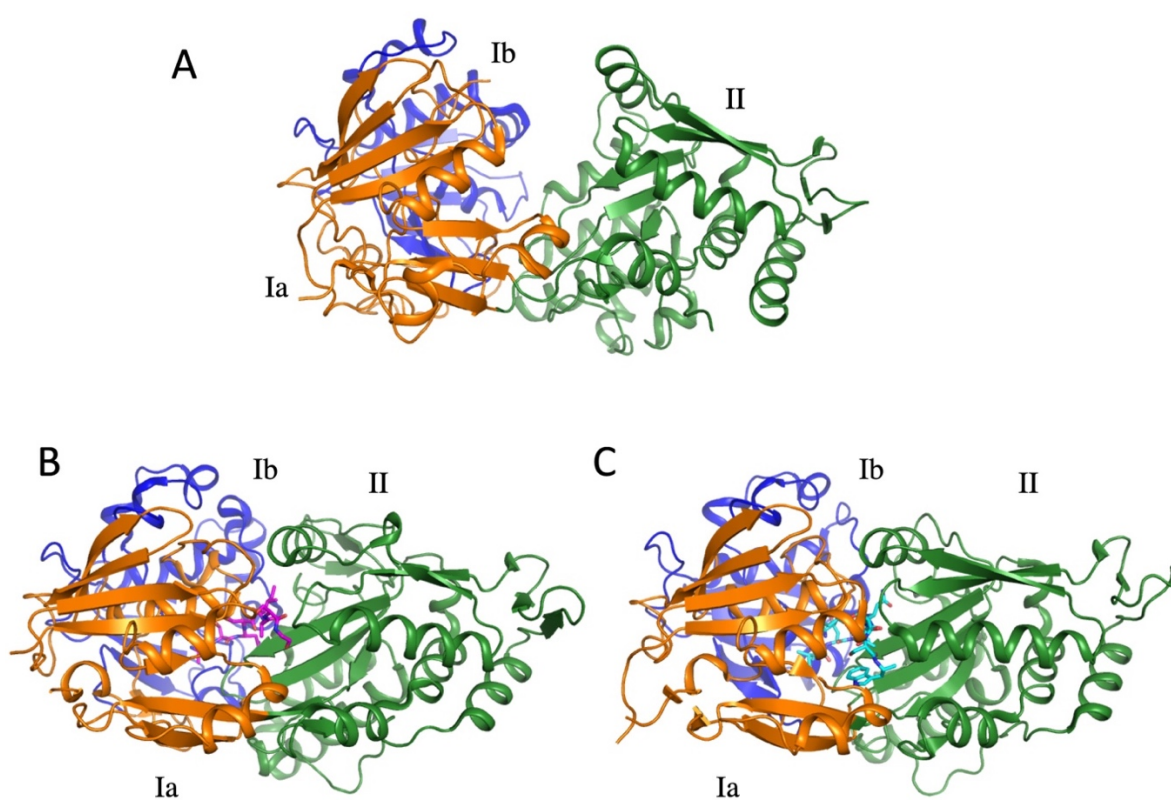


Figure 3.3.4.5. Structures of YejA homologues. Side views of MbnE from *M. parvus* OBBP (PDB ID: 5ICQ, (Dassama *et al.*, 2016)) (A), AppA from *B. subtilis* (PDB ID: 1XOC, (Levdikov *et al.*, 2005)) (B) and OppA from *L. lactis* (PDB ID: 3DRF, (Berntsson *et al.*, 2009)) (C). Domain Ia, Ib and domain II are coloured in orange, blue and green, respectively. The bound peptides inside the cavity of AppA and OppA are shown in magenta and cyan stick format, respectively. Oxygen and nitrogen atoms are coloured in red and blue, respectively.

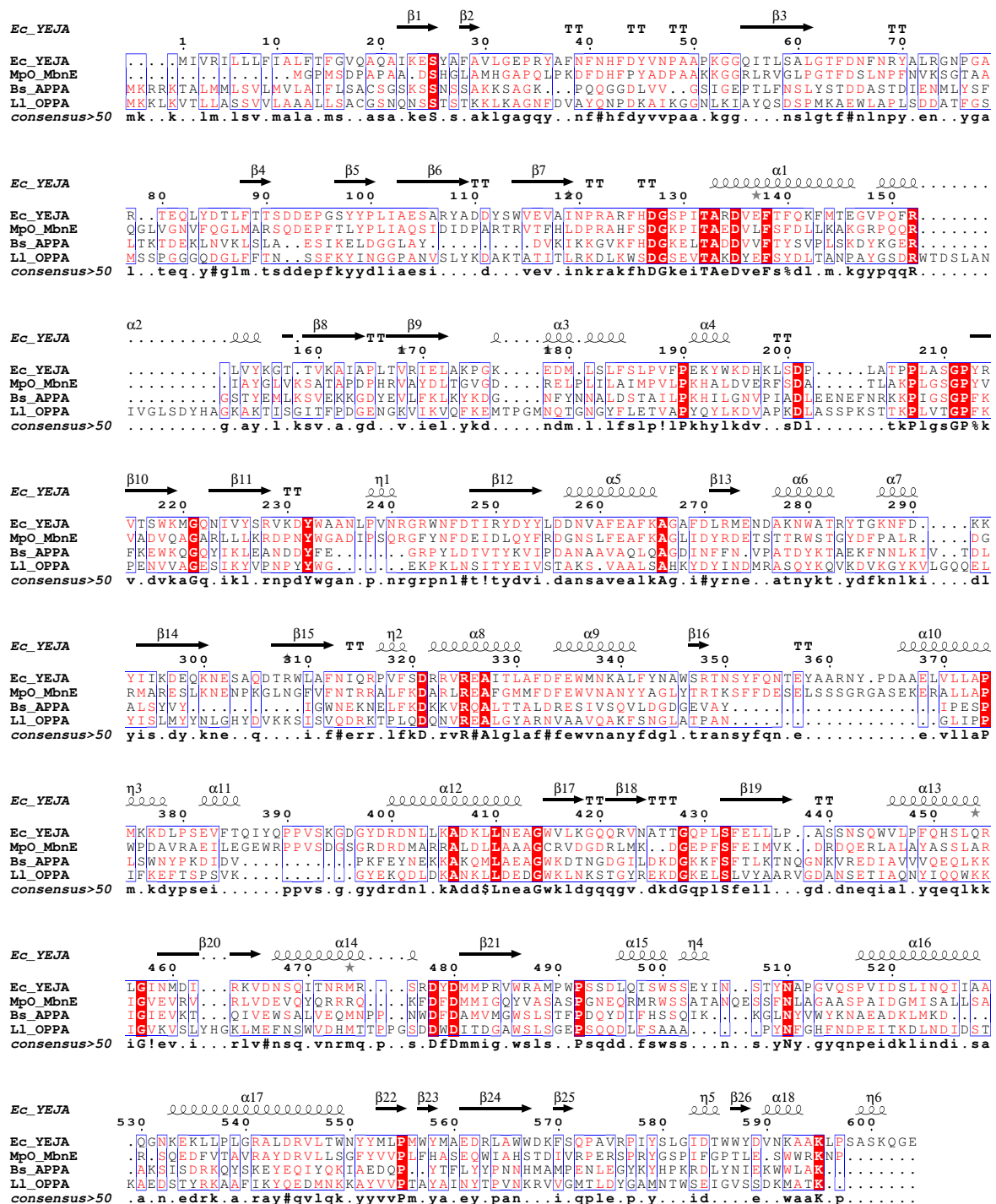


Figure 3.3.4.6. Sequence alignment of YejA from *E. coli* and its homologues. YejA from *E. coli* (Ec_YejA), MbN from *M. parvus* OBBP (MpO_MbN), AppA from *B. subtilis* (Bs_AppA) and OppA from *L. lactis* (Ll_OppA). Conserved and similar residues are shown in red and blue boxes, respectively. Residue numbers and secondary structure elements of YejA are indicated at the top of the protein sequence. The sequence alignment was performed by MultAlin (Corpet, 1988) and the figure was prepared by ESPrpt 3 (Robert & Gouet, 2014).

3.3.5. YejA in complex with a polypeptide

YejA is the SBP component of the ABC importer YejABEF. YejABEF is hijacked by MccC7/C51 to enter sensitive bacterial cells (Novikova *et al.*, 2007). Based on sequence similarity, YejABEF is probably involved in oligopeptide internalisation and YejA is probably an oligopeptide-binding protein (Vondenhoff *et al.*, 2011). The physiological substrate of YejA is yet to be confirmed. The presence of a hendecapeptide in the obtained crystal structure supports the notion that YejA is an oligopeptide-binding protein.

During refinement, additional electron density was observed in the substrate-binding pocket of YejA (Figure 3.3.4.3). The strong electron density allowed for straightforward model building of a hendecapeptide backbone. Electron density for peptide side chains was more ambiguous. This indicates that a population of different peptides might have bound to the active site, resulting in an averaged electron density and low occupancy for the peptide side chains. Assignment of residue side chains was more challenging. The electron density for side chains of residues at positions 1, 4, 6, 9, 10 and 11 are ambiguous and could not be readily interpreted. The electron density at positions 2, 3 and 7 was small, suggesting the presence of residues with short side chains. Therefore, glycine, alanine and alanine were assigned to positions 2, 3 and 7, respectively. However, the electron density at positions 3 and 7 can also be attributed by disordered side chains of bulkier residues such as lysine. Based on the density, side chains of residues at positions 5 and 8 were bulky. Specifically, we could observe the density for a benzene ring at position 8. Arg5' and Phe8' were assigned to the peptide accordingly. The hendecapeptide has been modelled as Asp'-Gly'-Ala'-Ser'-Arg'-Glu'-Ala'-Phe'-Asp'-Lys'-Lys' (where underlines indicate interpretation with higher confidence) (Figure 3.3.5.1). The current peptide sequence is derived on the basis of density of the peptide, potential interactions with the environment and iterative cycles of model building and refinement. The YejA-peptide

structure also showed some weak electron density protruding from the C-terminal end of the modelled hendecapeptide, suggesting that longer peptides might be present in some proteins in the crystal. However, the density was not well defined thus limiting assignment of additional residues. Assignment of the underlined residues indicates that YejA from *E. coli* may have a preference for peptides with such residues.

Interactions within the bound peptide stabilise the peptide conformation (Figure 3.3.5.1A). The extensive interactions between Gly2'/Ala3' and Arg5'/Glu6' form two 90° turns in the peptide backbone (Figure 3.3.5.1A). The backbone amide groups of Gly2', Glu8', Asp9' and Lys10' interact with the side chains of Asp1', Ala7', Phe8' and Asp9', respectively (Figure 3.3.5.1A, blue circles). A van der Waals contact formed by the backbone carboxyl group of Gly2' has stabilised the aliphatic chain of Arg5' (Figure 3.3.5.1A, red circles). Interactions between the bound peptide and YejA/solvent contribute to the coordination of the peptide in the substrate-binding pocket of YejA (Figure 3.3.5.1B). Binding of the hendecapeptide to YejA is almost exclusively contributed by hydrogen bonds between the peptide backbone and the protein, apart from the side chains of Arg5' and Phe8'. Carbonyl groups of Ser4', Ala7', Phe8', Asp9' form hydrogen bonds with the carbonyl nitrogen of Gln150, the hydroxyl group of Tyr509, and the guanidinium groups of Arg69 and Arg475, respectively. The amide groups of Gly2' and Asp9' are hydrogen bonded to the carboxylate group of Asp495 and the hydroxyl group of Thr508, respectively. Carbonyl groups of Asp1' and Glu6' and the amide group of Phe8' are hydrogen bonded to three different water molecules. The side chains of Arg5' and Phe8' form hydrogen bonds and van der Waals interactions with YejA. Three guanidinium nitrogen atoms of Arg5' form four hydrogen bonds with the carboxylate groups of Asp179 and Asp495 and the hydroxyl group of Ser494. The benzene ring of Phe8' forms extensive van der Waals interactions with Leu72, Val148, Gln150 and Ser507. Side chains of peptide residues at positions 1, 9 and 10 might have formed polar contacts with Glu178, Ser182 and Arg484 and

Asn75, but it was not possible to distinguish the exact interactions due to the unclear definition of the residues.

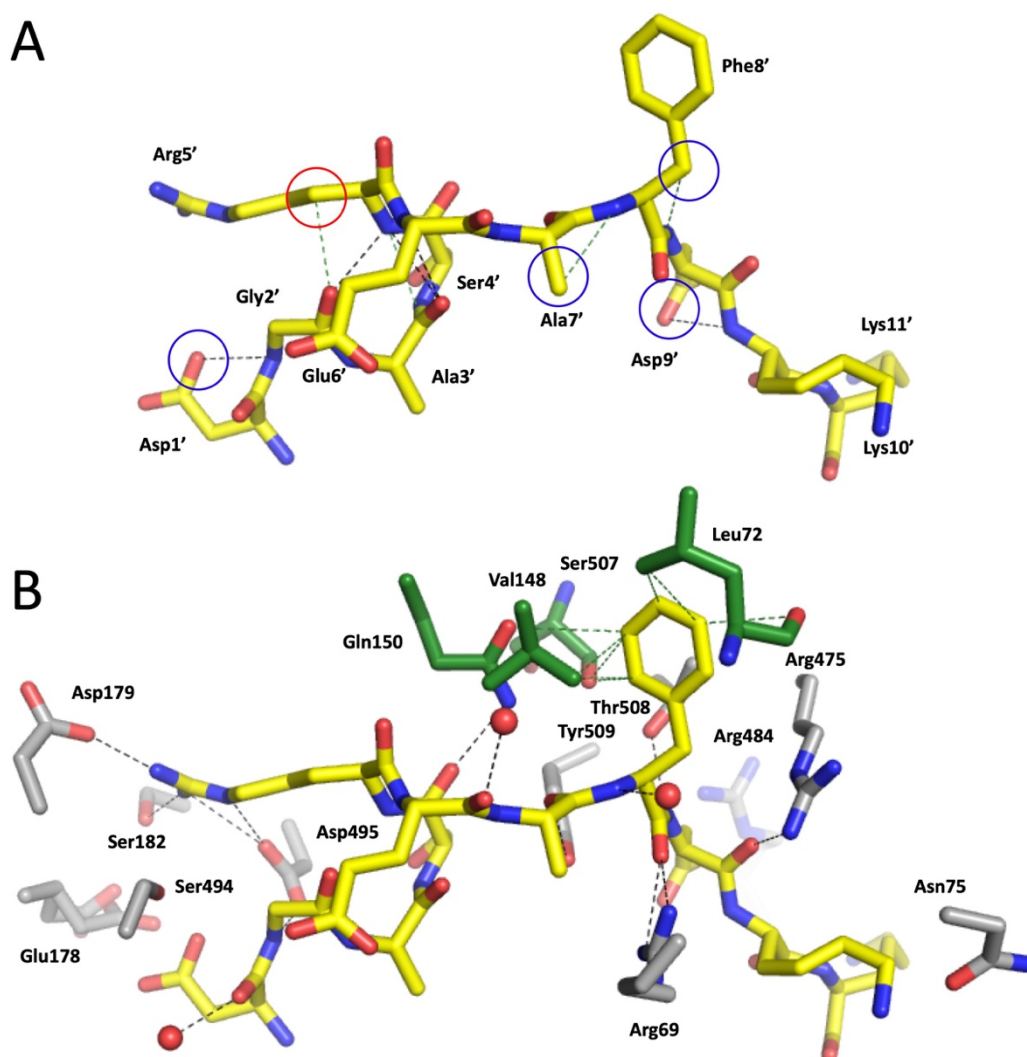


Figure 3.3.5.1. Peptide binding to YejA. (A) Interactions within the polypeptide. Side chains that form critical interactions with backbone amide and carboxyl groups are indicated by blue and red circles, respectively. (B) Interactions between the bound peptide and YejA/solvent. Carbon atoms of residues that form van der Waals contact with Phe8' are coloured in green. The bound peptide and residues that interact with the peptide are shown in stick format (yellow and grey, respectively). Oxygen and nitrogen atoms are coloured in red and blue, respectively. Water molecules are shown as red spheres. Interactions between peptide residues Gly2', Ala3', Arg5', Ala7' and Phe8', and protein residues are shown as dotted lines, with hydrogen bonds and van der Waals interactions coloured in black and green, respectively.

3.3.6. Characterisation of the bound peptide

MALDI and nLC-ESI MS have identified an abundance of peptides that are present in the YejA sample (Figure 3.3.6.1 and Figure 3.3.6.4). Some of the peptides can be seen in both the MALDI and nLC-ESI mass spectra, for instance the ion signal at a mass-to-charge ratio (m/z) of 1426.7 in MALDI MS (Figure 3.3.6.1) that corresponds to the signal at 714 m/z (2^+) in nLC-ESI MS (Figure 3.3.6.4). Mass spectra of BSA/angiotensin II are a lot clearer. The masses at 1046 m/z dominate the spectra in MALDI MS (Figure 3.3.6.2), which accounts for peaks at 349 m/z (3^+) and 523 m/z (2^+) in nLC-ESI MS (Figure 3.3.6.5). Other peaks observed in the MALDI mass spectra such as peptides at 1051.7 m/z and 1165.7 m/z (Figure 3.3.6.1) are also seen in the BSA/angiotensin II control (Figure 3.3.6.2), suggesting that they are likely to be matrix adduct peaks. In fact, this is confirmed by comparison with the spectra of the α -cyano-4-hydroxycinnamic acid matrix alone (Figure 3.3.6.3). It is notable that a large number of peaks are observed in the mass spectra of the matrix (Figure 3.3.6.3). The matrix α -cyano-4-hydroxycinnamic acid matrix is a small compound (MW: 189.17 kDa), however it can form clusters which form a major source of background interference (Keller & Li, 2000). Formation of matrix-clusters is dependent on many experimental factors including the type of matrix used, pH, the salt content, the amount of analyte and the laser power. The signal at 1449 m/z is unique to the fraction eluted by 10% of acetonitrile but is also seen in the matrix control. The better performance of the BSA/angiotensin II control is possibly owing to angiotensin II being a known peptide in the database and partially the peptide being abundant with a low background of other peptides.

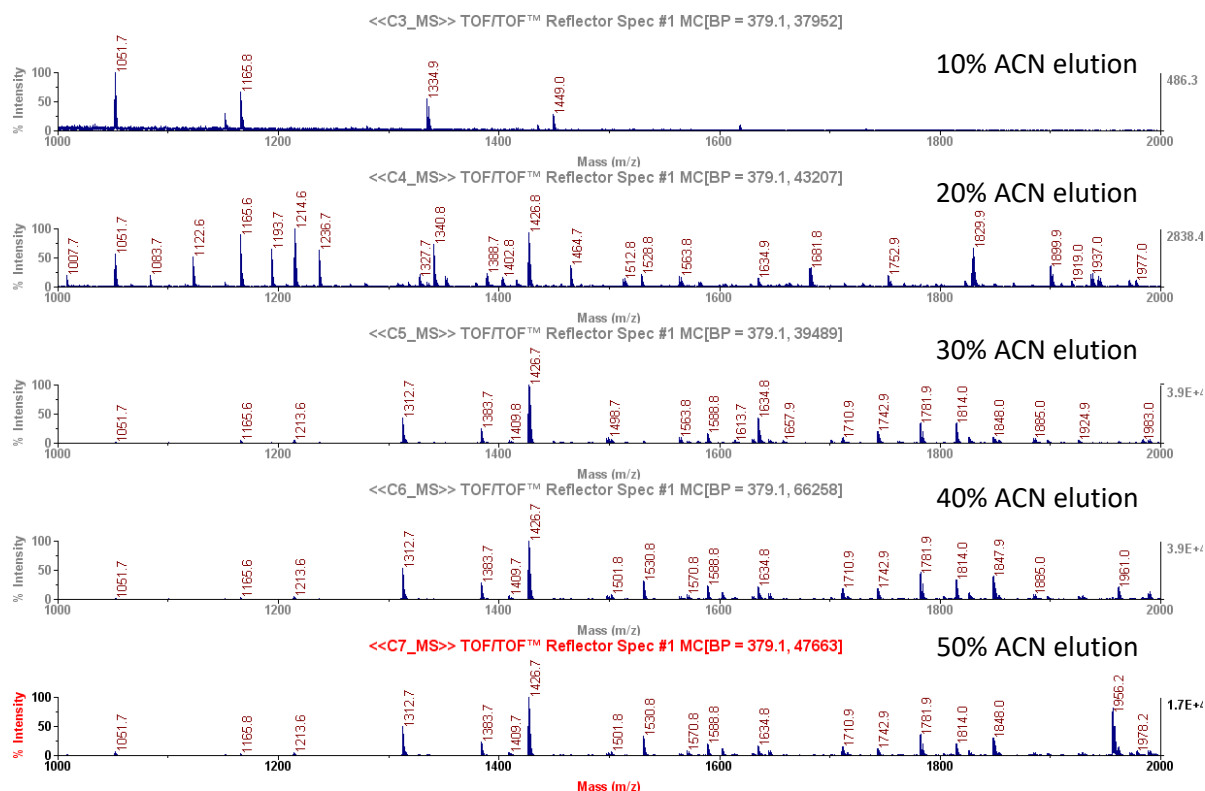


Figure 3.3.6.1. The MALDI mass spectra of the YejA protein mixture. The sample was denatured with urea and loaded onto a C₁₈ ZipTip column. The spectra were obtained from eluates of the C₁₈ ZipTip with 10-50% of acetonitrile. The sample was mixed with the matrix and spotted on the plate. The sample was allowed to dry prior to MALDI MS analysis. The chromatogram ranges from 1000-2000 m/z. The most right-side peak indicates the 100% intensity.

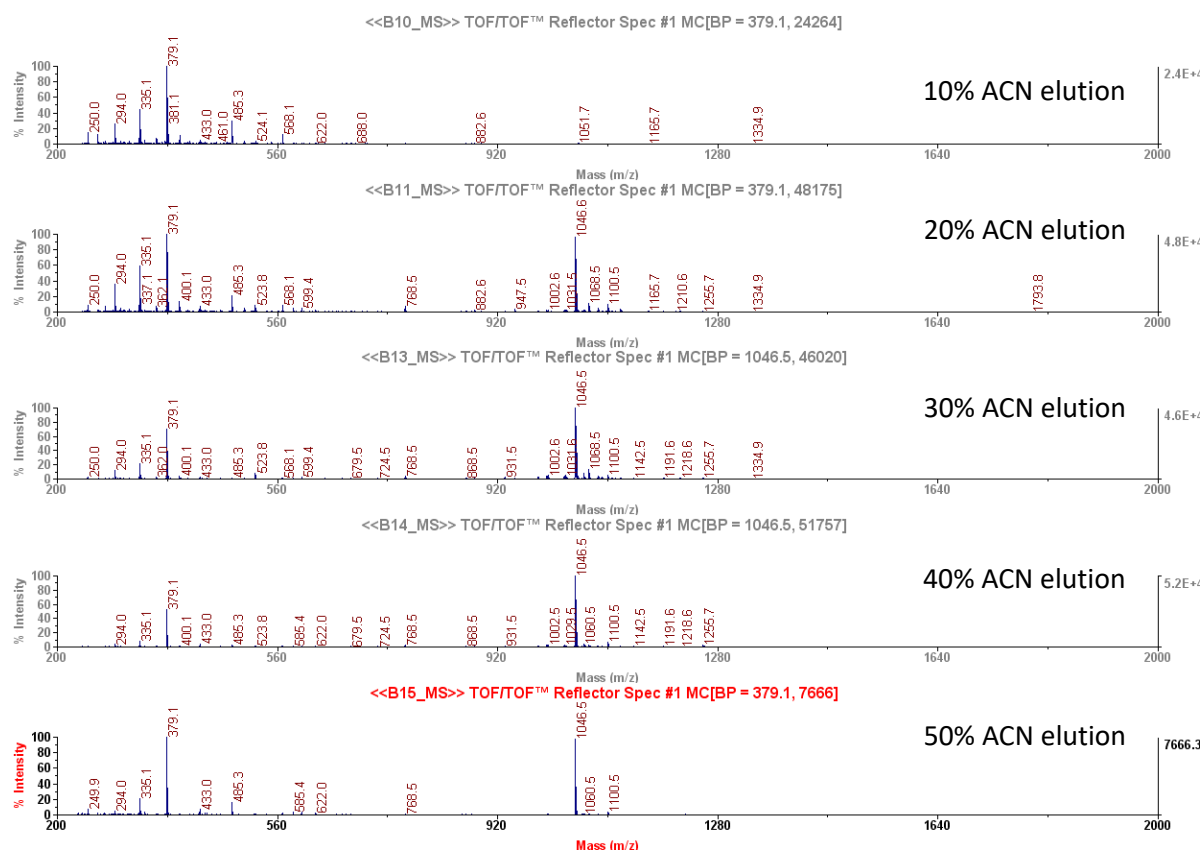


Figure 3.3.6.2. The MALDI mass spectra of BSA and angiotensin II. BSA and angiotensin II (1:1 ratio) were denatured with urea and loaded onto a C₁₈ ZipTip column. The spectra were obtained from eluates of the ZipTip with 10-50% of acetonitrile. The sample was mixed with the matrix, spotted on the plate and allowed to dry prior to MALDI MS analysis. The chromatogram ranges from 200-2000 m/z. The most right-side peak indicates the 100% intensity.

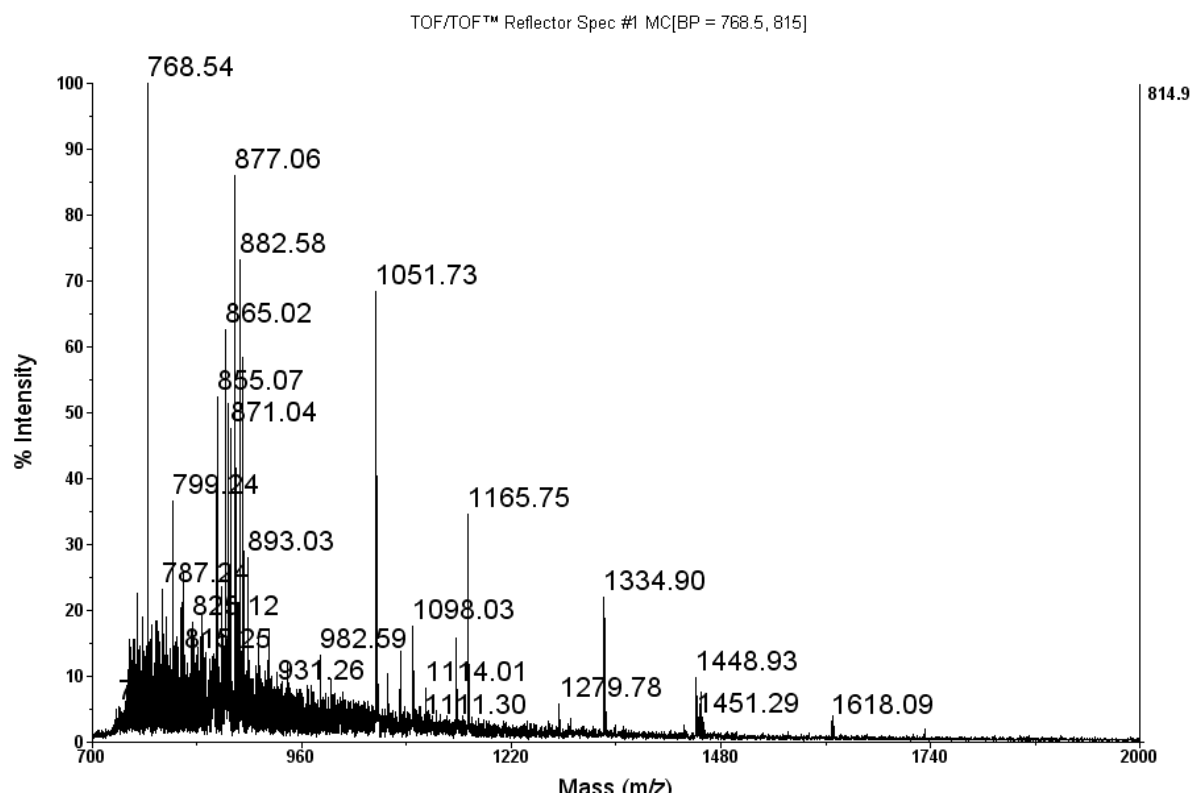


Figure 3.3.6.3. The MALDI mass spectra of the matrix. The α -cyano-4-hydroxycinnamic acid matrix was spotted directly onto the sample plate, allowed to dry and analysed by MALDI MS. The chromatogram ranges from 700-2000 m/z. The formation of matrix clusters is a major reason for the presence of the many peaks in the mass spectra. The most right-side peak indicates the 100% intensity.

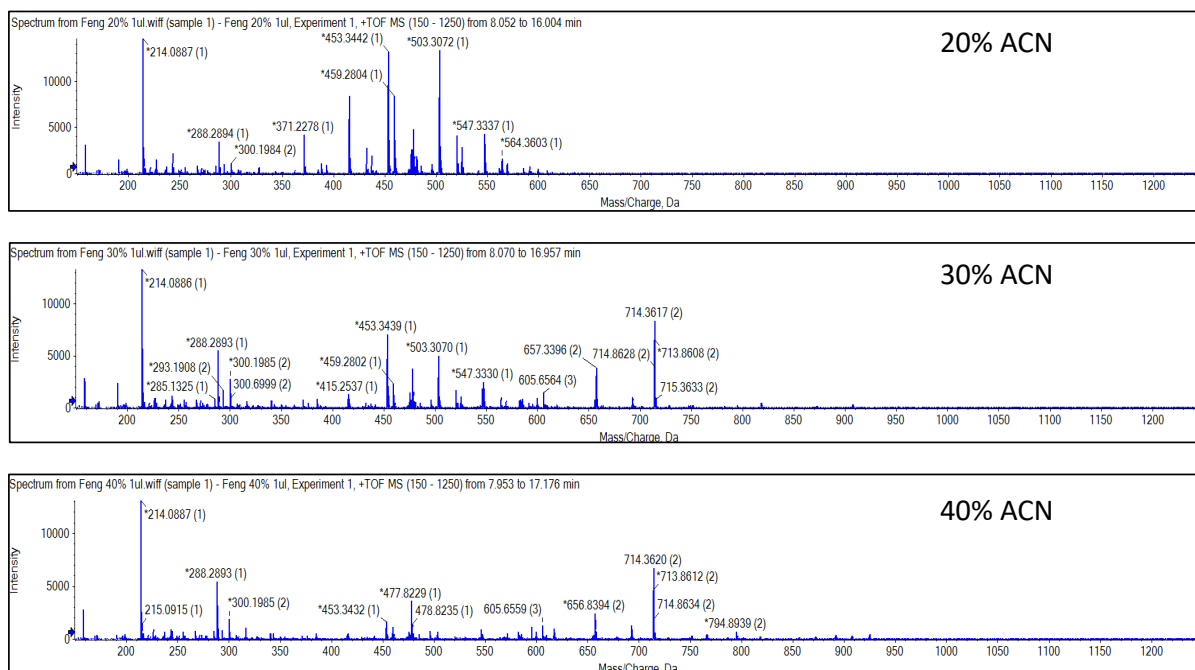


Figure 3.3.6.4. The nLC-ESI mass spectra of the YejA protein mixture. The YejA sample was denatured with urea and loaded onto a C₁₈ ZipTip column. The mass spectra were obtained from eluates of the ZipTip with 20%, 30% and 50% of acetonitrile. The MS spectra range from 150-1250 m/z. Ion peaks in this chromatogram are a mixture of 1⁺, 2⁺ and 3⁺ charged states.

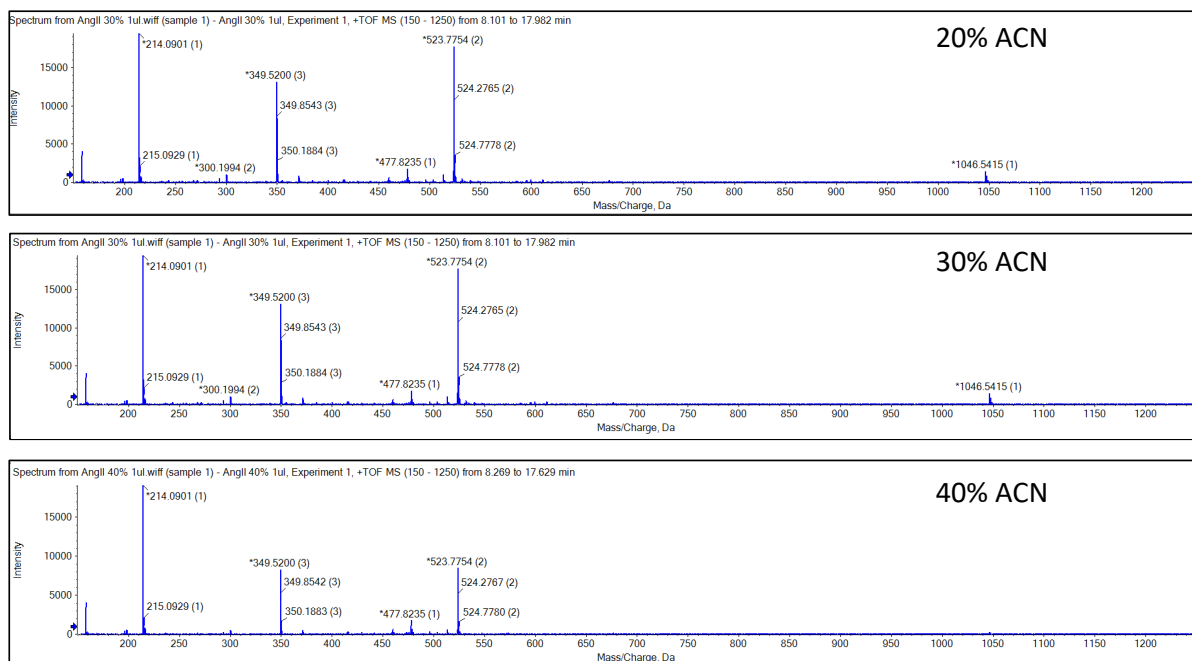


Figure 3.3.6.5. The nLC-ESI mass spectra of BSA and angiotensin II. BSA and angiotensin II (1:1 ratio) were denatured with urea and loaded onto a C₁₈ ZipTip column. The spectra were obtained from eluates of the ZipTip with 20%, 30% and 50% of acetonitrile. The MS spectra range from 200-1250 m/z. Ion peaks in this chromatogram are a mixture of 1⁺, 2⁺ and 3⁺ charged states. Angiotensin can be clearly seen at 349 m/z (3⁺) and 523 m/z (2⁺).

The ion signal at 1426 m/z from MALDI MS was the most intense peak and was observed in samples eluted with 20%, 30%, 40% and 50% of acetonitrile (Figure 3.3.6.1). This peptide is absent in the spectra of the BSA/angiotensin II control and corresponds to a signal at 714 m/z (2⁺) in nLC-ESI MS, making it a promising hit for further investigation. The MALDI MS/MS fragmentation, however, indicated that this peptide is derived from YejA with its sequence interpreted as Val'-Leu'-Gly'-Glu'-Pro'-Arg'-Tyr'-Ala'-Phe'-Asn'-Phe'-Asn' (Figure 3.3.6.6 and Figure 3.3.6.7). There were other ion signals that were found to be YejA fragments.

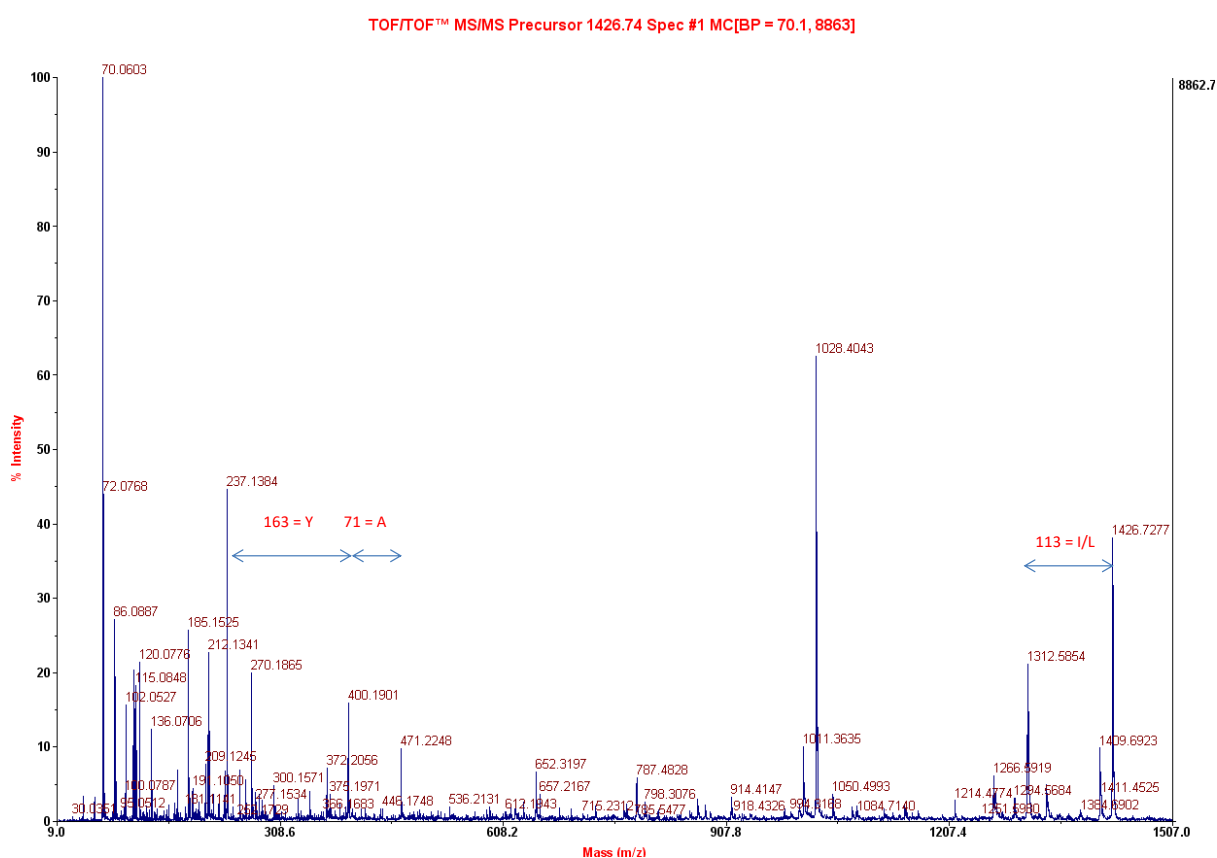


Figure 3.3.6.6. The MALDI MS/MS fragmentation of the peptide at 1426.7 m/z. The most right-side peak indicates the 100% intensity.

>sp|P33913|YEJA_ECOLI (strain K12)

MIVRILLLLFIALFTFGVQAQAIKESYAFAVLGEPRYAFNFNHFDYVNPAAAPKGGQITLSA
LGTFDNFNRYALRGNPGARTEQLYDTLFTTSDDPEGSYYPLIAESARYADDYSWVEVAIN
PRARFHDGSPITARDVEFTFQKFMTEGVPQFRLVYKGTTVKAIAPLTVRIELAKPGKEDM
LSLFSLPVFPEKYWKDHLSDPLATPPLASGPYRVTSWKMGQNIVYSRVKDYWAANLPVN
RGRWNFDTIRYDYYLDDNVAFEAFKAGAFDLRMENDAKNWATRYTGKNFDKKYIIKDEQK
NESAQDTRWLAFNIQRPVFSDDRRVREAITLAFDFEWMNKALFYNAWSRTNSYFQNTHEYAA
RNYPDAAELVLLAPMKKDL PSEVFTQIYQPPVSKGDGYDRDNLLKADKLLNEAGWVLKGQ
QRVNATTGQPLSFELL LPASSNSQWVLPFQHSLQRLGINMDIRKVDNSQITNRMRSRDYD
MMPRVWRAMPWPSSDLQISWSSEYINSTYNAPGVQSPVIDSLINQIIAAQGNKEKLLPLG
RALDRVLTWNYYMLPMWYMAEDRLAWWDKFSQPAVRPIYSLGIDTWYDVNKA AKLP SAS
KQGE

Figure 3.3.6.7. Sequence of an identified YejA peptide. The sequence of full-length YejA of *E. coli* with the peptide sequence highlighted in red. The peptide is observed at 1426 m/z in MALDI MS and 714 m/z (2⁺) in nLC-ESI MS, which accounts for a YejA fragment.

The fact that the YejA sample contains a range of different YejA peptide fragments is a problem for identification of the bound peptide. It is necessary to remove the abundant YejA peptides in order to eliminate the prominent background in MS analysis. The YejA sample was loaded onto a 10 kDa MW cut-off spin filter and the flow-through was collected. The YejA sample before spin filtration, the flow-through and supernatant were analysed by MALDI MS. Mass spectra of the initial YejA sample and the supernatant of spin filter were almost identical, suggesting that these YejA peptides did not go through the 10 kDa filter (Figure 3.3.6.8). The data indicate that the abundant YejA peptides are bound to YejA. These YejA peptides may derive from degradation of YejA. For example, the identified YejA peptide (Figure 3.3.6.6 and Figure 3.3.6.7) is located on a loop of YejA. During protein purification, it is likely that this flexible loop was cleaved by certain protease and the resulting peptide bound to the surface of YejA. We did not observe extra electron density at surface of the YejA structure that could be accounted for by a peptide, suggesting that the peptide might be lost during crystallisation. It

is unlikely that the YejA peptides are created during the MALDI MS process, since MALDI is a soft ionisation technique that does not cause protein/peptide fragmentation.

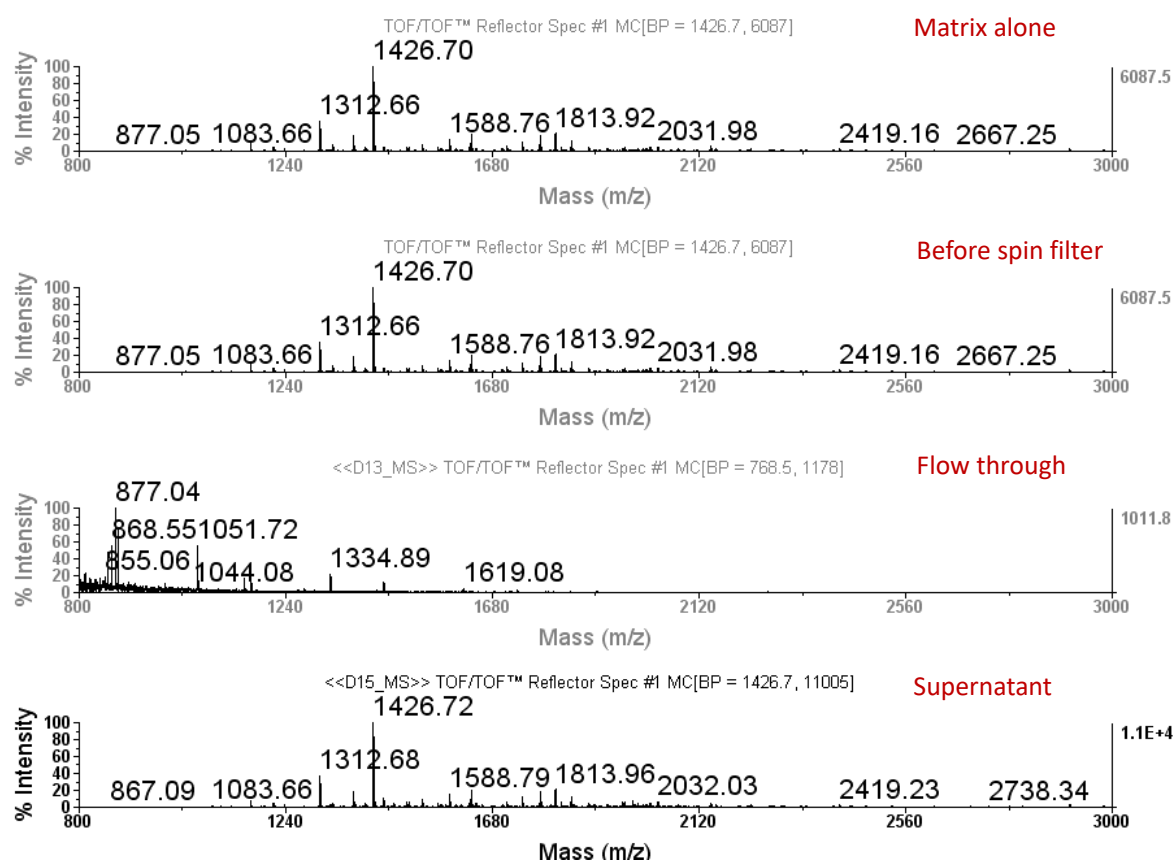


Figure 3.3.6.8. The MALDI mass spectra of the YejA sample with spin filtration. The initial YejA sample, the flow through and supernatant were mixed with the matrix and spotted on the plate. The sample was allowed to dry prior to MALDI MS analysis. Analysis with the matrix alone was performed as a control. The chromatogram ranges from 800-3000 m/z. The most right-side peak indicates the 100% intensity.

YejA crystals were dissolved by protein buffer and loaded on a PD SpinTrap column to separate the crystallised YejA-peptide complex from the abundant peptides and PEG molecules. In total, three fractions were collected and analysed by MALDI MS. MALDI mass spectra of all the three fractions exhibit strong ion signals at 3000-4000 m/z using α -cyano-4-hydroxycinnamic acid as a matrix (Figure 3.3.6.9), which correspond to PEG3350 (Hong *et al.*, 2014). Presence of PEG molecules may have suppressed the ion signals of other fragments in MALDI MS. However, it is unlikely that PEG3350 can occlude the signal from the peptide of interest, since the signal of a ~11-residue peptide, as interpreted from the crystal structure, would be at ~1100 m/z. The MALDI analysis was still releasing YejA peptides as before, for instance, peptides at 1165.7 m/z, 1426.7 m/z and 1449 m/z. The presence of PEG molecules has restricted the usage of nLC-ESI MSMS analysis.

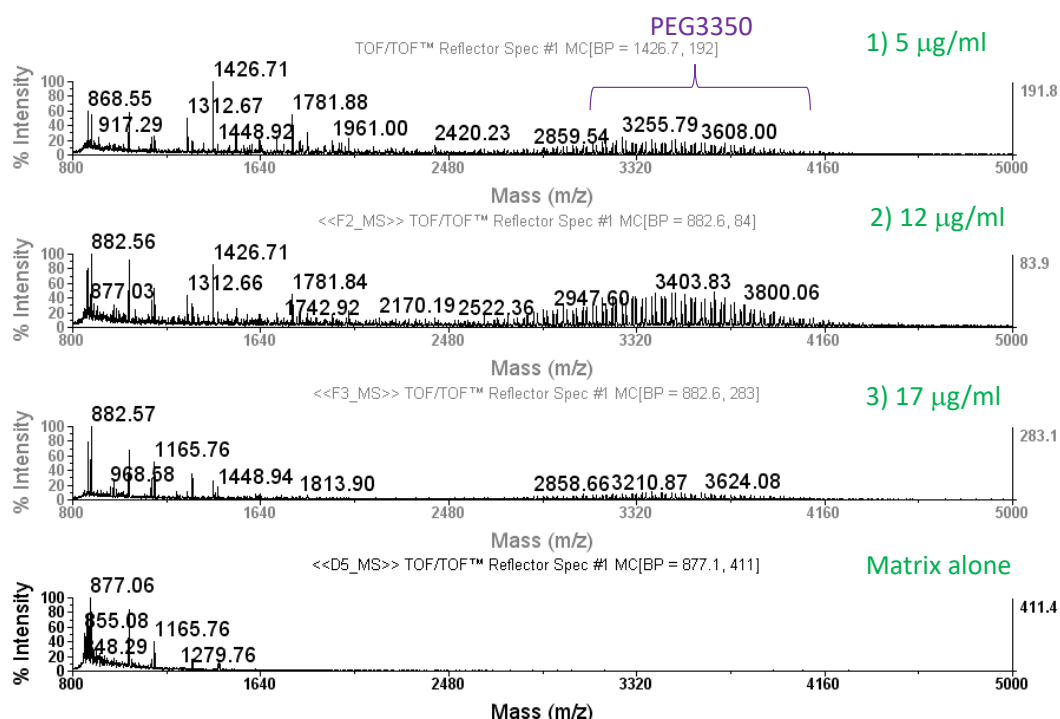


Figure 3.3.6.9. The MALDI mass spectra of sample dissolved from YejA crystals. The three fractions of 5 µg/mL, 12 µg/mL and 17 µg/mL were mixed with the matrix and spotted on the plate. The sample was allowed to dry prior to MALDI MS analysis. Analysis with the matrix alone was performed as control. The chromatogram ranges from 800-5000 m/z. The most right-side peak indicates the 100% intensity.

To date, we have been unable to identify the sequence of the bound peptide by MS. There are three possible reasons for this discrepancy. First, the bound peptide does not ionize MALDI MS or nLC-ESI MS and therefore it does not reach the MS detector. Second, ions from the bound peptide are not stable and fragment in the flight tube. Third, the bound peptide generates a peak in the mass spectra, but it was not detected due to the strong background signals from the abundant YejA-derived peptides that are bound to YejA. Presence of the YejA peptides has begged the question whether the strong electron density in the active site of YejA is a side effect of YejA overexpression or is owing to the presence of a biologically relevant peptide. However, none of the YejA peptides match the sequence interpreted from the electron density in the crystal structure, Asp'-Gly'-Ala'-Ser'-Arg'-Glu'-Ala'-Phe'-Asp'-Lys'-Lys'. This means that the YejA peptides bind to the surface of YejA rather than to the active site. We subjected to the modelled peptide sequence to BLAST, but we could not find a good match to the peptide sequence.

3.4. Discussion

3.4.1. Nature of the bound peptide

Bacterial cells can be exposed to a range of different oligopeptides. As the ligand is trapped inside the binding cavity of SBP, it is possible to grow crystals of oligopeptide-binding proteins harbouring heterogeneous peptides. Consequently, the resultant crystal structure may be an average of different peptides in the crystal lattice. In the YejA structure, the electron density associated with the bound peptide allowed for straightforward model building of a hendecapeptide backbone, while the side-chain density is more ambiguous, indicating a population of different peptides might be present in the binding site. The polypeptide observed in the YejA structure has three possible sources: overexpression of the recombinant protein (*i.e.* YejA), the growth medium (*i.e.* LB) and the expression cell line (*i.e.* *E. coli* BL21 (DE3)). We have identified multiple YejA peptides that are present in the YejA sample by MS. However, none of the YejA peptides match the peptide sequence interpreted from the electron density. This suggests that the bound peptide is not derived from YejA. From our work, there is not enough evidence to define whether this peptide is exogenous or endogenous. We subjected the modelled peptide sequence to BLAST, which compares the known sequence to sequence databases and calculates the statistical significance of matches. However, we could not find a good match to the peptide, suggesting that the interpreted peptide sequence might not be 100% correct. MS analysis failed to identify the amino acid sequence of the bound peptide. Based on the electron density map, we modelled the bound peptide as Asp'-Gly'-Ala'-Ser'-Arg'-Glu'-Ala'-Phe'-Asp'-Lys'-Lys'.

3.4.2. Phe8' might be a critical residue

Structures of AppA from *B. subtilis* and OppA from *L. lactis* were superimposed to the YejA structure based on protein C α atoms (Figure 3.4.2.1). Superimposition shows that backbones of the nonapeptide in AppA and the octapeptide in OppA are almost perfectly overlapped. Both peptides are completely buried in the oligopeptide-binding proteins with no access to solvent. By contrast, the polypeptide in YejA shifts away from where the nonapeptide and octapeptide are located and binds to a region that is close to the surface of YejA. The distinct location of the hendecapeptide in YejA might be owing to the presence of Phe8' in the bound hendecapeptide. Phe8' inserts to the surface of YejA and its benzene ring forms extensive van der Waals contact with Leu72 (domain Ia), Val148 (domain Ib), Gln150 (domain Ib) and Ser507 (domain II) of YejA. We propose that Phe8' is an important residue for YejA recognition and binding.

We compare the YejA-peptide structure with the structure of FhuA in complex with MccJ25 (Mathavan *et al.*, 2014). MccJ25 hijacks the outer-membrane receptor FhuA by mimicking the interactions of ferrichrome (Locher *et al.*, 1998; Mathavan *et al.*, 2014). Structures of FhuA-ferrichrome and FhuA-MccJ25 reveal that minimal conserved interactions are sufficient for MccJ25 to hijack FhuA. The benzene ring of His5' of MccJ25 is buried more deeply inside the plug domain compared to other residues. Similarly, in the YejA-peptide structure, Phe8' inserts to the surface of YejA. Using the phage T5 competition assays and native MS, Mathavan *et al.* (2014) have shown that His5' is essential for MccJ25 recognition, binding and transport.

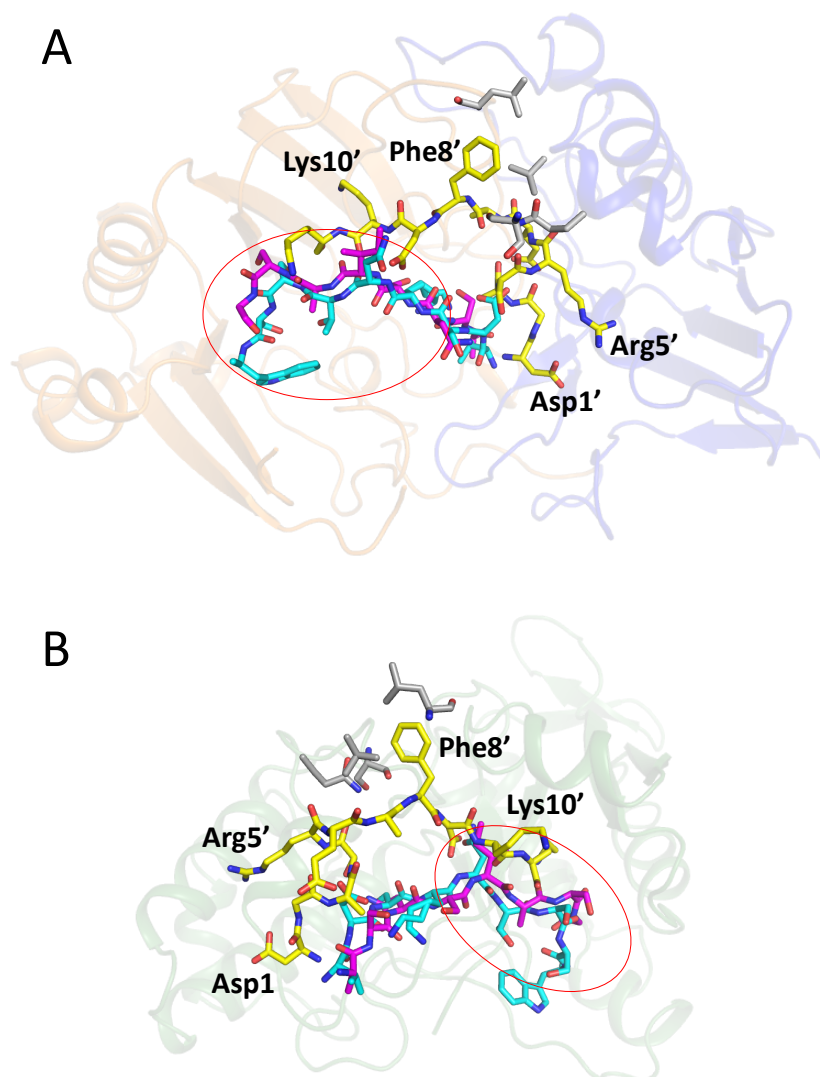


Figure 3.4.2.1. Comparison of peptides binding to YejA, AppA and OppA. Domain Ia, domain Ib and domain II of YejA are coloured in orange, blue and green, respectively. The nonapeptide (Val'-Asp'-Ser'-Lys'-Asn'-Thr'-Ser'-Ser'-Trp') in *B. subtilis* AppA (PDB ID: 1XOC, (Levdikov *et al.*, 2005)) and the octapeptide (Ala'-Ser'-Asn'-Ser'-Ile'-Ala'-Ser'-Gly') in *L. lactis* OppA (PDB ID: 3DRF, (Berntsson *et al.*, 2009)) are superimposed to the YejA structure based on protein C α atoms. The polypeptides in *E. coli* YejA, *B. subtilis* AppA and *L. lactis* OppA are shown in stick format (yellow, cyan and magentas, respectively). Residues that form van der Waals contact with Phe8' are shown in stick format (grey). Oxygen and nitrogen atoms are coloured in red and blue, respectively. The red circle highlights the expanded cavity at the C-terminus of the hendecapeptide.

3.4.3. Binding of the hendecapeptide to YejA

MccC7/C51 is a peptide-based microcin (Figure 3.1.1.1A). Novikova *et al.* (2007) have revealed that MccC7/C51 is a Trojan horse-type microcin and the peptide chain is essential for its active uptake. It is proposed that the peptide chain of MccC7/C51 hijacks YejA by mimicking the natural substrate of YejA to induce Yej-dependent transport. We compare the sequence of the bound peptide in YejA with the heptapeptide in MccC7/C51 and observe no obvious sequence similarity. Binding of the hendecapeptide to YejA is almost exclusively contributed by hydrogen bonds between the peptide backbone and the protein, apart from the side chains of Arg5' and Phe8'. Therefore, we propose that YejA may recognise the ligand based on the peptide backbone. Similar cases have been observed in oligopeptide-binding proteins from *B. subtilis*, *L. lactis* and *S. typhimurium*, where no side chain-specific interactions are formed with the protein (Levdikov *et al.*, 2005; Berntsson *et al.*, 2009; Tame *et al.*, 1994). Previous studies on OppA specificity have shown that backbone recognition and binding site desolvation are the determinants for the protein's promiscuity towards peptides of various amino acid sequences (Figure 2.4.2.1) (Lanfermeijer *et al.*, 2000; Doeven *et al.*, 2004). YejA also contains a hydrated binding site, suggesting that YejA may be able to recognise different peptides. This is in agreement with the voluminous binding cavity of YejA and the lack of constraints at the C-terminus of the bound peptide, which probably allow YejA to accommodate larger peptides. The peptide chain length of MccC7/C51 has been characterised by Vondenhoff *et al.* (2011). YejABEF can transport MccC7/C51 analogues that possess peptide of at least six residues in length. The maximum length of the peptide chain is yet to be confirmed due to the difficulties in their synthesis. Results from our work agree with this conclusion and indicate that the Yej system can recognise and bind a peptide of 11 residues in length.

3.4.4. The “Venus fly-trap” mechanism of YejA

As discussed in section 1.0.5.6, hinge regions of class I and class II SBPs are composed of three and two connecting segments, respectively, whereas the hinge region of class III SBPs is typically formed of a single α -helix. Ligand binding of class I and class II SBPs follows a “Venus fly-trap” mechanism, where the ligand-bound SBPs are structurally distinct from their apo forms (Figure 1.0.5.6) (Berntsson *et al.*, 2010). The distinct structures allow the inner-membrane components to distinguish between ligand-bound and apo SBPs, thereby increasing substrate transport efficiency. However, due to the rigidity of α -helix, class III SBPs undergo a smaller domain movement upon substrate binding. The two domains of the YejA structure are connected by two hinge segments, suggesting that YejA is a class II SBP and follows a “Venus fly-trap” mode upon ligand binding. We have determined the structure of YejA in complex with an oligopeptide. The dynamic behavior of apo YejA can be investigated by MD simulation of YejA with its bound peptide removed. The domain motions between open and closed conformation of YejA can also be calculated. Similar work has been done by Krewulak *et al.* (2005) on the class III SBP FhuD. Their work shows that the C-terminal domain of FhuD closes by 6° upon release of the substrate, which allows the cognate TMDs (FhuB) to distinguish the ligand-bound FhuD.

3.4.5. Conclusion and outlook

MccC7/C51 is an antibacterial peptide and is known to hijack the YejABEF system for internalisation. However, the physiological function and substrate of the YejABEF system is unknown. In this work, I have determined the structure of YejA in complex with an oligopeptide. Due to the presence of YejA-derived peptides, we failed to identify the sequence of the bound peptide by MS. However, based on the electron density, we assigned eleven residues to the bound oligopeptide. Our structure indicates that YejA is probably an oligopeptide-binding protein. MccC7/C51 probably hijacks the YejABEF system by mimicking the interactions between the peptide backbone and YejA to induce the Yej-dependent import.

In order to verify sequence of the bound peptide, future work will focus on synthesising the obtained peptide and examine the binding affinity to YejA. YejABEF is a putative oligopeptide transporter. Knowledge of the structure and physiological substrate of YejA provides precious guidance for designing prodrugs of novel antibiotics. Prodrugs are derivatives of drug molecules that release the active parental drugs upon *in vivo* transformation via enzymatic or chemical cleavage (Jornada *et al.*, 2016). We are interested in exploiting this peptide to develop antibiotics that can penetrate the bacterial membranes through the naturally occurring nutrient importer YejABEF. By conjugating this peptide to known antibiotics, we want to examine the ability of these compounds to hijack the YejABEF system and enable delivery of active parental drugs.

Chapter 4

Inhibitor screening against *E. coli*

AcrAB-TolC

We may write up a paper regarding the flow scheme of XChem fragment screening platform for membrane proteins.

4.0. Introduction

4.0.1. Bacterial resistance

The inaccurate and over prescription of antibacterial drugs has accelerated the development of antibiotic resistance amongst bacteria. Since penicillin-resistant strains were first reported in the 1940s, bacterial antibiotic resistance has been recognised as a global threat to modern medicine (Otto, 2010). Consequently, many of the conventional antibiotics have lost their efficacy as treatments for bacterial infections (Piddock, 2006). Moreover, a growing number of bacteria have developed resistance to more than one antimicrobial, that is multidrug resistance (MDR) (Zechini & Versace, 2009).

Bacteria acquire antibiotic resistance through several mechanisms. The three main types of strategy are: (1) reducing intracellular antibiotic accumulation, (2) inactivating antibiotics, and (3) switching the active sites of intracellular targets (Figure 4.0.1.1) (Hawkey, 1998). For antibiotics to be active, they must reach their intracellular targets and accumulate to certain concentrations. Some bacteria acquire antibiotic resistance by preventing an antibiotic from entering the bacterial cell and/or actively exporting an antibiotic (Figure 4.0.1.1A). The first case of antibiotic efflux was described in 1980, when a tetracycline resistance mechanism was identified in *Enterobacteriaceae* (McMurry *et al.*, 1980). Tetracycline inhibits protein synthesis by preventing the charged aminoacyl-tRNA from attaching to ribosomes (Connell *et al.*, 2003). The protein-synthesis machinery is located in the cytoplasm so tetracycline must penetrate cell membranes and accumulate to a certain concentration to enable activity. Tetracycline-resistant bacteria overexpress efflux transporters that can pump out tetracycline faster than it diffuses in, so the intracellular tetracycline concentration is kept at low

concentrations. Some resistant bacteria protect the target of action by destructing or modifying the chemical property of the intracellular antibiotic (Figure 4.0.1.1B). A classic case is *Staphylococci* strains acquiring resistance to aminoglycosides owing to the presence of aminoglycoside-modifying enzymes (Shaw *et al.*, 1993). Aminoglycosides represent a group of aminocyclitol-containing antibiotics that target the ribosome and disrupt protein synthesis. Aminoglycoside-resistant bacteria produce aminoglycoside-modifying enzymes to modify the aminoglycosides by either adenylation, phosphorylation or acetylation. The modified antibiotics do not bind adequately to RNA and therefore cannot interfere with protein synthesis. Penicillin is a β -lactam antibiotic that inhibits formation of the bacterial cell wall. The four-membered β -lactam ring of penicillin binds to the DD-transpeptidase enzyme. The enzyme loses its catalytic activity to form polymer peptidoglycan cross-links, resulting in an imbalance between cell wall production and degradation. Penicillin resistance arises from two mechanisms; the first being antibiotic inactivation (Figure 4.0.1.1B), and the second involving alternation of the target structure (Figure 4.0.1.1C). In the latter case, the penicillin-binding proteins (PBPs) develop into a form with a lower affinity to penicillin and therefore are not inhibited (Garcia-Bustos & Tomasz, 1990).

A. Decreased antibiotic concentration

B. Antibiotic inactivation

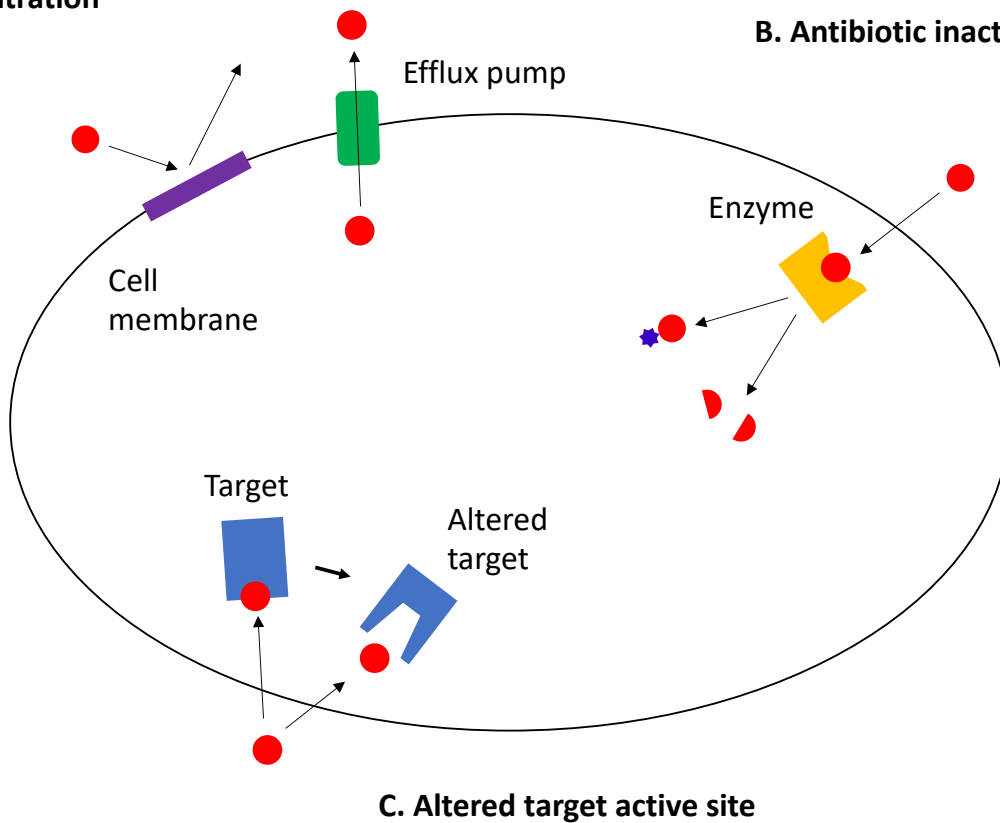


Figure 4.0.1.1. Mechanisms of antibiotic resistance. (A) Decreased intracellular antibiotic concentration owing to reduced influx and/or increasing efflux of antibiotics. (B) Inactivation of intracellular antibiotics can be achieved by either adding an inactivating group (blue asterisk) or by disrupting the molecular structure of an antibiotic. (C) Alternation of the target active site. The antibiotic molecule is shown as a red sphere. Adapted from (Hawkey, 1998).

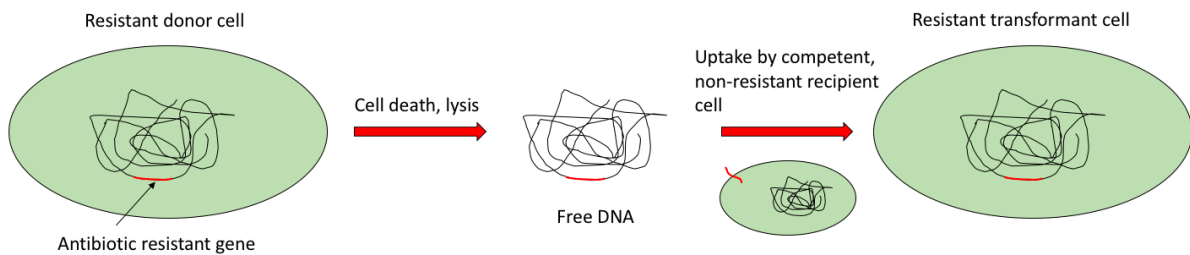
Antibiotic efflux pumps are constitutively expressed in bacterial genomes and are the main reason for intrinsic bacterial resistance (Piddock, 2006). Efflux pumps that are absent in wild-type bacterial strains but expressed after mutations or gene transfer, are responsible for the acquired antibiotic resistance phenomenon (Poole, 2005). To date, the importance of the efflux mechanism in bacterial antibiotic resistance has been reported in most conventional antibiotics.

Bacteria acquire resistance through either sequential mutations or the transfer of drug resistance genes (Shaw *et al.*, 1993). Mutations occur spontaneously in all organisms and do not require the presence of an antibiotic. If one of these mutations confers antibiotic resistance, the bacterium carrying this gene is selectively advantageous over neighbouring bacteria that are sensitive to this antibiotic. In the presence of such antibiotic, the resistant bacteria survive while the sensitive bacteria are killed. The resistant bacterium will divide, fulfil the space and grow into the dominant population in its microbial community. Apart from mutations, antibiotic resistance genes can be transferred between bacteria through transformation, conjugation and transduction (Figure 4.0.1.2) (Levy & Marshall, 2004; Johnston *et al.*, 2014).

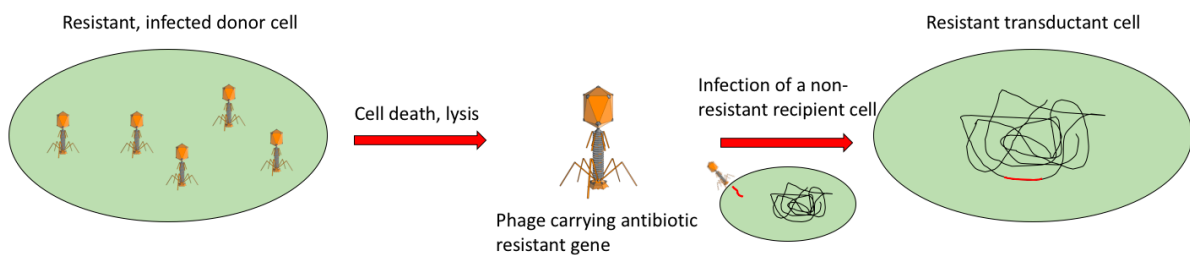
Transformation is the direct uptake and integration of an exogenous gene from the surrounding medium into the genetic system of the recipient bacterium (Figure 4.0.1.2A). An example is *Streptococcus pneumoniae* that acquires penicillin resistance through the incorporation of genes from *Streptococcus viridans*, a naturally occurring penicillin resistant strain, leading to the production of mosaic PBPs (Spratt, 1994). Transduction is the process of gene uptake mediated by a bacteriophage (Figure 4.0.1.2B). For example, two antibiotic resistance genes, *bla*_{TEM} and *bla*_{CTX-M}, once transferred to susceptible bacterial cells through a bacteriophage, confer ampicillin-resistance to the recipient bacteria (Colomer-Lluch *et al.*, 2011). Conjugation is the gene transfer from one bacterium to another through direct cell-cell contact or pilus-mediated connection (Figure 4.0.1.2C). For example, *Campylobacter jejuni* has been

established as a major cause of bacterial diarrheal disease in animals and human. Tetracycline resistance in *Campylobacter* has been shown to be mediated by plasmids and can be transferred between tetracycline-resistant *Campylobacter* bacteria via conjugation (Taylor *et al.*, 1981; Taylor *et al.*, 1983; Tenover *et al.*, 1985).

A. Transformation



B. Transduction



C. Conjugation

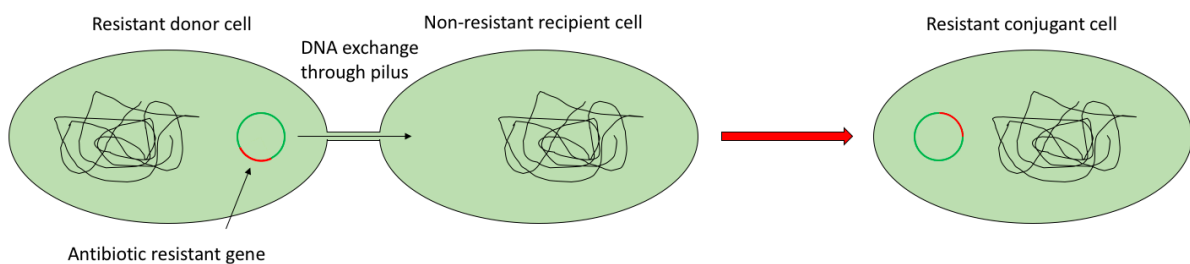


Figure 4.0.1.2. Horizontal gene transfer between bacteria. (A) Transformation occurs when naked DNA is taken up by the recipient organism. (B) In transduction, gene transfer is mediated by a bacteriophage. (C) Conjugation occurs by direct contact or pilus transfer between two bacteria. Adapted from (Furuya & Lowy, 2006).

4.0.2. RND-type efflux pumps

Overexpression of bacterial efflux pumps is a major reason for multidrug resistance (Du *et al.*, 2018). The RND superfamily is responsible for translocation of a diverse range of antimicrobial compounds and is associated with a clinically significant multidrug resistance phenomenon. Gram-negative bacteria tend to have higher resistance to antimicrobials than Gram-positive bacteria. This was initially explained by the presence of the outer membrane barrier that minimizes access of antibiotics to their intracellular targets (Nikaido, 1988). The outer leaflet of outer membrane is composed of LPSs containing almost exclusively saturated fatty acid residues, resulting in an unusual asymmetry of the outer membrane lipid bilayer and a low fluidity of its outer leaflet (Kamio & Nikaido, 1976). Therefore, the permeation rates of hydrophobic compounds decrease by nearly 2-fold compared to conventional phospholipid bilayers (Plesiat & Nikaido, 1992). Porins (pore size: 6-15 Å) are molecular sieves in the outer membrane that allow hydrophilic drugs to penetrate (Nikaido & Vaara, 1985). *P. aeruginosa* displays the highest antibiotic resistance among Gram-negative bacteria, whose porin pore sizes are the smallest, highlighting the contribution of narrow porins to drug resistance (Yoshihara & Nakae, 1989). Nevertheless, permeability studies have shown that all β -lactams equilibrate across the *E. coli* outer membrane within 60 seconds, a time span much shorter than the *E. coli* generation time. This indicates that the outer membrane alone cannot generate a sufficient level of resistance, and therefore another resistance mechanism must be working synergistically with the outer membrane barrier (Nikaido, 1989).

The MexB exporter was first identified as a multidrug transporter in *P. aeruginosa* (Poole *et al.*, 1993). Mutations in this transporter were found to restore activity of multiple antibacterial drugs against *P. aeruginosa* (Nikaido, 1994). Following studies have shown that efflux transporters are more important to drug resistance than porins (Li *et al.*, 1994). MexB works in

conjunction with the MFP MexA and the outer membrane channel OprM to form a tripartite system. Similar tripartite complexes have later been identified including MexCD-OprJ (Poole *et al.*, 1996), MexEF-OprN (Kohler *et al.*, 1997) and MexXY-OprM (Mine *et al.*, 1999) in *P. aeruginosa* as well as AcrAB-TolC and AcrAD-TolC in *E. coli* (Nishino & Yamaguchi, 2001). All these systems belong to the RND superfamily. Members of the RND superfamily are ubiquitously found in all major kingdoms of life, but have only been shown to contribute to multidrug resistance in Gram-negative bacteria (Tseng *et al.*, 1999). Phylogenetic analyses have defined 10 families within the RND superfamily (Table 4.0.2.1) (Reddy *et al.*, 2016; Nikaido, 2018).

RND family	Domain
SecDF protein-secretion accessory protein (SecDF) family	Bacteria and Archaea
Hydrophobe and amphiphile efflux-1 (HAE1) family	Bacteria
Heavy metal efflux (HME) family	Bacteria
Nodulation factor exporter (NFE) family	Bacteria
Hydrophobe and amphiphile efflux-3 (HAE3) family	Archaea and Bacteria
Aryl polyene pigment extrusion (APPE) family	Bacteria
Hydrophobe and amphiphile efflux-2 (HAE2) family	Bacteria and Eukaryota
Eukaryotic sterol transporter (EST) family	Eukaryota
Dispatched family found in eukaryotes	Eukaryota
An uncharacterised family	Bacteria

Table 4.0.2.1. The 10 families of the RND superfamily. Classification based on the Transporter Classification Database (Reddy *et al.*, 2016).

There are six families of multidrug transporters/efflux pumps in Gram-negative bacteria; these are MFS transporters, MATE transporters, SMR transporters, PACE transporters, ABC transporters and RND transporters (Li & Nikaido, 2009). Different energy sources are utilised among these multidrug transporters. Most of these transporters reside in the plasma membrane and translocate drugs from the cytoplasm to the periplasm. Only the RND-type pumps (with a few exceptions) exist in a tripartite complex composed of an inner-membrane RND transporter, a periplasmic adaptor protein from MFP family and an outer-membrane protein from the OMF family, expanding both the inner and outer membrane (Figure 4.0.2.1). Each of these components are essential for formation of a functional complex. Due to their tripartite composition, RND pumps export substrates directly into the extracellular medium bypassing the periplasm, after which the drugs must traverse the low-permeable outer membrane barrier to re-enter the bacteria (Nikaido, 1996). Wild-type *E. coli* strains are resistant to most lipophilic antibiotics such as penicillin G and cloxacillin and this was thought to be due to the existence of the outer membrane barrier. Nevertheless, researchers have shown that even in the presence of an intact outer membrane, disruption of a major RND transporter, AcrB, almost completely sensitizes the *E. coli* cells to cloxacillin, a lipophilic penicillin, with the MIC decreasing from 512 µg/mL in the wild type to 2 µg/mL (Nikaido, 1998).

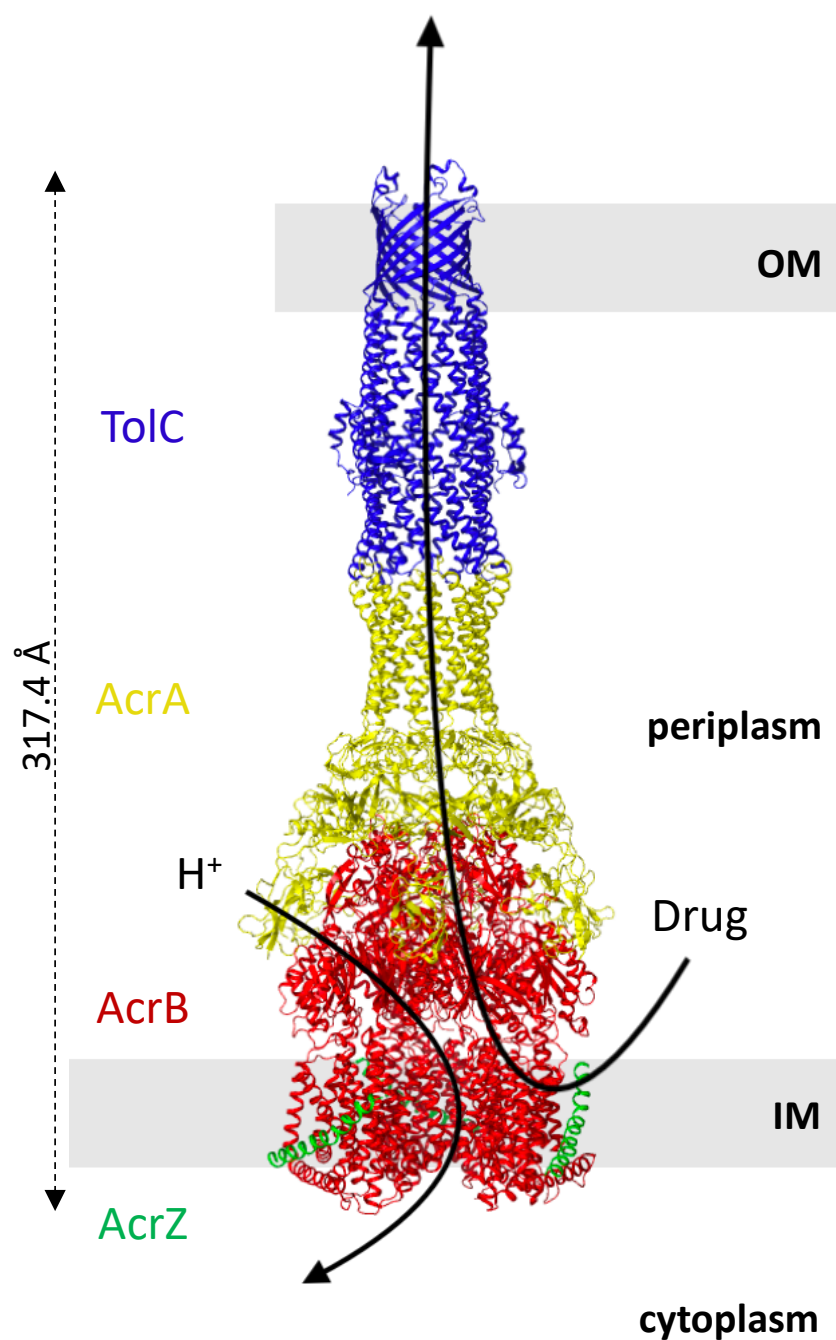


Figure 4.0.2.1. The AcrA₆B₃Z₃-TolC₃ efflux pump. The assembly is composed of the homotrimers AcrBZ-TolC and the hexamer AcrA, and traverses both the inner- and outer-membranes of *E. coli* (PDB ID: 5O66, (Wang *et al.*, 2017)). The overall mass of the AcrABZ-TolC complex is 771 kDa. The inner-membrane transporter AcrB captures substrates from either within the inner membrane or the periplasm, and extrudes them to the outer membrane protein TolC, where they are transported to the extracellular medium. AcrB is the major component of the system, responsible for substrate specificity and energy transduction. Drug efflux process is mediated by AcrB through a functional rotation mechanism coupled to proton influx. AcrA, AcrB, AcrZ and TolC are coloured in yellow, red, green and blue, respectively.

RND efflux pumps demonstrate a broad substrate spectrum (Elkins & Nikaido, 2002). Typical substrates of RND transporters include antibiotics, dyes, antiseptics, detergents, ions and small cytotoxic solvents. These RND exporters recognise a defined set of substrates through specific mechanisms and are therefore inhibited differently. For example, AcrB of *E.coli* and MexB of *P. aeruginosa* do not transport aminoglycosides such as streptomycin and kanamycin, while AcrD of *E.coli* and MexY of *P. aeruginosa* do (Elkins & Nikaido, 2002). Similarly, the potent inhibitor of AcrB and MexB, pyridopyrimidine derivative ABI-PP, does not affect MexY (Nakashima *et al.*, 2013).

To date, all functionally characterised RND-type pumps function as exporters using a substrate/proton antiport mechanism, where efflux of drugs is coupled to an influx of protons down the electrochemical gradients (Tseng *et al.*, 1999). From amino acid sequences, all RND-type transporters are expected to contain 12 TM helices and two large periplasmic domains between TM1-TM2 and TM7-TM8 (Tseng *et al.*, 1999). In the best characterised RND efflux pumps (*i.e.* *E. coli* AcrAB-TolC and CusCBA, *P. aeruginosa* MexAB-TolC), the inner-membrane RND transporter exists as a homotrimer. The periplasmic domains contain the ligand-binding pockets, while the transmembrane domains are responsible for energy transduction.

Six RND efflux transporters have been identified in *E. coli*. These transporters can be classified into two subfamilies: the hydrophobic and amphiphilic efflux RND (HAE-RND) family and the heavy metal efflux RND (HME-RND) family (Nies, 2003). Specifically, multidrug efflux pumps Acr(A)B (Tikhonova & Zgurskaya, 2004; Pos, 2009), Acr(A)D (Rosenberg *et al.*, 2000; Elkins & Nikaido, 2002), Acr(E)F (Nishino & Yamaguchi, 2001), Mdt(A)BC (Baranova & Nikaido, 2002; Nagakubo *et al.*, 2002) and Mdt(E)F (Kobayashi *et al.*, 2006) belong to the HAE-RND subfamily. By contrast, only one transporter, Cus(CB)A (Chacón *et al.*, 2014),

which confers resistance to Cu(I) and Ag(I), belongs to the HME-RND subfamily. Outer membrane protein and/or MFP components are indicated in brackets. To date, only two of these RND pumps have been structurally characterised, which are AcrB (Murakami *et al.*, 2002; Murakami *et al.*, 2006; Seeger *et al.*, 2006) and CusA (Long *et al.*, 2010).

During the past decade, clinical studies have shown that efflux pumps play an increasingly important role in emergence and dissemination of antibiotic resistance among Gram-negative bacteria (Piddock, 2006). It is often the case that during antibiotic therapy of patients with bacterial infections, initial susceptible isolates gradually lose susceptibility to a variety of different antimicrobial molecules that are structurally unrelated (Giamarellou & Poulakou, 2009; Rice, 2009; Maragakis, 2010). This phenomenon agrees with the description of poly-specific efflux pumps (usually RND transporters) that can recognise and export a wide range of cytotoxic substances (Piddock, 2006; Li & Nikaido, 2009).

4.0.3. AcrABZ-TolC multidrug efflux pump

Multidrug efflux pumps play an essential role in bacterial intrinsic and acquired resistance to a range of different antibiotics (Piddock, 2006). In Gram-negative bacteria, multi-protein assemblies are formed by an inner-membrane transporter, an outer membrane channel and a periplasmic adaptor protein that connects the two membrane-spanning components (Tseng *et al.*, 1999). These tripartite efflux pumps span the cell envelope; thus, drugs can be directly pumped into the extracellular medium, bypassing the periplasm. A major reason for antibiotic resistance in *E. coli* is overexpression of the multidrug efflux pump AcrAB-TolC (Okusu *et al.*, 1996) with its homologues found in most Gram-negative bacteria (Paulsen *et al.*, 2000). AcrAB-TolC is an RND-based tripartite efflux system, consisting of an inner-membrane RND transporter AcrB, an outer membrane protein TolC and a periplasmic MFP AcrA (Figure 4.0.2.1) (Du *et al.*, 2014; Wang *et al.*, 2017). Specifically, AcrB is a homotrimeric antiporter that is responsible for substrate recognition and energy transduction. TolC extends from the external medium into the periplasm by forming a pore in the outer membrane. AcrA is the adaptor protein that mediates the connection between AcrB and TolC. The drug efflux process is mediated by AcrB through a functional rotation mechanism coupled to proton influx (Figure 4.0.2.1) (Yamaguchi *et al.*, 2015). A 49-residue protein, AcrZ, has been recently identified by genetic screening as a binding partner of AcrB (Hobbs *et al.*, 2012). AcrZ is an inner-membrane protein that modifies the transport of a subclass of AcrB substrates, including puromycin, tetracycline and chloramphenicol.

Crystal structures of each of the four components of AcrABZ-TolC complex have been determined (Koronakis *et al.*, 2000; Murakami *et al.*, 2002; Mikolosko *et al.*, 2006; Du *et al.*, 2014; Wang *et al.*, 2017). AcrB and TolC are both homo-trimers; however, there was a debate over the number of AcrA and AcrZ components in the system. The recently determined cryo-

EM structures of the AcrABZ-TolC system have revealed that the AcrABZ-TolC system follows a 6:3:3:3 compositional stoichiometry (hexamer AcrA : homotrimer AcrB : homotrimer AcrZ : homotrimer TolC) (Figure 4.0.2.1) (Du *et al.*, 2014; Wang *et al.*, 2017). In contrast to previous crystal structures (Murakami *et al.*, 2002; Symmons *et al.*, 2009), the cryo-EM structures of the 771 kDa AcrABZ-TolC complex suggest that AcrB and TolC are entirely connected by AcrA and there is no direct interaction between AcrB and TolC (Du *et al.*, 2014; Wang *et al.*, 2017). The vertical length of the AcrAB-TolC complex is 317 Å, which is much longer than that of the earlier AcrAB-TolC model (270 Å) where AcrB and TolC interact directly with each other (Symmons *et al.*, 2009).

The physiological role of the AcrAB-TolC system is to remove bile salts and their derivatives present in intestinal tract (Thanassi *et al.*, 1997). Bile salts are produced by the animal liver and are secreted into the bile to assist fat digestion and dispersion in the small intestine. In humans, primary bile acids consisting of cholate and chenodeoxycholate are secreted in a form conjugated to either taurine or glycine via amide bonds (Franklund *et al.*, 1993). Intestinal microflora then deconjugate and transform the cholate and chenodeoxycholate into secondary bile acids deoxycholate and lithocholate, respectively. The bile acid concentration in the small intestine can be as high as 20 mM. Bile salts, as a type of detergent, show high levels of cellular toxicity since they disrupt cell membranes via rupture of the lipid bilayer. Since the animal intestinal tract is the natural habitat of *E. coli*, they must defend themselves against the toxicity of bile salts. *E. coli* acquire bile salt resistance by overexpressing the AcrAB-TolC and EmrAB-TolC systems. Specifically, AcrAB-TolC has been shown to play a more significant role in bile acid efflux (Thanassi *et al.*, 1997).

Moreover, the AcrAB-TolC system extrudes a wide variety of cytotoxic substances, including antibiotics including erythromycin, doxorubicin, rifampicin, novobiocin, minocycline,

carbenicillin and cloxacillin; detergents-like compounds like sodium dodecylsulfate (SDS) and Triton X-100; small solvents such as hexane, cyclohexane and benzene; and dyes including crystal violet, acriflavine, rhodamine 6G and ethidium bromide (Figure 4.0.3.1) (White *et al.*, 1997). These substrates include both aliphatic and aromatic compounds and are charged in diverse manners; cationic, anionic, non-ionic and amphiphilic. The only similarity between the substrates of AcrABZ-TolC is a certain degree of lipophilicity. A possible interpretation for this is that these substrates utilise lipophilic groups to insert into the lipid bilayer and are captured by AcrB from either within the inner membrane or from the periplasm/inner-membrane interface (Nikaido, 1996).

Knowledge of the kinetic behaviour of efflux pumps is essential to understanding their roles in drug resistance. Efflux of different antibiotics has been demonstrated using intact *E. coli* cells overexpressing AcrAB-TolC (Nagano & Nikaido, 2009; Lim & Nikaido, 2010). Results show that the AcrAB-TolC system pumps out drugs at a significant rate. The plot of the efflux rate (V_e) versus periplasmic concentration of the drug (C_p) is often sigmoidal. Specifically, for example, the V_{max} values for efflux of oxacillin, piperacillin, penicillin V and ampicillin were determined to be 0.99 nmol/mg/s, 0.38 nmol/mg/s, 0.76 nmol/mg/s and 0.37 nmol/mg/s, respectively (Lim & Nikaido, 2010).

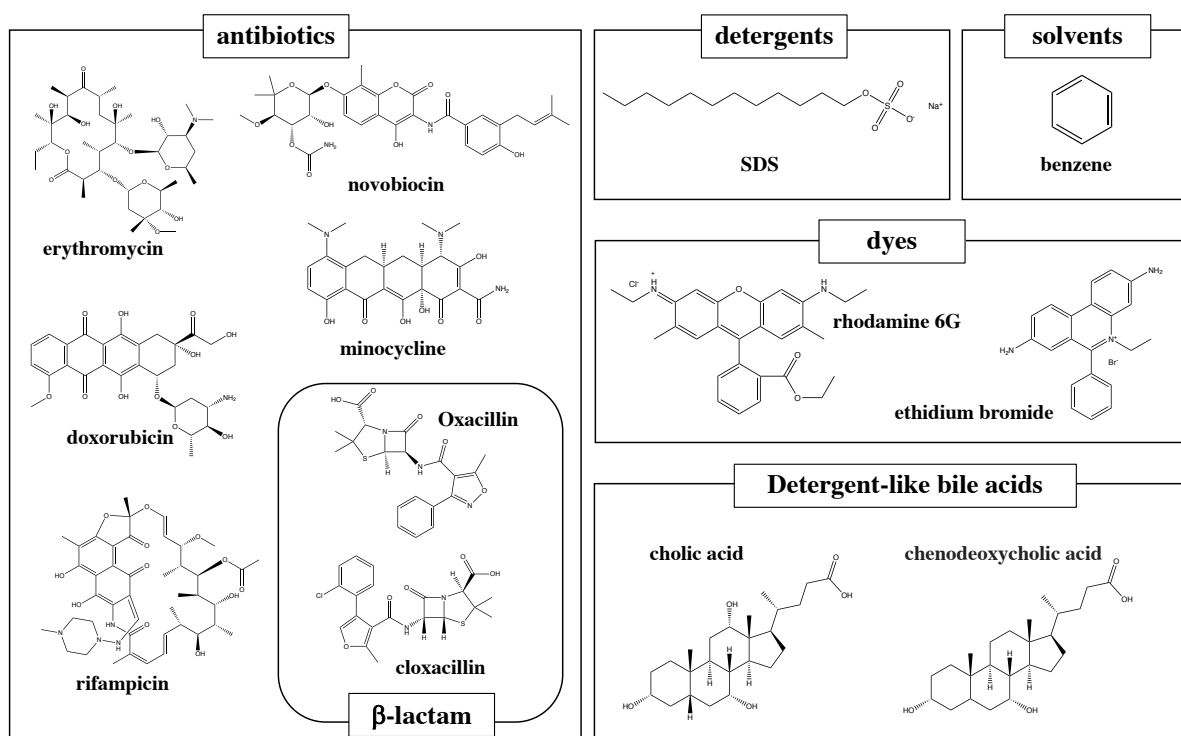


Figure 4.0.3.1. Examples of substrates of AcrAB-TolC. The AcrAB-TolC system transports a variety of substrates including antibiotics, detergents, solvents and dyes. The physiological role of this system is to remove bile salts and their derivatives present in the intestinal tract. Adapted from: (Yamaguchi *et al.*, 2015)

4.0.4. AcrB

4.0.4.1. Structure of AcrB

Acriflavine resistance protein B (AcrB) is a member of the RND superfamily (Nishino & Yamaguchi, 2001). AcrB is a major component of the AcrAB-TolC system, responsible for substrate selectivity and energy transduction. The first crystal structure of AcrB was determined at 3.5 Å resolution, comprising an apo AcrB homotrimer with a 3-fold symmetry (space group *R*32) (Murakami *et al.*, 2002). Each AcrB monomer consists of 1049 amino acids with a MW of 113,574 Da. AcrB constitutes a transmembrane domain of 50 Å in length and a periplasmic headpiece protruding 70 Å into the periplasm (Figure 4.0.4.1). The transmembrane domain is formed of 12 TM α -helices (TM1-TM12) and the hydrophilic headpiece is composed of the two hydrophilic loops between TM1-TM2 and TM7-TM8. The hydrophilic headpiece comprises two domains: the docking domain and the pore domain. The AcrB monomer possesses pseudo-2-fold structural symmetry with its N- and C-terminal halves each comprising six TM α -helices, one subdomain (DC and DN) from the docking domain and two β - α - β subdomains (PC1 and PC2 or PN1 and PN2) from the pore domain (Murakami *et al.*, 2002). This suggests that AcrB may have evolved from gene duplication. An elongated horizontal β -hairpin from the DN subdomain inserts into the DC domain of a neighbouring monomer, which stabilises the AcrB homotrimer and forms a jelly-fish-like structure. A short vertical hairpin protrudes towards the periplasm.

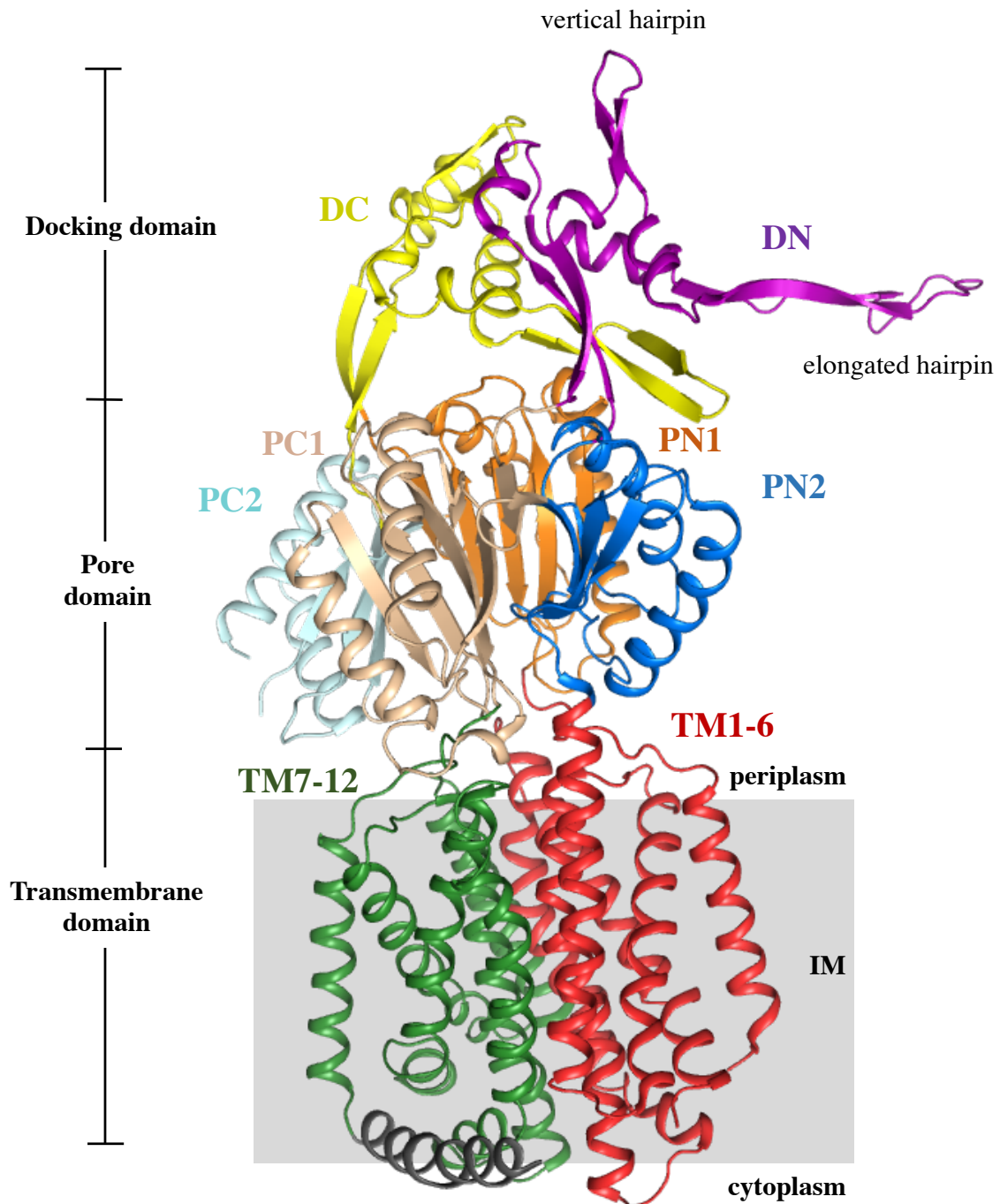


Figure 4.0.4.1. Structure of the AcrB monomer. Ribbon representation of an AcrB monomer in the R32 crystal form (PDB ID: 1IWG, (Murakami *et al.*, 2002)). Each monomer of AcrB is composed of a docking domain, pore domain and transmembrane domain. The elongated β -hairpin from the DN subdomain connects a neighbouring monomer and stabilises the AcrB trimer. A short vertical hairpin protrudes towards the periplasm. Subdomains in the docking domain and pore domain, and each half of the transmembrane domain are indicated by different colours.

The three elongated hairpins enable the three monomers of AcrB to interact tightly with each other and hereby form a stable trimeric structure (Figure 4.0.4.2A) (Murakami *et al.*, 2002; Murakami *et al.*, 2006; Seeger *et al.*, 2006). In the AcrB trimer, the pore domains and docking domains form a trimeric periplasmic domain, 70 Å in length and 80 Å in diameter. The three transmembrane domains form a trimeric transmembrane domain that is composed of 36 TM α -helices (12 from each monomer), which is 50 Å in length and 100 Å in diameter (Figure 4.0.4.2C). The transmembrane domain is open to the cytoplasmic side of the plasma membrane. At the pore domains, the three central α -helices (designated pore helices), donated by the three PN1 subdomains of each monomers, form a central pore at the centre of the trimer (Figure 4.0.4.2B). The central pore has a small diameter, which prevents drugs from passing through (Pos, 2009). The PC1 and PC2 subdomains constitute a periplasmic cleft. The PC2 subdomain and the PN2 subdomain from the neighbour monomer form a vestibule at the monomeric interface. A funnel-like opening of 30 Å in diameter is located at the top of the three docking domains and is closed at the bottom through strong interactions with the three pore helices. In the transmembrane domain, a phospholipid-filled central hole is formed by three 12- α -helix clusters with a diameter of ~ 30 Å (Figure 4.0.4.2C) (Murakami *et al.*, 2006; Seeger *et al.*, 2006). A central cavity is located between the central pore and the phospholipid surface of the central hole. It is connected to the periplasm through three vestibules between the interfaces of three monomers.

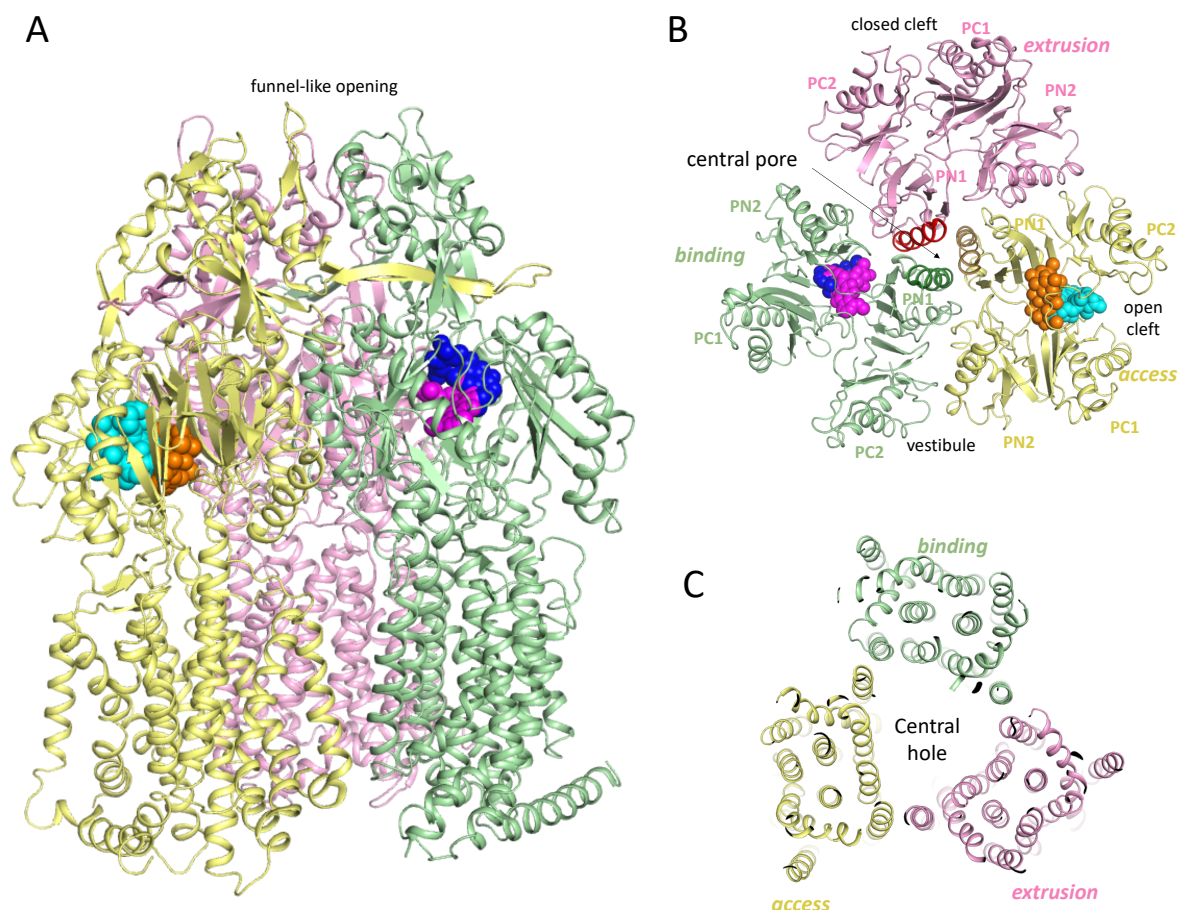


Figure 4.0.4.2. Crystal structure of the asymmetric AcrB in a ligand-bound state. (A) Asymmetric structure of the doxorubicin-bound AcrB trimer from the C2 crystal form (doxorubicin is shown as blue spheres; PDB ID: 2DR6), in which minocycline (magenta; PDB ID: 2DRD), rifampicin (cyan; PDB ID: 3AOB) and erythromycin (orange; PDB ID: 3AOC) are overlapped (Murakami *et al.*, 2006; Nakashima *et al.*, 2011). Structures of AcrB in complex with these four antibiotics have been determined independently. A funnel-like 30-Å opening is located at the top of the trimer. Yellow, green and pink indicate the access, binding and extrusion monomers, respectively. (B) Bottom view of the pore domain. A central pore is formed by three pore helices (highlighted in dark colours) donated by the PN1 subdomains of each of the monomers. The PC1 and PC2 subdomains constitute a periplasmic cleft. The three vestibules between interfaces of the three monomers connect the periplasm to the central cavity. (C) Bottom view of the transmembrane domain of the AcrB trimer. Three 12-helical bundles in the TM domain confine a central hole with a diameter of ~ 30 Å.

4.0.4.2. Mechanism of transport

The physiologically-relevant structure of AcrB was determined from crystals belonging to the space group *C2* (Figure 4.0.4.2) (Murakami *et al.*, 2006; Seeger *et al.*, 2006). This structure has no 3-fold symmetry and is therefore termed as the asymmetric structure. Each monomer adopts a different conformation: access, binding and extrusion states, representing three steps in the substrate transport cycle. The access monomer of the asymmetric structure resembles the symmetric AcrB structure.

In the asymmetric structure, bound doxorubicin and minocycline molecules are observed in a voluminous pocket that is located between the β -sheets of PC1 and PN2 in the pore domain of the binding protomer (Figure 4.0.4.2) (Murakami *et al.*, 2006; Seeger *et al.*, 2006). The doxorubicin/minocycline-binding pocket is connected by a hydrophilic channel to an inner-membrane entrance (entrance 1), which alternates between open and closed states during drug efflux (Figure 4.0.4.3, B and C). The inner-membrane entrance is proximal to but distinct from the previously described vestibule, which stays constitutively open (Figure 4.0.4.3A). In the binding monomer, this channel opens at the inner-membrane entrance, but is blocked at the distal end of the binding pocket by an inclined central helix of the extrusion protomer. Therefore, the binding monomer is in an inside-open but outside-closed conformation. By contrast, the extrusion protomer exhibits an inside-closed but outside-open conformation. In the extrusion monomer, the N-terminal end of TM8 elongates and blocks the inner-membrane entrance (Figure 4.0.4.3C). The periplasmic entrance (the cleft, entrance 2) is also closed due to the movement of subdomain PC2 towards subdomain PC1. The swing of PC1 towards PN2 results in shrinking of the distal binding pocket. The exit of the distal binding pocket opens because the PN1 subdomain moves away from the PN2 subdomain and the central helix inclines away from the exit towards the binding protomer. The bound drug is thus squeezed

into the central funnel. The access monomer represents a resting state. The conformation of the access monomer is similar to that of the binding monomer with both entrances open, but the distal binding pocket remains shrunk. Site-directed mutagenesis showed that both the entrance 1 and entrance 2 play a role in substrate export (Husain *et al.*, 2011; Nakashima *et al.*, 2011). Conversion from the extrusion to the access conformation is proposed to be the proton motive force-dependent step, which requires binding of a proton from the periplasm to the proton translocation site (Asp407, Asp408, and Lys940) (Seeger *et al.*, 2006). Conversion from the binding to the extrusion state involves release of a proton from the proton translocation site to the cytoplasm. Molecular simulations suggest that the inner-membrane entrance and the periplasmic entrance are responsible for the transport of hydrophobic and hydrophilic substrates, respectively (Yao *et al.*, 2013). Nakashima *et al.* (2011) have identified a third putative entrance from the central cavity (entrance 3), but evidence for its function remains to be confirmed.

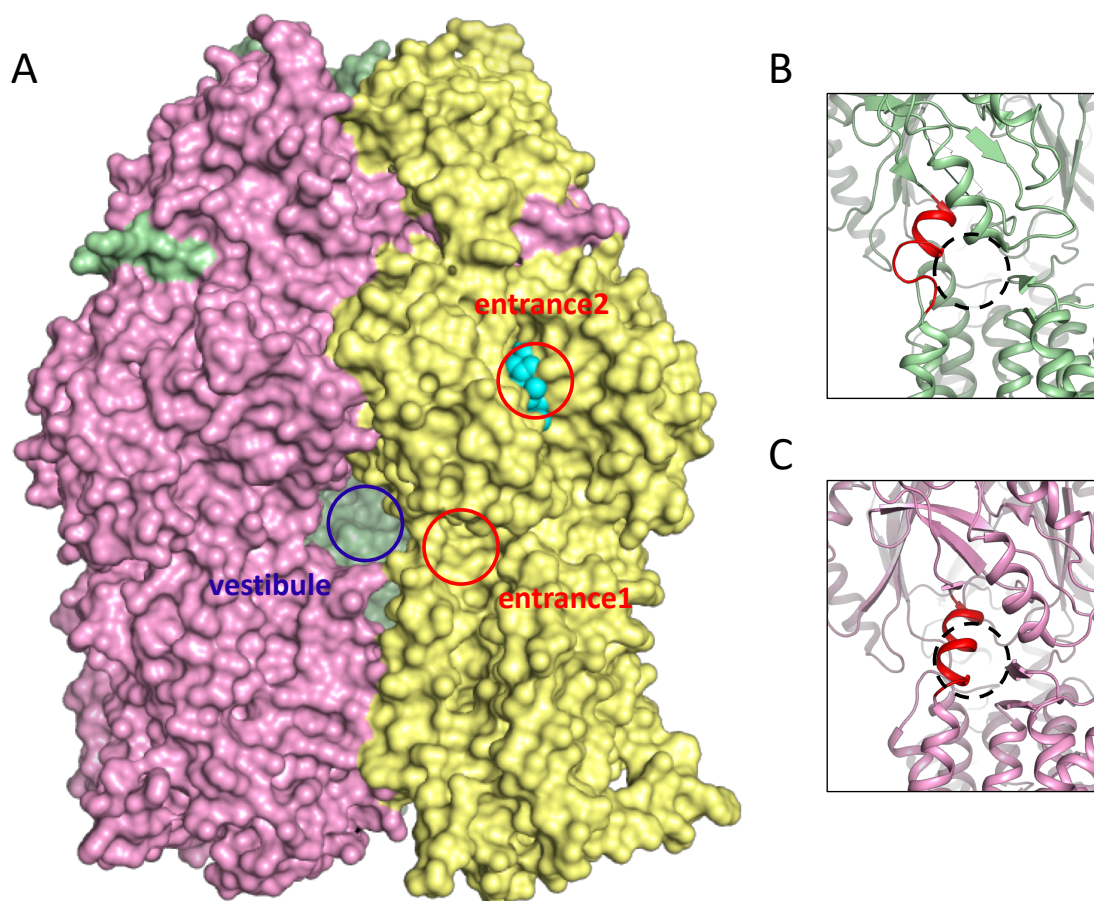


Figure 4.0.4.3. Entrances of the AcrB drug-transport channel. (A) Surface model of the AcrB trimer. The inter-membrane entrance (entrance 1) and periplasmic entrance (entrance 2) are shown as red circles. The vestibule and rifampicin molecule are illustrated by blue circle and cyan spheres. Yellow, green and pink indicate the access, binding and extrusion monomers, respectively. (B) The open inner-membrane entrance of the binding monomer. (C) The closed inner-membrane entrance of the extrusion monomer. The untied and extended N-terminal TM8 is coloured in red. PDB ID: 2DR6 and 3AOB, (Murakami *et al.*, 2006; Nakashima *et al.*, 2011).

Following the doxorubicin and minocycline-binding structures, Nakashima *et al.* reported structures of AcrB in complex with rifampicin and erythromycin (Figure 4.0.4.2) (Nakashima *et al.*, 2011). Compared with doxorubicin (MW: 544 Da) and minocycline (MW: 457 Da), rifampicin (MW: 823 Da) and erythromycin (MW: 734 Da) are larger in size and bind to the access monomer. Doxorubicin and minocycline are referred to as low molecular mass drugs (LMMDs), while Rifampicin and erythromycin are high molecular mass drugs (HMMDs). The LMMD- and MMMD-binding pockets are referred to as the distal and proximal pockets,

respectively. The voluminous distal pocket is located between the β -sheets of PC1 and PN2 in the pore domain of the binding protomer (Figure 4.0.4.2). The distal binding pocket contains multiple binding sites and different sets of residues are involved in binding of different substrates. For example, doxorubicin interacts with Gln176, Phe615 and Phe617 (Figure 4.0.4.4C), while minocycline contacts Phe178, Asn274 and Phe615 (Figure 4.0.4.4D). The presence of such multi-site binding pocket in AcrB allows recognition of a wide range of substrates (Murakami *et al.*, 2006; Seeger *et al.*, 2006). The proximal binding pocket for rifampicin and erythromycin resides between the β -sheets of the PC1 and PC2 subdomains and is located in the drug transport channel between the entrance(s) and the binding pocket for LMMDs. The proximal binding pocket is also a multi-site binding pocket, which uses different sets of residues to coordinate different substrates (Figure 4.0.4.4, A and B). All the three putative entrances described above are merged at the proximal binding pocket. The two pockets are separated by a switch loop at the position F617. In the access protomer, the proximal binding pocket is larger than the distal binding pocket, allowing multisite binding of the HMMDs rifampicin and erythromycin. Upon transition of the access state to the binding state, the F617 loop swings, which allows translocation of HMMDs from the proximal pocket to the distal pocket. Simultaneously, the proximal binding pocket shrinks and the distal binding pocket expands (Nakashima *et al.*, 2011). The importance of the F617 loop is verified by double proline mutations (G614P-G621P and G616P-G619P) and the G616N mutation (Cha *et al.*, 2014) that completely abolished or significantly reduced the substrate transport activity. Crystal structures have revealed that these mutations fix the switch loop at a point between the access state and the binding state (Nakashima *et al.*, 2011). Site-directed mutagenesis and doxorubicin efflux assays show that LMMDs and HMMDs are transferred through both pockets (Nakashima *et al.*, 2011). However, the proximal pocket is activated in the access state and the distal pocket is activated in the binding state (Nakashima *et al.*, 2011). The presence

of two voluminous multisite binding pockets and three entrances significantly expands the substrate spectrum of AcrB.

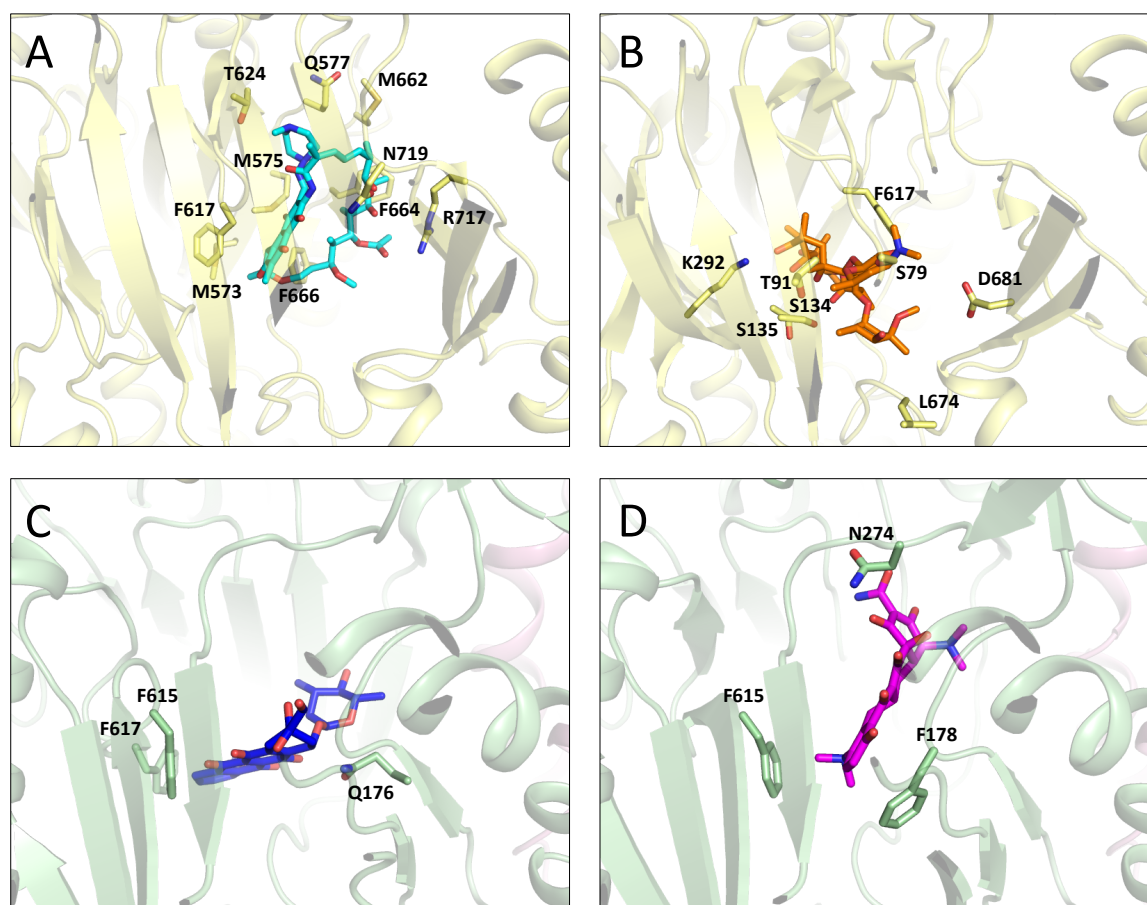


Figure 4.0.4.4. Proximal and distal binding pockets of AcrB. Binding site of rifampicin (cyan; PDB ID: 3AOB, (Nakashima *et al.*, 2011)) (A), erythromycin (orange; PDB ID: 3AOC, (Nakashima *et al.*, 2011)) (B), doxorubicin (blue; PDB ID: 2DR6, (Murakami *et al.*, 2006)) (C), and minocycline (magenta; PDB ID: 2DRD, (Murakami *et al.*, 2006)) (D). Carbon atoms from the access and binding monomers are coloured in yellow and green, respectively. Sulfur, oxygen and nitrogen atoms are coloured in gold, red and blue, respectively.

In summary, through inter-dependent sequential conformational transitions between the access, binding to extrusion monomers, AcrB harnesses the proton motive force to export substrates from the inner-membrane or periplasmic entrance to the extracellular medium through the TolC channel (Murakami *et al.*, 2006; Seeger *et al.*, 2006). Substrates enter the proximal pocket in

the access state (Nakashima *et al.*, 2011). The proximal pocket recognises and binds HMMDs but not LMMDs. Upon transition from the access to binding state, substrates are translocated to the distal pocket via swing of the F617 loop and relative movement between the PC1, PC2 and PN2 subdomains, which results in the expansion of the distal binding pocket and shrinking of the proximal binding pocket. The distal binding pocket recognises and binds LMMDs but instead occludes HMMDs due to the small volume of the proximal binding pocket. Finally, during conformational change to the extrusion monomer, substrates are squeezed into the funnel-like opening at the docking domain and sequentially through the TolC channel. This transport process is known as the functional rotation mechanism (Figure 4.0.4.5) (Murakami *et al.*, 2006; Seeger *et al.*, 2006).

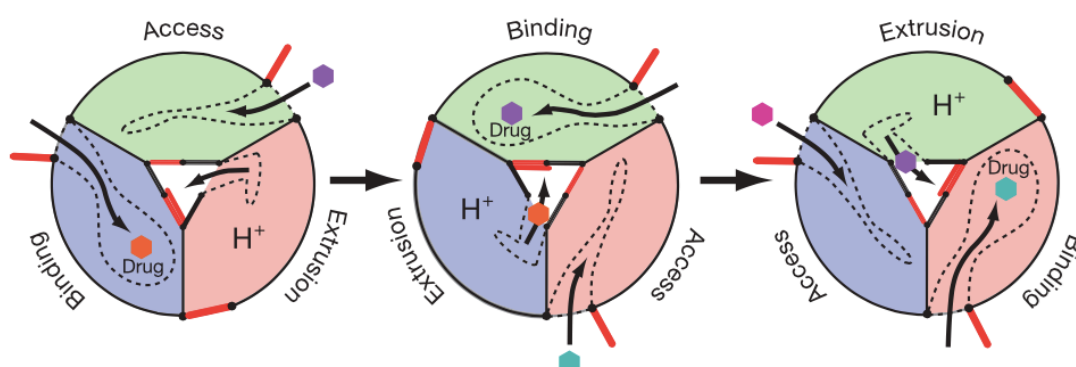


Figure 4.0.4.5. Functional rotation mechanism mediated by AcrB. In the access state, the entrances are open for substrate entry. In the binding state, substrates bind to the expanded distal binding pocket. In the extrusion state, the distal binding pocket shrinks, the exit opens, and drugs are squeezed into the central funnel and TolC. Conversion from the extrusion to the access conformation is the proton motive force-dependent step, which requires binding of a proton from the periplasm to the proton translocation site (Asp407, Asp408, and Lys940) (Seeger *et al.*, 2006). Conversion from the binding to the extrusion state involves release of a proton from the proton translocation site to the cytoplasm. The green, blue and red colours depict three different monomers of AcrB and their conformational changes during the drug transport cycle. Taken from (Murakami *et al.*, 2006) with permission.

4.0.5. TolC

TolC is an outer membrane efflux channel, members of which function in conjunction with three types of transport systems (RND-, ABC- and MFS-type) (Paulsen *et al.*, 1997). TolC and homologues are involved in efflux of large proteins (type I secretion system), such as enzymes and toxins, as well as small toxic molecules, including antibiotics, detergents and solvents. The association between TolC and inner-membrane transporters is mediated through members of periplasmic adaptor fusion proteins (Zgurskaya, 2009).

Extending from the extracellular space into the periplasm, the homotrimer TolC is a 140-Å-long tapered cylinder comprising a 40-Å-long β -barrel embedded in the outer membrane (the channel), a 100-Å long α -helical barrel located in the periplasm (the tunnel), and a mixed α/β equatorial domain that forms a belt at the equator of the α -helical barrel (Figure 4.0.5.1) (Koronakis *et al.*, 2000). The overall cavity of TolC encloses a solvent-exposed region of $\sim 43,000 \text{ \AA}^3$ and can accommodate substrates as large as 160 kDa. The β -barrel of TolC is composed of 12 β -strands (four antiparallel β -strands from each monomer, $\beta 1$, $\beta 2$, $\beta 4$, $\beta 5$), while β -barrels of other outer membrane proteins are often formed by a single protomer (Locher *et al.*, 1998; Pautsch & Schulz, 1998; Buchanan *et al.*, 1999). The extracellular exit of TolC is open towards the external medium. Unlike other outer membrane proteins, whose interior channels are typically either partially or entirely occluded (Locher *et al.*, 1998; Pautsch & Schulz, 1998; Buchanan *et al.*, 1999), the β -barrel of TolC exhibits an open state and is solvent accessible (Koronakis *et al.*, 2000). This is due to the lack of either an inward-folded loop or a plug protein as observed in OmpF and FhuA, respectively, that restricts the β -barrel (Saint *et al.*, 1996; Locher *et al.*, 1998). In TolC, the three monomers each contribute four α -helices to form the 12-antiparallel-helix α -barrel. The α -helical domain of each TolC protomer comprises two long helices that extend 100 Å from the bottom end of the β -barrel into the

periplasm ($\alpha 3$ and $\alpha 7$), and two pairs of shorter helices that form pseudocontinuous helices by end-to-end tacking interactions ($\alpha 2$ and $\alpha 4$; $\alpha 6$ and $\alpha 8$). The stabilisation of the barrel is proposed to be achieved by a “knobs-into-holes” packing (Koronakis *et al.*, 2000), where side chains of each helix intermesh laterally with the neighbouring helices. Throughout the entire α -barrel, the helices follow a left-handed superhelical twist, while exhibit a partial untwist at the β -barrel-proximal end compared to conventional coiled coil helices. In the β -barrel-distal end, each monomer contributes an inward-folded coil, forming a resting “closed” state at the periplasmic entrance. During drug export, this channel opens for substrate entry. The allosteric opening mechanism envisages that the three inner-coiled coils (comprising α -helices $\alpha 7$ and $\alpha 8$) undergo an iris-like movement to realign with the outer-coiled coils (comprising α -helices $\alpha 3$ and $\alpha 4$), thereby enlarging the pore in a “twist-to-open” transition (Andersen *et al.*, 2002).

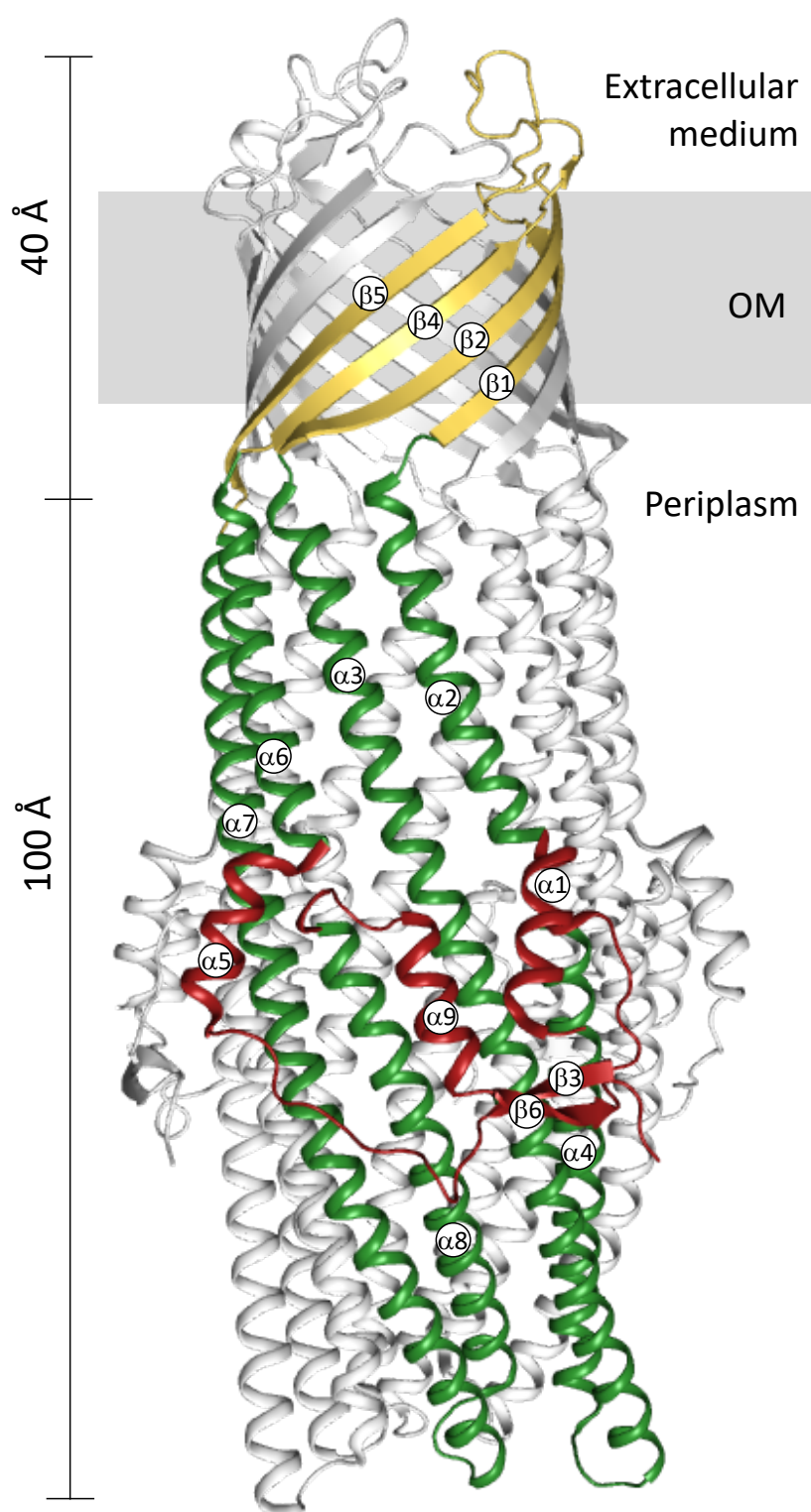


Figure 4.0.5.1. Structure of the TolC trimer. A TolC protomer is coloured with the β -barrel domain in yellow, the α -helical domain in green and the α/β equatorial domain in red. The other two protomers are shown in grey. Secondary structure elements are indicated: α -helices (α) and β -strands (β). (PDB ID: 1EK9, (Koronakis *et al.*, 2000))

4.0.6. AcrA

A third essential component of the RND efflux systems is a periplasmic membrane MFP (Dinh *et al.*, 1994). MFPs function mainly as adaptors to connect the inner-membrane transporters of the Gram-negative bacteria with the OMF proteins. Due to the tripartite composition, RND transporters pump out drugs directly into the extracellular medium, bypassing the periplasm. MFP proteins are also essential components in certain ABC transporter (*e.g.* MacAB) and MFS transporter (*e.g.* EmrAB and EmrKY) tripartite efflux pumps (Kobayashi *et al.*, 2001; Furukawa *et al.*, 1993; Kato *et al.*, 2000).

AcrA forms the periplasmic MFP component in the AcrABZ-TolC multidrug efflux system (Ma *et al.*, 1995). Structure of the stable core of AcrA (residues 45-312) has been determined by X-ray crystallography with protease digestion to trim the flexible N- (residues 26-44) and C-termini (residues 313-397) (Figure 4.0.6.1) (Mikolosko *et al.*, 2006). Later cryo-EM structures of the AcrAB(Z)-TolC system have revealed the overall shape of the pump and the detailed interactions between the components, as well as the structure of full-length AcrA (Du *et al.*, 2014; Wang *et al.*, 2017). The elongated structure of full-length AcrA consists of four types of domains: a coiled coil α -helical hairpin formed of two α -helices at the TolC-proximal end, a centrally located lipoyl domain, a β -barrel domain, and a membrane-proximal domain at the AcrB-proximal end (Figure 4.0.6.1 and Figure 4.0.6.2) (Mikolosko *et al.*, 2006; Symmons *et al.*, 2009; Du *et al.*, 2014; Wang *et al.*, 2017).

In AcrA, the coiled coil α -helical hairpin comprises two α -helices with five heptad repeats per helix, packed together by hydrophobic side chains (Mikolosko *et al.*, 2006). The cryo-EM structures show that the α -helical hairpins of AcrA interact in a tip-to-tip manner with the α -helical coiled coils of TolC in the periplasm (Figure 4.0.6.2) (Du *et al.*, 2014; Wang *et al.*,

2017). Presence of this 58-Å long intervening α -helical hairpin separates the lipoyl domain, which can be subdivided into two lipoyl half-motifs each containing four β -strands (β 2- β 5 and β 6- β 9, respectively) that intertwine to form a β -sandwich structure. The lipoyl domains do not make direct contact with either AcrB or TolC (Du *et al.*, 2014; Wang *et al.*, 2017). The β -barrel domain, composed of six antiparallel β -strands and a short α -helix (α 3), interacts with the docking domain of AcrB (Figure 4.0.6.2) (Wang *et al.*, 2017). The short helix α 3 lies proximal to a flexible C-terminal loop region (residues 313-397) that forms the membrane-proximal domain and interacts with both the docking domain and the top of the pore domain of AcrB (Figure 4.0.6.2) (Wang *et al.*, 2017).

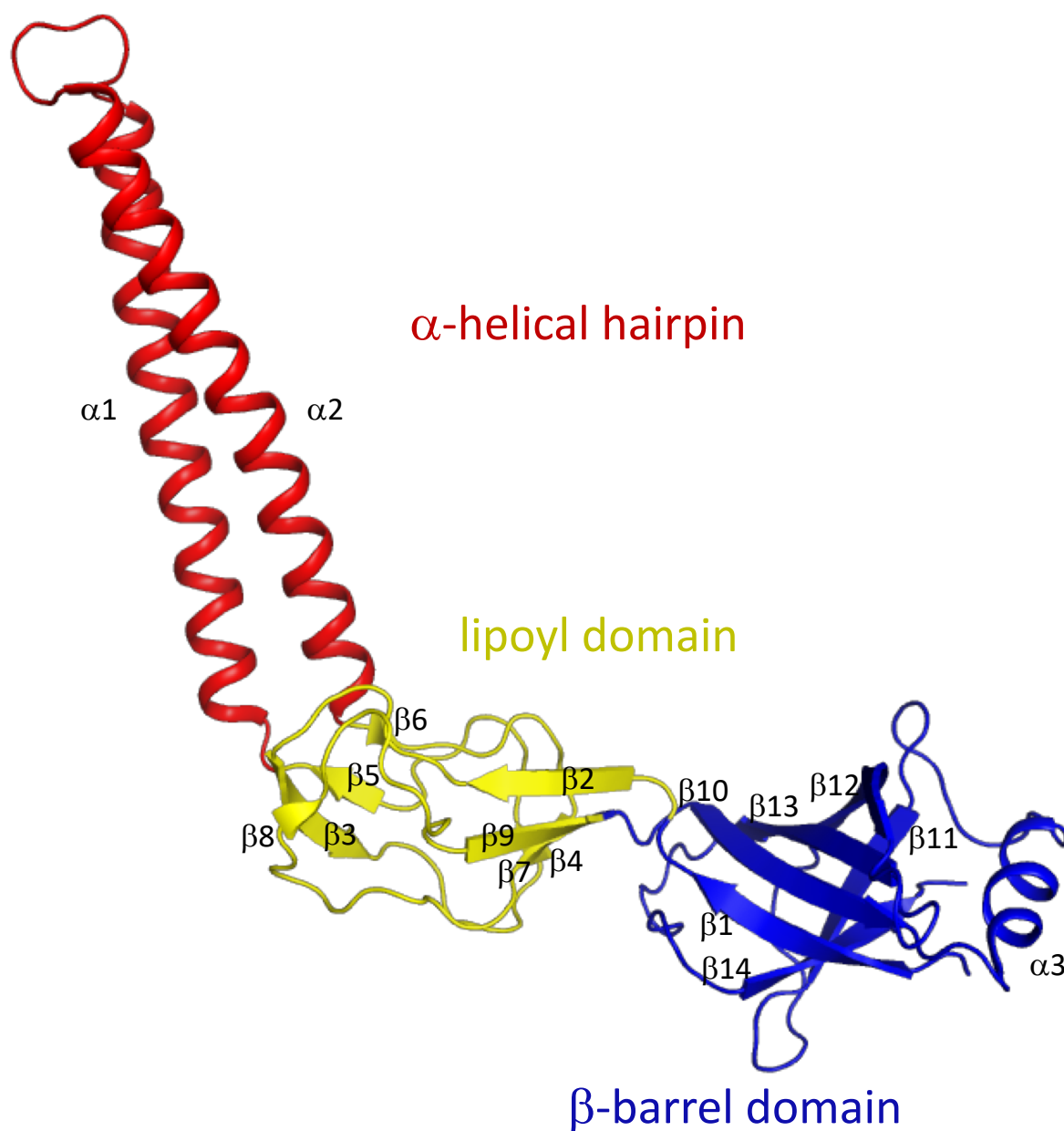


Figure 4.0.6.1. Structure of the AcrA monomer. The α -helical hairpin domain, the lipoyl domain and the β -barrel domain are coloured in red, yellow and blue, respectively (PDB ID: 2F1M, (Mikolosko *et al.*, 2006)). Secondary structure elements are indicated: α -helices (α) and β -strands (β).

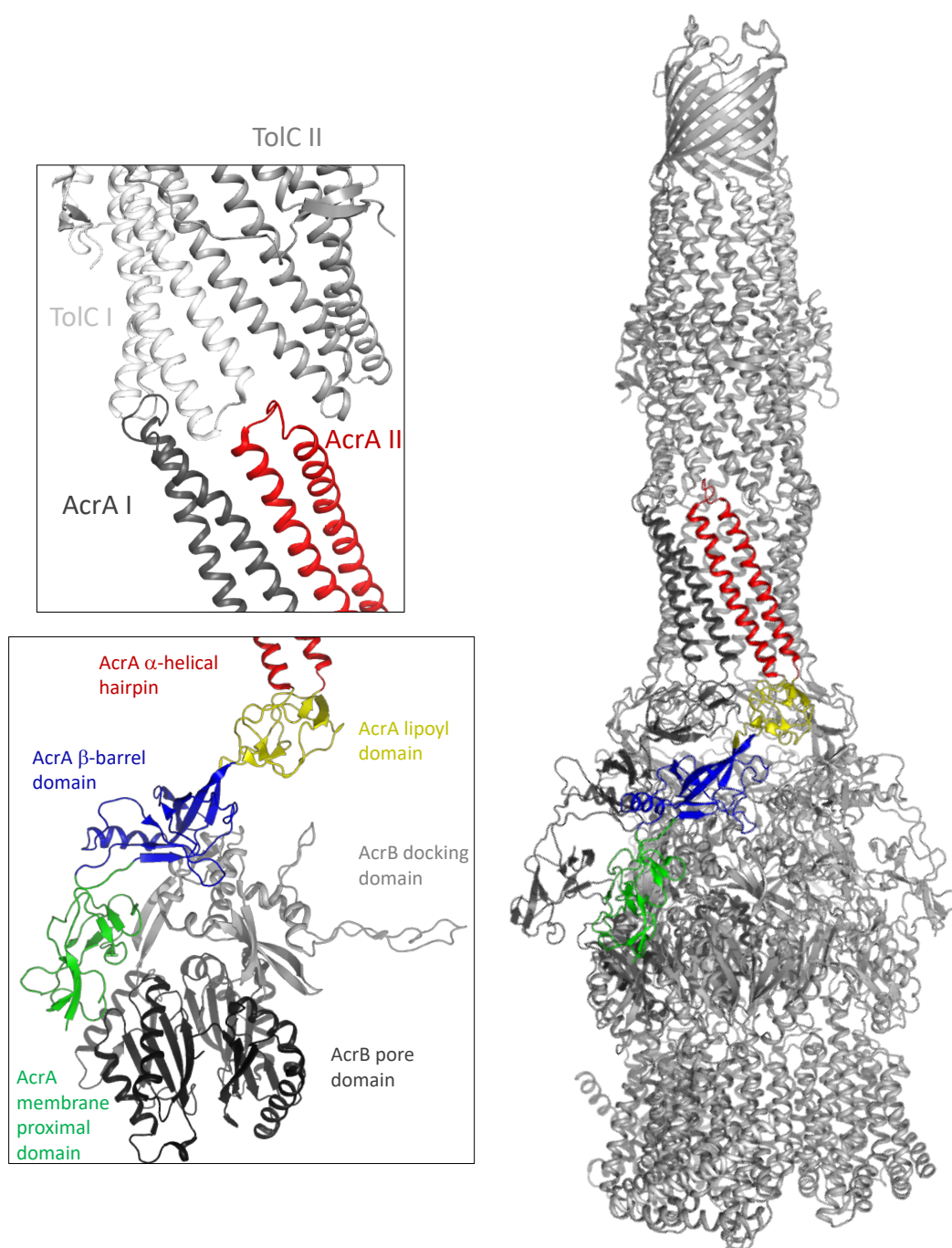


Figure 4.0.6.2. Interactions between components of the AcrA₆B₃-TolC₃ complex. Top inset: the tip-to-tip interactions between two protomers of TolC (white for TolC I and light grey for TolC II) and two α -helical hairpin domains of AcrA (grey for AcrA I and red for AcrA II). Bottom inset: the contacts between an AcrA protomer and an AcrB monomer. The β -barrel domain (blue) interacts with the docking domain of AcrB (light grey). The membrane-proximal domain (green) interacts with both the docking domain and the top of the pore domain (grey) of AcrB. The lipoyl domains (yellow) and the α -helical hairpin domains do not make direct contact with AcrB (PDB ID: 5O66, (Wang *et al.*, 2017)).

4.0.7. AcrZ

AcrZ is an inner-membrane protein of 49 amino acid in length and is highly conserved among Enterobacteria and other Gram-negative strains (Hobbs *et al.*, 2012). AcrZ associates with the AcrAB-TolC efflux pump via direct interactions with the inner-membrane protein AcrB (Figure 4.0.2.1; Figure 4.0.7.1). *acrZ* is located between *modEF* and *modABC*, the two operons responsible for uptake of molybdenum and synthesis regulation of molybdopterin, the molybdenum cofactor. Although *acrZ* gene exhibits high synteny with molybdenum operons, level of molybdenum does not affect AcrZ expression (Hobbs *et al.*, 2012). *acrZ* is co-regulated with genes encoding AcrA, AcrB and TolC by transcription factors MarA, Rob and SoxS, which regulate cellular responses to antibiotics, detergents, and redox-cycling drugs (Hobbs *et al.*, 2012). Competition assays show that Δ AcrZ cells are sensitive to a subset of AcrB substrates, such as chloramphenicol, puromycin and tetracycline (Hobbs *et al.*, 2012), while the Δ AcrB mutant almost loses its antibiotic resistance (Sulavik *et al.*, 2001). Hobbs *et al.* proposed that AcrZ binding to AcrB may initiate conformational alternations in AcrB by allosteric modulation, thus affecting the recognition and binding of certain substrates (Hobbs *et al.*, 2012; Du *et al.*, 2014). Sequence alignment shows that residues of AcrZ that directly contact AcrB are highly conserved among different *Enterobacteriaceae* species, suggesting that similar interactions are likely to be observed in other efflux pumps (Du *et al.*, 2014).

AcrZ is a hydrophobic α -helical peptide that enters a wide groove in the transmembrane domain of AcrB (Figure 4.0.7.1) (Du *et al.*, 2014). AcrZ expands the inner membrane, with its N-terminus located in the periplasm and C-terminus in the cytoplasm, contacting the hydrophilic surface of AcrB. The helical axis of AcrZ is oriented by nearly 45° with respect to the normal vector of the lipid bilayer, which promotes its interfacial complementarity with AcrB. Other examples of membrane-associated 45°-tilted α -helical proteins are provided by

structures of the AcrB-YajC complex and the SecY channel (Törnroth-Horsefield *et al.*, 2007; Li *et al.*, 2016). The potential AcrB partner YajC from the AcrB-YajC complex and the *in vivo* inserted signal sequence from the SecY channel both tilt $\sim 45^\circ$ relative to the plane of the membrane. A common characteristic of these tilted α -helices is amphipathicity, with the hydrophilic residues interacting with the negatively charged head groups of lipid bilayer and the lipophilic residues inserting into phospholipid tails. Another example of a titled α -helix is the a-subunit in F-type ATP synthase (Allegretti *et al.*, 2015).

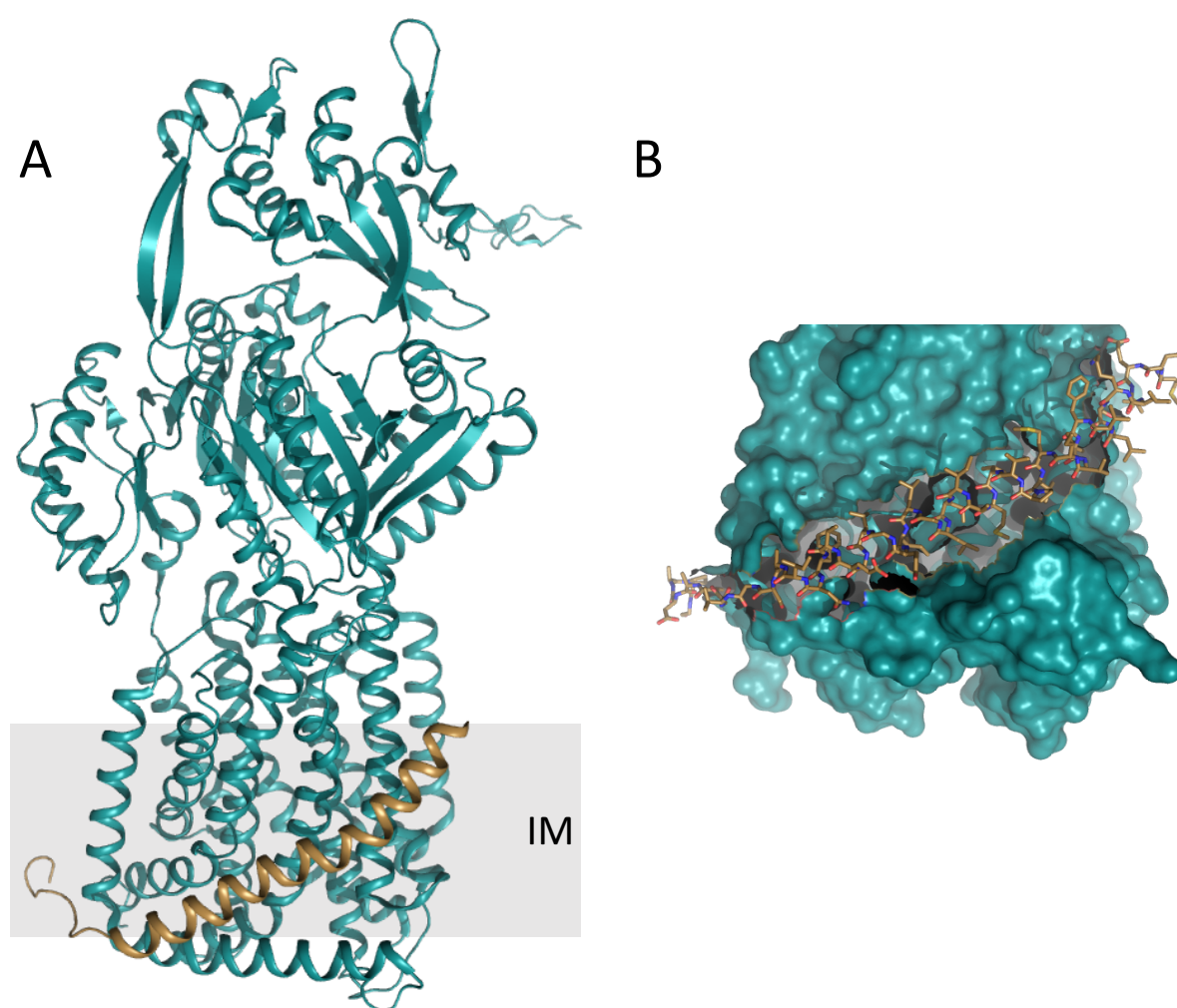


Figure 4.0.7.1. Interactions between AcrB and AcrZ. (A) Structure of AcrB (cyan) in complex with AcrZ (sand). The protein was incubated with DARPin prior to crystallisation. DARPin is bound to AcrB monomer at 1:1 ratio in the crystallographic asymmetric unit. DARPin molecule is not shown for clarity. (B) AcrZ fits into a wide groove in the transmembrane domain of AcrB. (PDB ID: 4C48, (Du *et al.*, 2014))

4.1. Efflux pump inhibitor development

4.1.1. Efflux pump inhibitors

Nowadays, two main strategies have been employed to overcome bacterial antibiotic resistance mediated by efflux pumps. One strategy aims to develop antimicrobials that can resist being exported by efflux pumps through optimising existing antibiotics or by designing new ones. However, resistance has occurred shortly after deployment of new molecules (Nikaido & Pagès, 2012). Another strategy consists of inactivating efflux pumps, especially RND pumps, one of the major reasons for bacterial antibiotic resistance (Zechini & Versace, 2009). Three possible modes of efflux pump inactivation have been proposed, which include (1) inhibitors of the proton motive force, (2) biological inhibitors and (3) efflux pump inhibitors (EPIs). Proton motive force inhibitors, such as dinitrophenol (DNP) and carbonylcyanide *m*-chlorophenylhydrazone (CCCP), inactivate efflux pumps by disrupting the electrochemical gradient across the cell membrane (Zechini & Versace, 2009; Pagès *et al.*, 2005; Mahamoud *et al.*, 2007). These compounds can effectively abolish efflux of all toxic compounds mediated by secondary transporters. Therefore, these inhibitors show severe side effects, since they generally affect both bacterial and host cells, limiting their clinical application (Zechini & Versace, 2009). Biological inhibitors include small interfering RNAs that interfere with the pump-encoding genes, neutralising antibodies that interact with and block the efflux pumps or other antisense molecules that inactivate gene transcription or translation. EPIs are therapeutic agents designed to circumvent antibiotic resistance by inhibiting activity of efflux pumps and therefore increasing intracellular antibiotic concentration (Lomovskaya & Bostian, 2006). An EPI in combination therapy is expected to potentiate the antibacterial activity of conventional antibiotics (Pagès & Amaral, 2009).

4.1.2. EPIs against the AcrAB-TolC system

The AcrAB-TolC (or AcrAB-TolC-like) pump has been characterised as a major mechanism of MDR in Gram-negative pathogens isolated from antibiotic treatment (Piddock, 2006; Li & Nikaido, 2009). This family of transporters show low homology between bacterial and human genomes (Tseng *et al.*, 1999). Although RND transporters exhibit a broad spectrum of substrates, these proteins recognise substrates through specific mechanisms and are therefore inhibited differently (Yamaguchi *et al.*, 2015). These features of RND transporters make them an attractive target for design of EPIs. The resulting EPIs, when applied in combination therapy, are expected to demonstrate improved potency, a broadened spectrum of inhibition and reduced emergence of acquired resistance (Mahamoud *et al.*, 2007).

AcrB is the best characterised example of an RND transporter with its structures determined in different conformations and its substrate efflux mechanism well elucidated (Murakami *et al.*, 2002; Murakami *et al.*, 2006; Seeger *et al.*, 2006; Nakashima *et al.*, 2011). Several functional inhibitors of AcrB have been identified through high-throughput screening assays (Lomovskaya *et al.*, 2001; Bohnert & Kern, 2005) and the structure of AcrB in complex with an inhibitor, ABI-PP, has been determined (Nakashima *et al.*, 2013). However, given the in-depth structural and functional studies of AcrB, to date no inhibitor against the AcrAB-TolC system or other efflux pumps has been clinically licensed as a cure of bacterial infections.

4.1.2.1. Peptidomimetic EPIs

Peptidomimetic EPIs include Phe-Arg- β -naphthylamide (PA β N) (Figure 4.1.2.1) and its derivative MC-04,124 (Lomovskaya *et al.*, 2001; Lomovskaya & Bostian, 2006). PA β N and derivatives are the first identified EPIs and have provided a valuable model for drug discovery. PA β N was identified by screening a variety of natural and synthetic compounds against *P. aeruginosa* overexpressing either of the three RND-type multidrug efflux pumps (MexAB-OprM, MexCD-OprJ, MexEF-OprN) in the presence of the quinolone antibiotic levofloxacin at subinhibitory concentrations (Lomovskaya *et al.*, 2001). It is proposed that PA β N is probably a competitive inhibitor that recognises and binds to the substrate-binding pocket of the RND transporter (Yu *et al.*, 2003; Yu *et al.*, 2005; Lomovskaya & Bostian, 2006). PA β N has also been shown to alter the integrity of the outer membrane and therefore to increase antibiotic permeation into the periplasm (Lomovskaya *et al.*, 2001). Secondary assays have shown that PA β N has a broad-spectrum inhibitory activity on efflux pumps and demonstrates significant potentiation on different classes of antibiotics including fluoroquinolones, chloramphenicol and macrolides (Lomovskaya *et al.*, 2001; Lomovskaya & Bostian, 2006). PA β N has significantly reduced the emergence of intrinsic and acquired resistance of *P. aeruginosa* (Lomovskaya *et al.*, 2001; Lomovskaya & Bostian, 2006). In addition to inhibiting the Mex RND pumps in *P. aeruginosa*, PA β N has also been shown to be a potent inhibitor against the AcrAB-TolC system from *E. coli* (Mazzariol *et al.*, 2000), *Enterobacter aerogenes* (Malléa *et al.*, 2002), *Klebsiella pneumoniae* (Hasdemir *et al.*, 2004), *S. typhimurium* (Baucheron *et al.*, 2002) and a range of different Mex homologues found in *A. baumannii* (Pannek *et al.*, 2006), *Campylobacter coli* and *Campylobacter jejuni* (Hannula & Hänninen, 2008). However, the high toxicity to human tissues has prevented the clinical application of PA β N and its derivatives (Lomovskaya & Bostian, 2006).

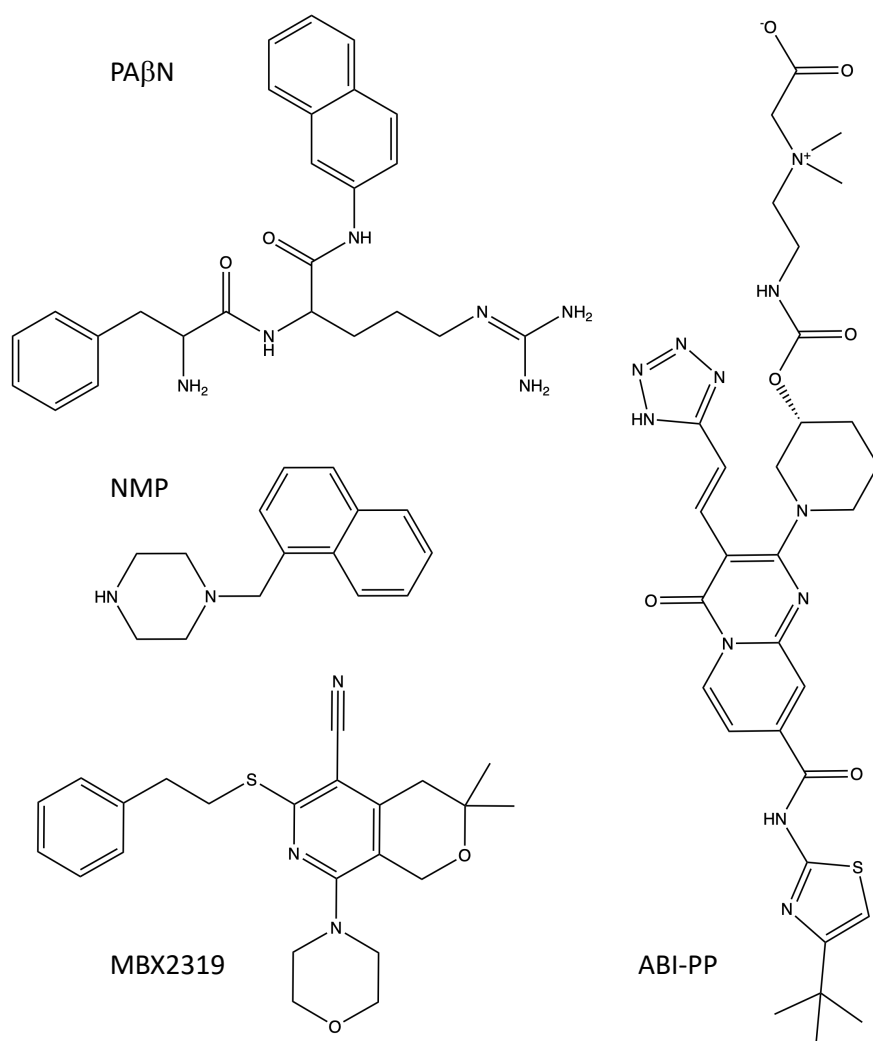


Figure 4.1.2.1. EPIs exemplified by PAβN, NMP, ABI-PP and MBX2319.

4.1.2.2. Piperazines

Screening of a library of organic N-heterocyclic compounds has resulted in identification of several arylpiperazines that can reverse MDR in *E. coli* overexpressing the RND-type efflux pumps. 1-(1-naphthylmethyl)piperazine (NMP) (Figure 4.1.2.1) is one of the most effective unsubstituted arylpiperazines, which demonstrates a dose-dependent enhancement of susceptibility to fluoroquinolones, chloramphenicol, tetracycline, linezolid and macrolides (Bohnert & Kern, 2005; Kern *et al.*, 2006). MIC tests on different bacterial species have shown that NMP is also a potent EPI in *A. baumannii* and several *Enterobacteriaceae* species, but not in *P. aeruginosa* (Bohnert & Kern, 2005; Kern *et al.*, 2006; Pannek *et al.*, 2006; Coban *et al.*, 2009). Due to the serotonin agonist properties, these arylpiperazines are possibly toxic for clinical application in humans. The mechanism of inhibition of NMP remains unknown.

4.1.2.3. Pyridopyrimidines

Unlike PA β N and NMP, another class of EPIs, pyridopyrimidine derivatives, exhibit low animal toxicity and low intrinsic antibacterial activity (Askoura *et al.*, 2011). These compounds are specific EPIs for the AcrAB-TolC and MexAB-OprM pumps, but they do not inhibit MexXY-OprM, which has hindered their clinical development (Nakayama *et al.*, 2003). Among these pyridopyrimidine derivatives, [[2-({[(3R)-1-{8-{{(4-tert-butyl-1,3-thiazol-2-yl)amino}carbonyl}-4-oxo-3-[(E)-2-(1H-tetrazol-5-yl)vinyl]-4H-pyrido[1,2-a]pyrimidin-2-yl}piperidin-3-yl)oxy]carbonyl}amino) ethyl](dimethyl) ammonio]acetate (referred to as ABI-PP) (Figure 4.1.2.1) demonstrates high *in vivo* efficacy, high water solubility and low animal toxicity (Yoshida *et al.*, 2007; Nakashima *et al.*, 2013). ABI-PP binds to AcrB and MexB, but not MexY (Nakashima *et al.*, 2013). ABI-PP exhibits significant inhibition towards binding of

LMMDs (*i.e.* doxorubicin and minocycline), but it has no effect on binding of HMMDs (*i.e.* erythromycin and rifampicin), which suggests that ABI-PP inhibits the AcrAB-TolC system by interacting with the distal binding pocket of AcrB. The presence of either drug does not affect ABI-PP binding to AcrB.

The structural basis for inhibition of RND exporters has been revealed by determination of the crystal structures of AcrB and MexB in complex with ABI-PP, the first physiologically relevant structures of an inhibitor-bound transporter. The binding sites and conformations of ABI-PP are similar in AcrB and MexB. Discussion here will be focused on structural analysis of the ABI-PP-bound AcrB structure. The binding site of ABI-PP is a phenylalanine-rich hydrophobic pocket that is located in the distal binding pocket of the binding monomer (Figure 4.1.2.2, A, B and C) (Nakashima *et al.*, 2013). The pyridopyrimidine ring and thiazolyl ring of ABI-PP bind tightly to residues F178 and F628, respectively, through π - π interactions (Figure 4.1.2.2C). The hydrophobic tail of ABI-PP inserts into the phenylalanine-rich pocket that branches off the substrate-translocation path of AcrB (Figure 4.1.2.2, D, E and F). The relatively hydrophilic head of ABI-PP superimposes with the binding site of doxorubicin and minocycline. The tight binding of ABI-PP to the hydrophobic trap sterically hinders the functional rotation of the monomer from the binding to the extrusion state. Therefore, ABI-PP can restore the activities of not only LMMDs but also HMMDs that are translocated by AcrB.

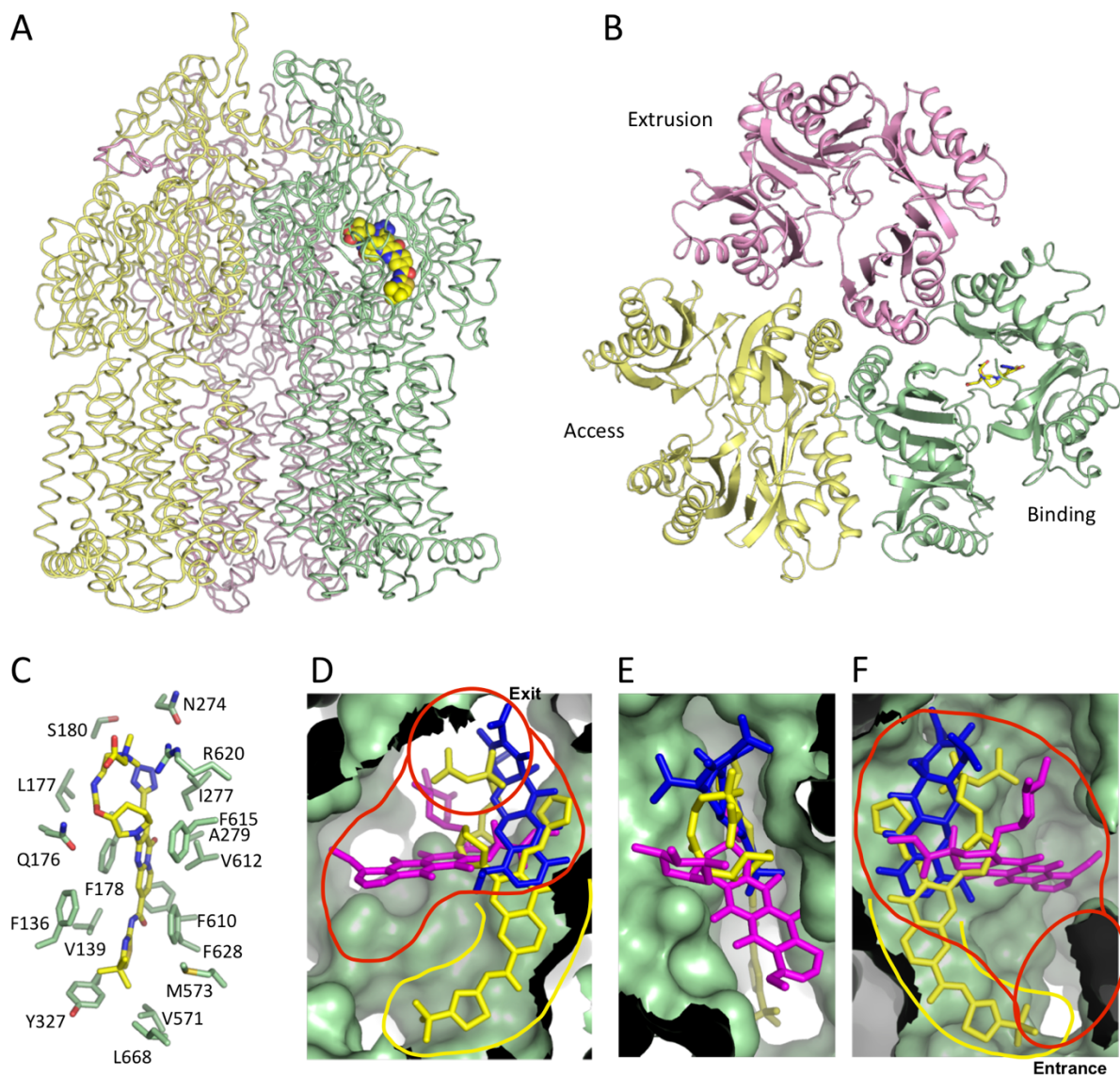


Figure 4.1.2.2. Crystal structure of AcrB in complex with ABI-PP. The access, binding and extrusion monomers are coloured in yellow, green and pink, respectively. (PDB ID: 3W9H, (Nakashima *et al.*, 2013)). (A) Side-view of the AcrB trimer in complex with ABI-PP. (B) Top view of the pore domain of AcrB in complex with ABI-PP. (C) Close-up view of the ABI-PP-binding site. Carbon atoms of ABI-PP and amino acid residues of AcrB are coloured in yellow and green. Nitrogen, oxygen and sulfur atoms are coloured in blue, red and gold, respectively. (D) Cut-away view of the distal binding pocket towards the exit. (E) The narrow hydrophobic pocket in the distal binding pocket (rotated 90° around the vertical axis from (D)). (F) Cut-away view of the distal binding pocket towards the entrance (rotated 90° around the vertical axis from (E)). The substrate translocation channel and the hydrophobic pocket are indicated by red and yellow curves, respectively. The bound doxorubicin (blue, PDB ID: 2DR6) and minocycline (magenta, PDB ID: 2DRD) are superimposed in the binding pocket (Murakami *et al.*, 2006).

4.1.2.4. Pyranopyridine derivatives

Pyranopyridine derivative MBX2319 is another potent EPI against *Enterobacteriaceae* (Opperman *et al.*, 2014). MBX2319 does not perturb cell membranes or exhibit intrinsic antimicrobial activity. More recently, derivatives of MBX2319 (MBX2931, MBX3132 and MBX3135) have been developed which demonstrate improved potency and stability (Nguyen *et al.*, 2015). Sjuts *et al.* determined high-resolution crystal structures of these inhibitors bound to the periplasmic domain of AcrB and hereby revealed the molecular mechanism that underpins the inhibitory effect (Sjuts *et al.*, 2016). The pyranopyridine inhibitors bind tightly to a hydrophobic phenylalanine-rich trap through hydrophobic interactions, which hinders substrate binding to AcrB and sterically prevents its functional rotation (Figure 4.1.2.3). The binding site of MBX inhibitors is similar to that of the pyridopyrimidine inhibitor ABI-PP, indicating that the hydrophobic pocket is a promising target for the development of EPIs.

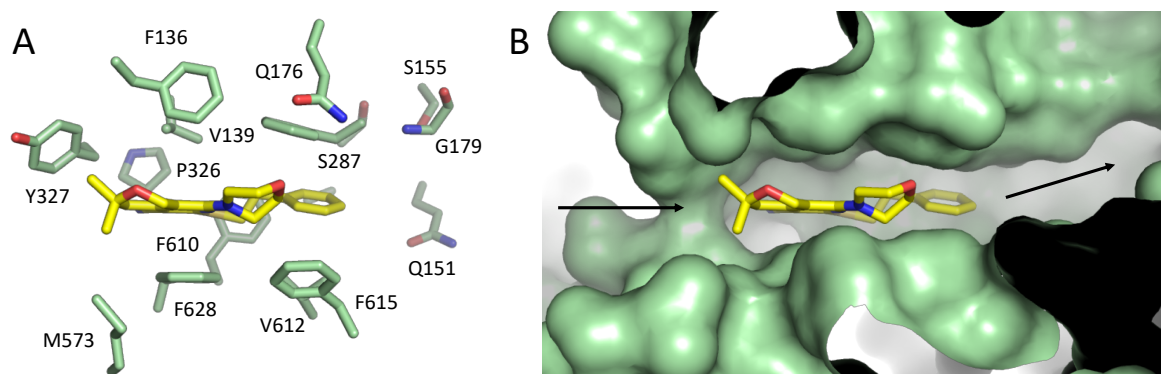


Figure 4.1.2.3. Binding of MBX2319 to AcrB periplasmic domain. Binding of MBX2319 is mainly mediated by hydrophobic stacking interactions with residues from the hydrophobic pocket. (A) Close-up view of the MBX2319-binding site. Carbon atoms of MBX2319 and residues of AcrB are coloured in yellow and green, respectively. Nitrogen, oxygen and sulfur atoms are coloured in blue, red and orange, respectively. (B) The narrow hydrophobic trap in the distal binding pocket. Arrows indicate the possible drug translocation pathway. (PDB ID: 5ENO, (Sjuts *et al.*, 2016))

4.1.3. Structure-based drug design

Structure-based drug design (SBDD) relies on the 3D structure of a biological target to assist development of novel drug molecules (Kuntz, 1992). SBDD is an iterative process and often involves multiple cycles to obtain a lead compound for phase I clinical trials. Once a suitable biological target has been identified, the target is cloned, overexpressed and purified for structure determination and biochemical assays. The first cycle aims to determine structures of the target proteins or nucleic acids by X-ray crystallography or NMR spectroscopy. If an experimental structure is not available, a homology model can be built based on structures of homologous proteins. Then libraries of druggable compounds or small fragments are docked into the potential drug binding site and scored based on their interactions with the active site. The best compounds are tested for binding affinity to the target with biochemical assays such as surface plasmon resonance (SPR). The second cycle aims to determine the structure of the target in complex with the resultant lead compound from the first cycle, which provides information for compound optimisation. In the final cycle, the optimised lead is synthesised, and the target-lead structure is determined for further optimisation. After iterative optimisation cycles, the obtained hits usually show an improved potency and specificity to the target. SBDD has been shown to be an efficient strategy for drug design and has witnessed great success over the recent years (*e.g.* Agenerase and Viracept against HIV protease (Kaldor *et al.*, 1997), Dorzolamide against carbonic anhydrase, (Grover *et al.*, 2006), Flurbiprofen against cyclooxygenase-2 (Dadashpour *et al.*, 2015)).

To date, no inhibitor against bacterial efflux pumps has been clinically licensed as a drug. In addition, a complete understanding of the molecular and conformational events that underpin drug efflux remains challenging. Development of novel EPIs and determination of high-resolution structures of inhibitor-bound transport proteins are urgently needed.

4.1.4. Aims and objectives

Inhibitors of multidrug efflux pumps can be used as an adjuvant therapy to raise the potency of clinically used antibiotics, hereby reducing emergence of MDR bacteria. The aim of the research presented here is to identify novel inhibitors of the AcrAB-TolC system and to determine structures of the protein in complex with inhibitors, hereby elucidating the mechanism of efflux inhibition and providing information for drug design.

The specific objectives of this research are as follows:

1. To identify potential EPIs of the AcrAB-TolC system using a fluorescence-based competition transport assay.
2. To validate the potentiation of the identified inhibitors on antibacterial effects of conventional antibiotics using MIC assays.
3. To purify and crystallise the AcrB protein and obtain structures of AcrB in complex with identified inhibitors.
4. To identify AcrB binders using the XChem fragment screening platform and develop AcrB inhibitors in light of the structures.

4.2. Materials and Methods

4.2.1. Doxorubicin efflux assay

Wild-type *E. coli* K-12 BW25113 (AcrB+) cells were plated on an LB-agar plate. The AcrB-deficient *E. coli* K-12 BW25113 (AcrB-) cells (Keio collection) were plated on an LB-agar plate containing 50 µg/mL kanamycin. Both plates were incubated for overnight at 37 °C. A single colony from each of the AcrB+ and AcrB- plates was picked and used to inoculate 50 mL of LB media (50 µg/mL final concentration of kanamycin for AcrB- culture). These cultures were grown at 37 °C with 220 rpm shaking overnight. Cells were then harvested by centrifugation at 3,000 g at 25 °C. The pellet was washed twice with PBS buffer (pH 7.4) containing 5 mM MgSO₄ and resuspended in the same buffer to achieve an OD₆₀₀ of ~18.

The doxorubicin efflux assay was performed in 96-well black clear-bottom plates with a final cell volume of 100 µL and a final doxorubicin concentration of 28.6 µM. The test compounds were dissolved and diluted to 10 mM in 100% dimethyl sulfoxide (DMSO). For each assay, 100 µM final concentration of the test compounds was added, resulting in a final DMSO concentration of 1%. The fluorescence of doxorubicin (excitation = 480 nm, emission = 600 nm) was measured every five minutes for a total of 90 minutes at 37 °C using a SpectraMax microplate spectrofluorometer. Due to the accumulation of doxorubicin in *E. coli* cells, the fluorescence emission of doxorubicin gets quenched in a time-dependent manner. The decrease in fluorescence intensity was prevented when doxorubicin was extruded to the extracellular medium by an efflux pump.

The doxorubicin efflux assay was used to screen a subset of the NCI diversity set V library (Table 4.2.1.1) for inhibitors of the AcrAB-TolC complex. The NCI diversity set V library was provided by the Chemotherapeutic Agents Repository, National Cancer Institute (NCI), USA. It contains 1597 compounds with the MWs ranging between 100-770 Da. A primary screen of 644 compounds with MWs > 250 Da was performed by intern students Frances Thomas and Chengzhi Guo, under the supervision of Feng Qu. 12 compounds were identified that demonstrated an enhancement of doxorubicin accumulation.

The doxorubicin efflux assay on the identified 12 compounds was repeated by Feng Qu. Each compound was tested in triplicate with cells from three different colonies. Data are represented as the mean and standard deviation.

Library	NCI diversity set V	DSi-Poised
NO. of compounds	1597	769
Stock concentration (mM)	10/20	500
MW range (Da)	100 - 770	109.13 - 334.35
Usage	Doxorubicin efflux assay	XChem fragment screening

Table 4.2.1.1. Libraries used in chapter 4.

4.2.2. Antibacterial activity assays

4.2.2.1. MIC determination of antibiotics

The MICs of the test antibiotics were determined using the broth microdilution method in accordance with the 2012 guideline of the CLSI protocol (M07-A9) (Clinical and Laboratory Standards Institute) (Waitz, 1990). *E. coli* K-12 BW25113 cells overexpressing the AcrAB-TolC efflux pump was used as the test strain. The culture was sub-cultured from a glycerol stock to LB-agar plate 24 h prior to any antimicrobial test. Firstly, dH₂O, DMSO or ethanol were tested against *E. coli* K-12 BW25113 cells to ensure that the solvents did not contribute to the antimicrobial activities of the test compounds. A final concentration of 1% dH₂O, DMSO or ethanol did not show inhibition of bacterial growth, therefore a 1% final concentration of solvent was applied for the antimicrobial test. Serial 2-fold dilutions of different antibiotic stocks (100x) were prepared in the appropriate solvent (dH₂O, DMSO or ethanol) according to the manufacturer's recommendations and dispensed onto a Nunc 384-well polypropylene storage microplate (Thermo Fisher Scientific).

A bacterial suspension was prepared by transferring 3-5 morphologically similar colonies from an LB-agar plate to 5 mL of cation adjusted Mueller-Hinton broth 2 (CA-MHB2) (Millipore). The culture was incubated at 37 °C until it achieved a turbidity of the 0.5 McFarland standard (1×10^8 colony forming units (CFU)/mL). Within 15 minutes of preparation, the 0.5 McFarland suspension was diluted 20-fold in CA-MHB2 to yield a 5×10^6 CFU/mL inoculum. The diluted cell inoculum was subsequently diluted 10-fold to yield a 5×10^5 CFU/mL cell suspension. 1 μ L of the antibiotic stock (100x) was added into 99 μ L of CA-MHB2 containing cell suspension (Figure 4.2.2.1). The assay was set up in clear 96-well sterile plates (Greiner Bio-One) with 100 μ L final volume per well. Columns 1-11 contained cell suspension and the

antibiotic with 2-fold dilutions, while column 12 was filled with only cell suspension and 1% solvent as a negative control. Different rows contained arrays of different antibiotics. The inoculated microdilution tray was incubated at 37 °C for 18 hours in an ambient air incubator. The lowest concentration of the antibiotic that prevented visible bacterial growth (either turbidity or precipitation) was defined as the MIC.

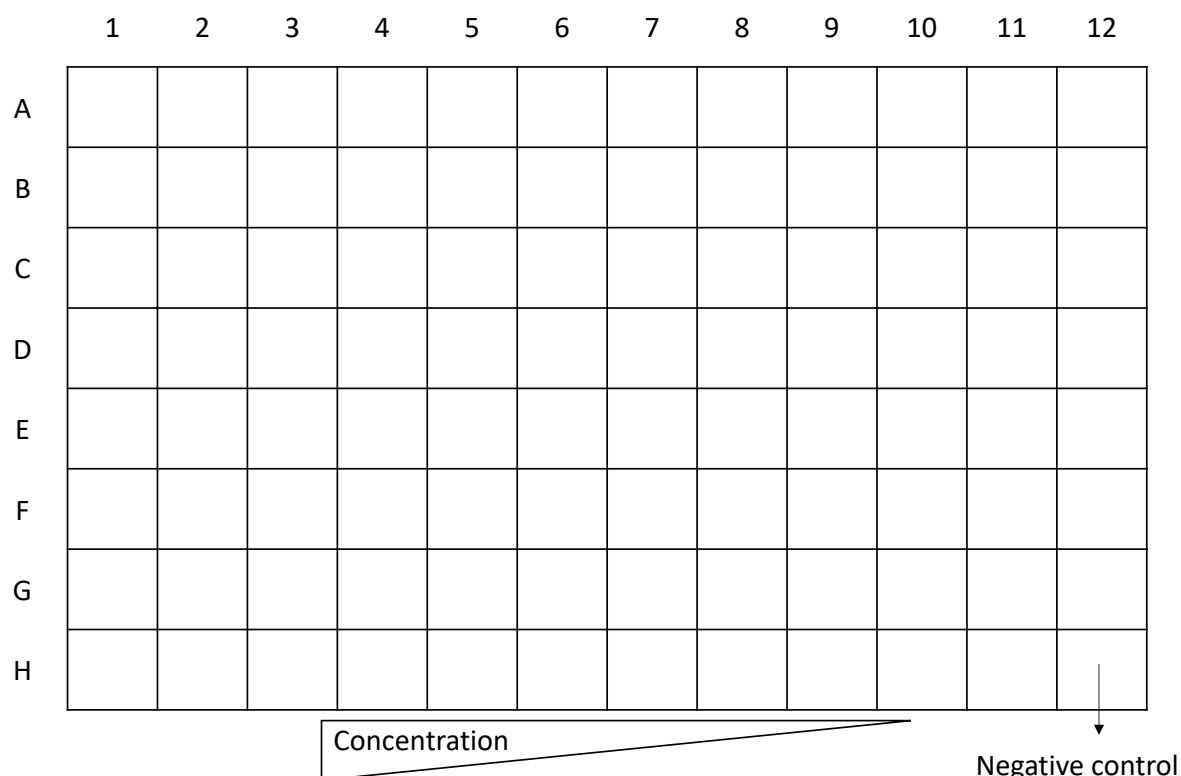


Figure 4.2.2.1. Layout of the MIC assay using the broth microdilution method. The triangle indicates the direction of the 2-fold serial dilutions of the test compound. The negative control contains cell culture with 1% dsH₂O/DMSO/ethanol.

4.2.2.2. MICs of antibiotics in the presence of an inhibitor

In order to verify the clinical potentials of the identified inhibitors, MICs were tested against antibiotics alone and in the presence of inhibitors.

According to the manufacturer's recommendations, dH₂O, DMSO or ethanol were used as solvents to dissolve the test compounds at the required concentrations. The intrinsic antibacterial activity of the six identified inhibitors and PA β N (a known EPI as positive control) was tested initially at a final concentration of 100 μ M. MICs of those compounds that prevented cell growth at 100 μ M were further determined by the 2-fold microdilution method. Then, MICs of different antibiotics in the presence of the identified inhibitors/PA β N were tested. A final concentration of 100 μ M was used for compounds that demonstrated MICs higher than 100 μ M, while for compounds whose MICs were equal to or lower than 100 μ M, a final concentration of 25 μ M was used. MICs were tested similarly to the protocol described in section 4.2.2.1, except that the antibiotic stock was diluted into CA-MHB2 containing cell suspension and the test compound.

4.2.3. AcrB protein production

4.2.3.1. AcrB overexpression

The gene fragment encoding full-length AcrB of *E. coli* (accession number: P31224, residues 1-1049) was previously cloned from *E. coli* chromosomal DNA into a pET-28b vector, resulting in a construct of AcrB protein with a C-terminal His₆-tag.

The AcrB-His₆ plasmid was transformed into *E. coli* C43 (DE3) cells, plated on LB-agar plate supplemented with 50 µg/mL kanamycin and incubated overnight at 37 °C. A single colony was picked and used to inoculate 100 mL of LB media supplemented with 50 µg/mL kanamycin. The starter culture was grown at 37 °C with 220 rpm shaking overnight. 10 mL of the overnight starter culture was used to inoculate 1 L of LB media supplemented with 50 µg/mL kanamycin. The culture was incubated at 37 °C with 220 rpm shaking. IPTG was added to a final concentration of 1 mM when the culture reached an OD₆₀₀ of ~0.6, and cell growth continued at 18 °C for 18 hours. Cells were harvested by centrifugation at 6,000 g for 10 minutes and stored at -80 °C until use.

4.2.3.2. AcrB membrane preparation

The whole cell pellet was resuspended in buffer containing 20 mM Tris-HCl (pH 7.5), 300 mM NaCl, PefaBloc (0.1 mg/mL), deoxyribonuclease (50 unit/mL) and 3 mM MgCl₂. The cell resuspension was passed through a cell disruptor for two cycles (22/25 kpsi), followed by centrifugation at 15,000 g for 10 minutes to remove unbroken cells and cellular debris. During cell breakage, the sample was kept on ice. Supernatant containing cell membranes was

recovered and subjected to an ultracentrifuge step at 180,000 g for 1 hour at 4 °C. Membrane pellets were resuspended in a buffer containing Tris-HCl (pH 7.5), 300 mM NaCl and PefaBloc (0.1 mg/mL) using a Dounce glass homogeniser (Fisher Scientific) on ice.

4.2.3.3. AcrB purification

The membranes were diluted in extraction buffer (Table 4.2.3.1) and the AcrB protein was solubilised from total membranes using 1% (w/v) DDM with 150 rpm stirring for 1 hour at 4 °C. After an ultracentrifugation step (180,000 g for 1 hour at 4 °C), supernatant containing AcrB was recovered for further purification. For the IMAC step, the AcrB protein was incubated with 10 mL of Ni-NTA SuperFlow resin and 50 mM imidazole (pH 7.5), stirring gently at 150 rpm for 3 hours at 4 °C. The protein-nickel slurry was then loaded onto a 50-mL glass polyprep column and washed with 10 CV of wash buffer A (Table 4.2.3.1), followed by protein elution with 5 CV of elution buffer (Table 4.2.3.1). The affinity chromatography was performed at 4 °C. The eluate was dialysed overnight at 4 °C in 2 L of dialysis buffer (Table 4.2.3.1). The dialysed sample was concentrated to ~28 mg/mL (250 µM) using a 100 kDa MW cut-off concentrator.

Different detergents such as Dodecanoyl Sucrose (DDS) (Anatrace) (Murakami *et al.*, 2006) and 6-Cyclohexylhexyl β-D-maltoside (Cymal6) (Anatrace) (Pos & Diederichs, 2002) were also used for purification. AcrB was extracted from total membranes using 1% (w/v) DDM, and subsequently purified using 0.1% DDS and 0.05% Cymal6. Before protein elution, an additional wash step was implemented with 10 CV of wash buffer B (Table 4.2.3.1) to remove residual DDM.

Buffer	DDM	DDM & DDS	DDM & Cymal6
Extraction buffer	20 mM Tris-HCl (pH 7.5) 300 mM NaCl 10% (v/v) glycerol 0.05% (w/v) DDM	20 mM Tris-HCl (pH 7.5) 300 mM NaCl 10% (v/v) glycerol 0.05% (w/v) DDM	20 mM Tris-HCl (pH 7.5) 300 mM NaCl 10% (v/v) glycerol 0.05% (w/v) DDM
Wash buffer A	20 mM Tris-HCl (pH 7.5) 300 mM NaCl 10% (v/v) glycerol 0.05% (w/v) DDM 50 mM imidazole	20 mM Tris-HCl (pH 7.5) 300 mM NaCl 10% (v/v) glycerol 0.05% (w/v) DDM 50 mM imidazole	20 mM Tris-HCl (pH 7.5) 300 mM NaCl 10% (v/v) glycerol 0.05% (w/v) DDM 50 mM imidazole
Wash buffer B		20 mM Na ⁺ phosphate (pH 6.4) 10% (v/v) glycerol 0.1% DDS 50 mM imidazole	20 mM Na ⁺ phosphate (pH 6.4) 10% (v/v) glycerol 0.05% Cymal6 50 mM imidazole
Elution buffer	20 mM Tris-HCl (pH 7.5) 300 mM NaCl 10% (v/v) glycerol 0.05% (w/v) DDM 250 mM imidazole	20 mM Na ⁺ phosphate (pH 6.4) 10% (v/v) glycerol 0.1% DDS 250 mM imidazole	20 mM Na ⁺ phosphate (pH 6.4) 10% (v/v) glycerol 0.05% Cymal6 250 mM imidazole
Dialysis buffer	20 mM Tris-HCl (pH 7.5) 150 mM NaCl 0.05% (w/v) DDM	20 mM Na ⁺ phosphate (pH 6.4) 10% (v/v) glycerol 0.1% DDS	10 mM Na ⁺ HEPES (pH 7) 0.05% Cymal6

Table 4.2.3.1. Buffers for AcrB purification. AcrB protein was purified using DDM as the detergent for extraction, and either DDM, DDS or Cymal6 for purification.

4.2.4. AcrB crystallisation

4.2.4.1. Crystallisation

Crystallisation trials were carried out using the vapour diffusion sitting drop method on 96-well CrystalQuick low profile plates (Grenier). Purified AcrB at a concentration of 28 mg/mL (250 μ M) in DDM (0.05% w/v), DDS (0.1% w/v) or Cymal6 (0.05% w/v) was screened against the MemGold and MemGold2 (Molecular Dimensions) screening plates. The reservoir volume was 45 μ L and the sitting drop consisting of 100 nL of AcrB and 100 nL of reservoir solution was dispensed by a Mosquito robot. Crystal plates were incubated at 20 °C and 4 °C. After overnight incubation, AcrB crystals appeared in multiple conditions with different sizes and morphologies. Once hit crystals were identified, crystallisation conditions were optimised on 96-well CrystalQuick low profile plates. Different drop sizes (200 nL, 250 nL and 300 nL), pH ($\pm$ 0.5) as well as protein (15 mg/mL, 30 mg/mL and 60 mg/mL), salt ($\pm$ 50 mM with 25 mM interval) and PEG concentrations ($\pm$ 4% (w/v) with 2% (w/v) interval) were screened against the original crystallisation conditions. Crystallisation optimisation was also performed on hanging drop 24-well XRL plates by mixing protein and reservoir solution at a ratio of 1:1 (1 μ L + 1 μ L). The quality of crystals was tested by *in-situ* or pin-mode X-ray experiments on I03, I04, I04-1 and I24 beamlines at the Diamond Light Source.

4.2.4.2. Crystal soaking

AcrB crystals from DDM, DDS or Cymal6 were used for compound soaking experiments. Crystals were prepared in 24-well XRL plates as described in section 4.2.4.1. For each test compound, 0.02 μ L of the 100 mM stock solution was added to the crystallisation drops. Crystals were harvested after 10 minutes, 30 minutes, 1 hour, 5 hours and overnight incubation.

4.2.4.3. Co-crystallisation

The test compound (100 mM) was mixed with purified AcrB protein at a molar ratio of 1:100. The mixture was incubated overnight at room temperature prior to crystallisation trials against the MemGold and MemGold2 screening plates using Mosquito. Once hit crystals were identified, crystallisation conditions were optimised on 24-well XRL plates and 96-well CrystalQuick low profile plates as described in 4.2.4.1.

4.2.4.3. XChem fragment screening

AcrB-DDS crystals were optimised and used for XChem fragment screening. The detailed crystal optimisation process is described in section 4.2.6.

4.2.5. AcrB data collection and structure determination

AcrB crystals were cryoprotected by transferring them to crystallisation buffer supplemented with 30% glycerol (v/v) and appropriate detergent and flash-cooled in liquid nitrogen for data collection. X-ray data were collected on the I03, I04-1 and I24 beamlines at the Diamond Light Source.

Data collection and structure determination from an AcrB-DDS crystal will be described here as an example. X-ray data were collected at a wavelength of 0.9763 Å on the I03 beamline at the Diamond Light Source. Data were indexed using Xia2 (Winter, 2010). The AcrB-DDS crystal belonged to the space group $P12_11$ and diffracted X-rays to 2.57 Å resolution. Data collection statistics are summarised in Table 4.2.5.1. Each asymmetric unit contains two molecules of the AcrB trimer with a solvent content of 89.02%. Initial phases were obtained by molecular replacement with Phaser (McCoy *et al.*, 2007) using PDB 2J8S as a search model (Sennhauser *et al.*, 2007). Iterative cycles of model building and refinement were completed using Refmac5 (Murshudov *et al.*, 2011) and COOT (Emsley & Cowtan, 2004). The final AcrB structure was refined at a resolution of 2.57 Å with a R-factor of 24.2% and R-free of 28.58%.

4.2.6. XChem fragment screening

The XChem fragment screening pipeline entails six steps: (1) crystal preparation, (2) crystal ranking, (3) acoustic crystal soaking, (4) crystal harvesting, (5) data collection and (6) data analysis. A graphic illustration of this process is shown in Figure 4.2.6.1.

Crystal preparation

Crystallisation of AcrB-DDS protein was optimised for the SwissCi/MRC 3-drop plate (Molecular Dimensions), the supported format for the XChem fragment screening platform. Each crystal drop consisted of 100 nL of the AcrB protein and 100 nL of the reservoir solution.

Crystal ranking

~500 optimised AcrB crystals were imaged using the Rock Imager 2 (Formulatrix) and the resulting images were imported to the program TeXRank (Ng *et al.*, 2014). Prior to compound deposition, it is necessary to identify the drops where crystals have grown and the exact location of crystals in each drop. The program TeXRank automatically ranks different crystal drops and queues these drops by the order of likelihood of presence of crystals. The desired dispensing coordinates were manually selected and exported for Echo (Labcyte).

Acoustic crystal soaking

Echo, an acoustic liquid handling system, was used to transfer compound solution or cryoprotectant to the crystallisation drops (<100 μm accuracy). The crystal plate was inverted over a compound plate or cryoprotectant plate. The Echo system uses ultrasonic pulses to add series of droplets (2.5 nL) to the selected dispensing coordinates until the required volume is

reached. Crystal plates with dispensed compound and cryoprotectant were resealed and incubated at 20 °C prior to crystal harvesting.

For XChem fragment screening, fragments from the Diamond-SGC-iNext (DSi)-Poised Library (Table 4.2.1.1) (Cox *et al.*, 2016) were tested against the AcrB crystals. Different final concentrations of DMSO (5%, 10%, 20% and 30%) and different incubation times (one hour, two hours and three hours) were tested to establish crystal tolerance. One-hour incubation with 5% final concentration of DMSO was determined to be the best condition for crystal soaking, which made up the final compound concentration of 23.8 mM. Glycerol, ethylene glycol, sucrose, PEG 400, nicotinamide with ethylene glycol and sodium p-toluenesulfonate with ethylene glycol at different concentrations (5%, 10%, 20% and 30%) were tested as cryoprotectants with different incubation times (one hour, two hours and three hours). The determined cryoprotectant contained the reservoir solution supplemented with 30% glycerol and 0.1% DDS. The best incubation time for cryoprotectant was found to be one hour. Once the DMSO tolerance and cryoprotectant were established, a checkerboard sensitivity test was performed to determine the best combinations of concentrations and incubation times of DMSO and cryoprotectant. The final crystal soaking protocol was established to firstly add fragment (5%) to crystal drops and incubate for 10 minutes, then add cryoprotectant (30% glycerol) and incubate for one hour at 20 °C. This was applied for subsequent experiments.

Crystal harvesting

Crystal harvesting was done using the Crystal Shifter Device (Oxford Lab Technologies) to quickly translate desired wells to the harvesting location, typically below the microscope. The crystal shifter platform consists of an x-y microscope stage and software that automatically tracks sample handling. The software is operated using a large touch screen where different

conditions (*e.g.* mounted or failed; clear, crystalline or precipitated; harvesting time; crystal ID; destinations) can be recorded. ~500 crystals were harvested using the shifter platform. Crystals were mounted using the Dual-Thickness MicroLoops (MiTeGen) and flash-cooled in liquid nitrogen.

Data collection

Data were mostly collected at a wavelength of 0.916 Å on the I04-1 beamline at the Diamond Light Source. For each dataset, in total 1500 images were collected with 0.12° oscillation. AcrB-DDS crystals belonged to the space group $P12_11$ and diffracted consistently to ~2.7 Å.

Data analysis

Data were integrated automatically by the XChemExplorer, a SQLite-based database, and analysed using PanDDA (Pan-Dataset Density Analysis), a density modification algorithm which facilitates identification of even weak binders (Krojer *et al.*, 2017; Pearce *et al.*, 2017).

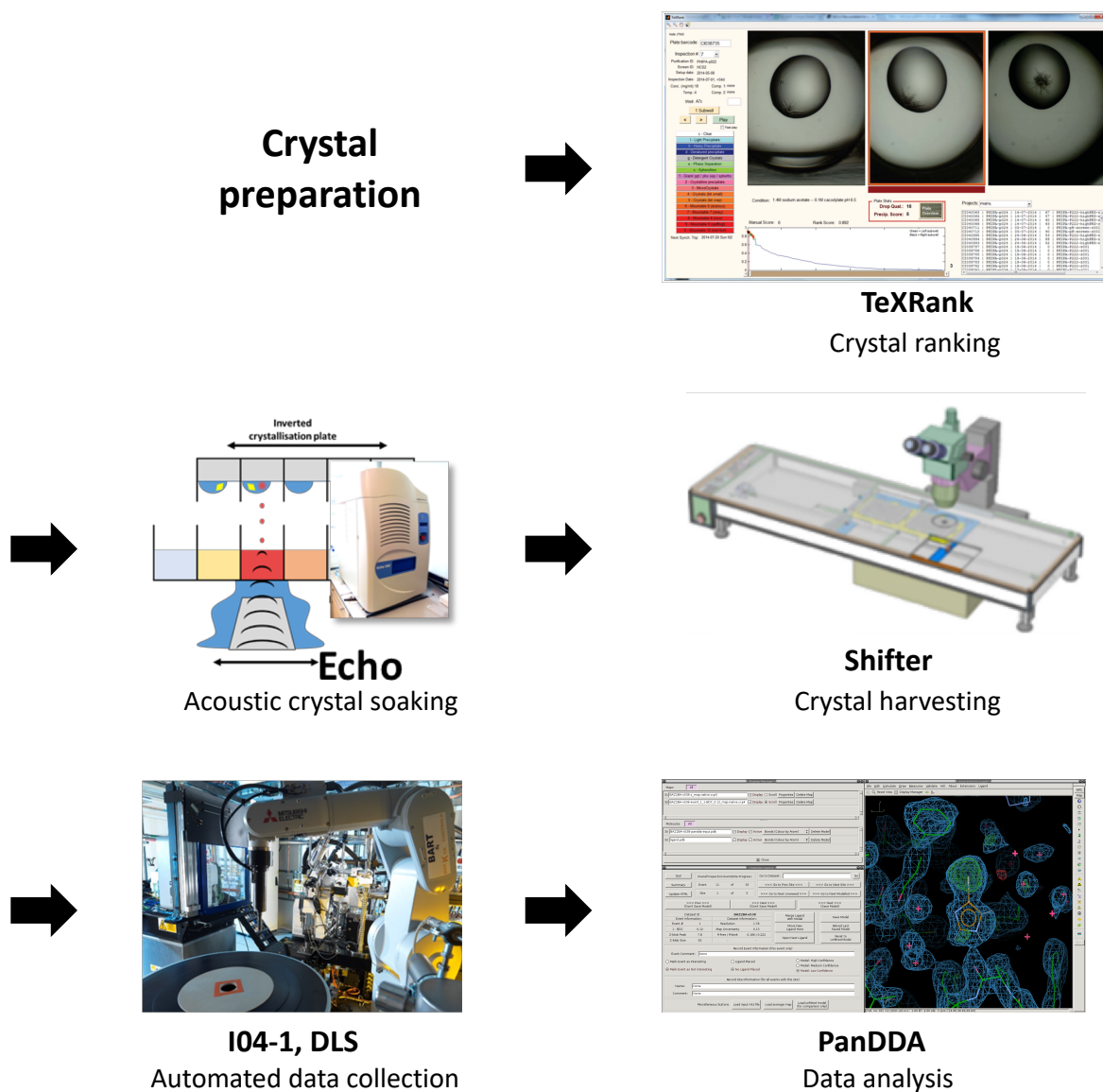


Figure 4.2.6.1. Graphic illustration of the XChem fragment screening pipeline. The XChem pipeline contains six steps: crystal preparation, crystal ranking, crystal soaking, crystal harvesting, automated data collection on the I04-1 beamline at the Diamond Light Source, and data analysis. Adapted from (<https://www.mitegen.com/lab-spotlight-XChem>).

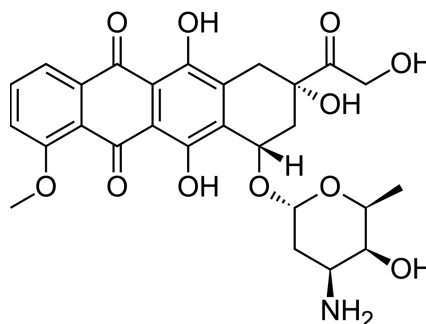
4.3. Results

4.3.1. Doxorubicin efflux assay

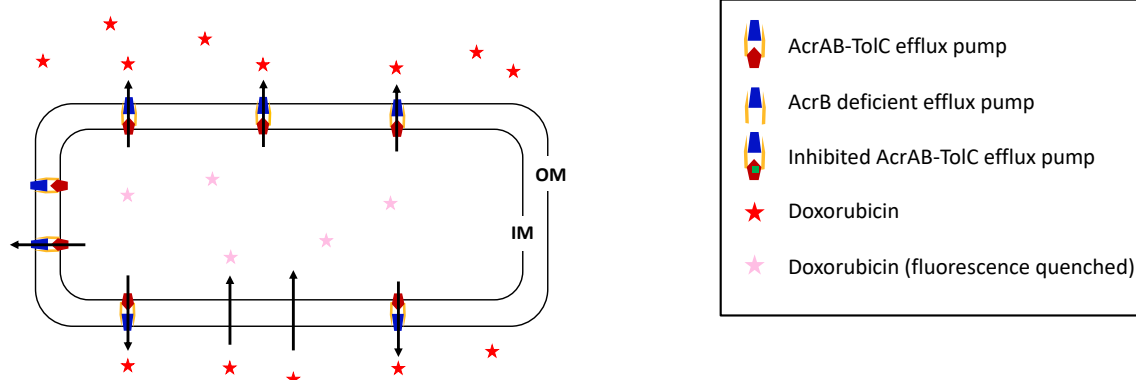
Doxorubicin, an anthracycline antibiotic, was initially discovered in *Streptomyces peucetius* (Figure 4.3.1.1A). It is widely used as a cytostatic drug in cancer chemotherapy, such as breast cancer, bladder cancer, lymphoma, Kaposi's sarcoma and bronchial carcinoma (Tacar *et al.*, 2012). Doxorubicin acts by binding to DNA-associated enzymes (such as topoisomerase enzymes I and II) and intercalating with DNA base pairs, therefore inducing a range of cytotoxic effects and antiproliferation, leading to DNA damage. Doxorubicin is also known to be fluorescent, a property often used in doxorubicin concentration characterisation.

Doxorubicin has been shown to be a substrate of the AcrAB-TolC efflux pump, with its fluorescence quenched when accumulated inside the cells. Together with determination of the first inhibitor-bound AcrB structure, Nakashima *et al.* (2013) have established a fluorescence-based doxorubicin efflux assay to enable rapid identification of possible AcrB inhibitors (Figure 4.3.1.1). AcrB⁺ cells extrude doxorubicin, resulting in an almost constant fluorescence emission in the extracellular medium due to the almost balanced influx and efflux of doxorubicin (Figure 4.3.1.1B). AcrB⁻ cells are efflux-deficient. Therefore, doxorubicin accumulates inside AcrB⁻ cells and gets quenched in a time-dependent manner (Figure 4.3.1.1C). In the presence of an effective EPI, doxorubicin efflux gets inhibited and fluorescence quenching is observed (Figure 4.3.1.1D). The level of efflux pump inhibition can therefore be quantified by measuring the quenching of doxorubicin in the presence of an inhibitor.

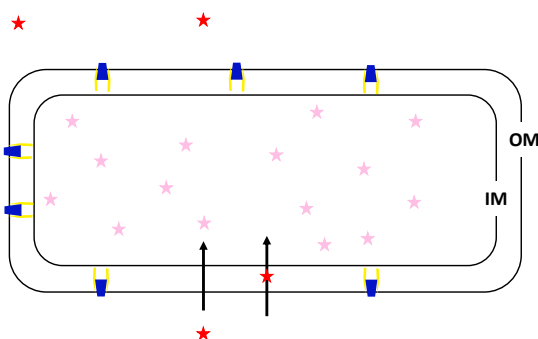
A. Structure of doxorubicin



B. Wild-type cells



C. AcrB deficient cells



D. Wild-type cells with an inhibitor

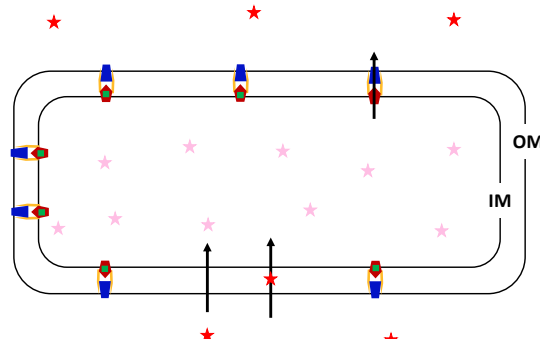


Figure 4.3.1.1. Doxorubicin efflux and accumulation. (A) Structure of doxorubicin. (B) AcrB⁺ cells extrude doxorubicin, resulting in an almost stable fluorescence emission in the extracellular medium. (C) AcrB⁻ cells accumulate doxorubicin, causing a significant fluorescence quenching. (D) An effective EPI inhibits doxorubicin efflux and fluorescence quenching is observed. AcrA, AcrB, TolC and an effective EPI are coloured in yellow, red, blue and green, respectively. Doxorubicin and fluorescence quenched doxorubicin molecules are shown as red and pink stars, respectively.

In this work, the fluorescence-based doxorubicin efflux assay was optimised and performed to screen the NCI diversity set V library (Table 4.2.1.1) against *E. coli* overexpressing the AcrAB-TolC efflux pump. The wild-type *E. coli* K-12 BW25113 was used as the test strain.

Control experiments were performed to establish the reliability and consistency of the doxorubicin efflux assay. The positive control consists of AcrB- *E. coli* cells, and the negative control is the AcrB+ *E. coli* cells. Since the test compounds were dissolved in DMSO, AcrB+ and AcrB- cells were tested in the presence of 1% DMSO. The fluorescence emission in AcrB+ *E. coli* cells remained almost stable (Figure 4.3.1.2, blue) while in the AcrB- cells, a significant fluorescence quenching was observed in a time-dependent manner (Figure 4.3.1.2, orange). This suggests that AcrB+ *E. coli* cells can export doxorubicin efficiently, while AcrB- *E. coli* cells do not have a functional pump for doxorubicin efflux. Since AcrB-deficient cells are incapable of doxorubicin efflux, the AcrAB-TolC system is the only efflux pump involved in doxorubicin export in the test *E. coli* K-12 BW25113 strain.

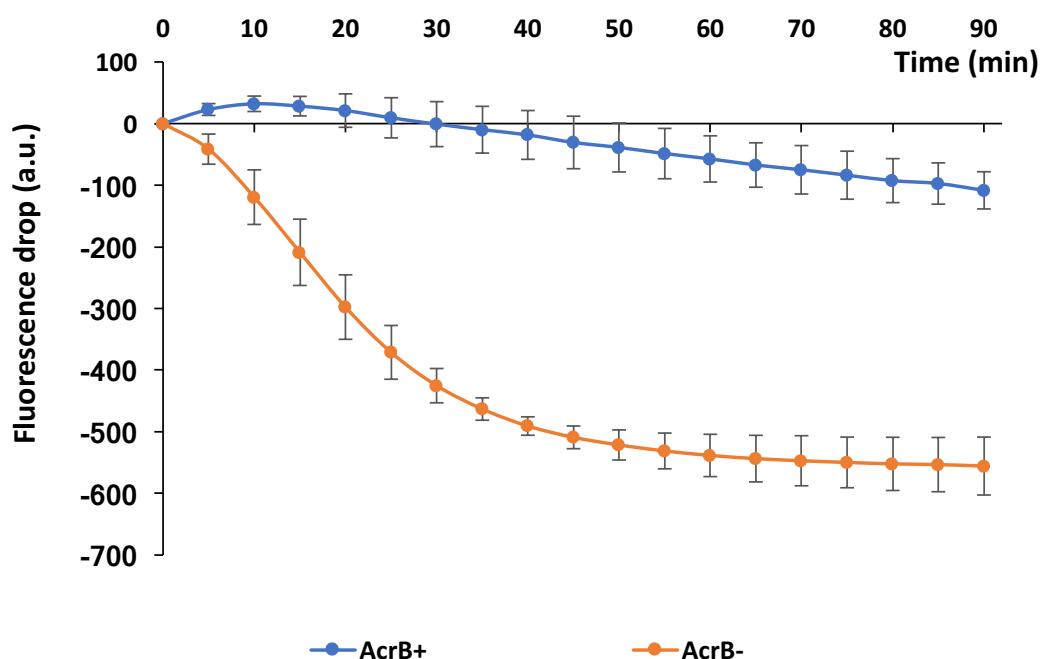


Figure 4.3.1.2. The effect of AcrB deletion on doxorubicin accumulation. Fluorescence changes of AcrB- and AcrB+ *E. coli* cells in the presence of 1% DMSO are shown in orange and blue, respectively. A significant fluorescence quenching was observed in the positive control (*i.e.* AcrB- cells), while the fluorescence emission quenched slowly and gradually in the negative control (*i.e.* AcrB+ cells). The data correspond to mean values of three independent experiments from different *E. coli* colonies (n=3). Error bars indicate the $\pm$ standard deviation of recorded values. a.u., arbitrary units.

4.3.2. Screening of the NCI diversity set V library

The NCI diversity set V library contains 1597 compounds with the MWs ranging between 100-770 Da (Table 4.2.1.1). Initial screening was performed on 644 compounds with MWs between 253-770 Da, which were thought to fit the binding pockets of AcrB. These compounds displayed different chemical structures. Initial screening resulted in 12 hits, which demonstrated an enhancement of fluorescence quenching. Following the preliminary screening, three independent transport experiments were performed with the identified compounds and six compounds were further selected (Table 4.3.2.1 and Figure 4.3.2.1). These compounds have relatively low MWs ranging between 313-408 Da. Doxorubicin is a LMMD (MW: 544 Da) that binds to the distal binding pocket, bypassing the proximal binding pocket. Therefore, it is not surprising that smaller compounds, which potentially also bind to the distal pocket, are more potent inhibitors for doxorubicin efflux than larger compounds. It is notable that the identified inhibitors are at least 136 Da lower in MW than doxorubicin, while compounds whose MWs are closer to doxorubicin are not identified as doxorubicin efflux inhibitors. This is most likely due to the flexibility of smaller compounds, so they fit more easily into the voluminous multisite binding pockets and inhibit substrate efflux.

Among the six compounds, Compound 1 showed the most significant fluorescence quenching (435 units) in the doxorubicin efflux assay, with compound 2 being the second (418 units) highest (Figure 4.3.2.2, A and B). At 100 μ M final concentration, these two compounds demonstrated an increased doxorubicin accumulation to a level that was ~four-times higher than that of the negative control (*i.e.* AcrB⁺ cells) and ~75% of the positive control (*i.e.* AcrB⁻ cells) (Figure 4.3.2.3). Structural comparison shows that both compounds contain a nitrophenyl-urea group (Figure 4.3.2.1), which might be critical for the similar level of high inhibitory effect on doxorubicin transport. These two compounds have relatively low MWs

(313 and 349 Da, respectively). Inhibitory effect of compounds 3, 4, 5 and 6 are not as severe as compounds 1 and 2, but still almost 3-fold higher fluorescence quenching has been observed compared to the negative control (Figure 4.3.2.2, C, D, E and F; Figure 4.3.2.3). Compounds 4 and 5 have similar scaffold with different substitution of chloride atoms, which might be the reason for the similar level of inhibition (Figure 4.3.2.1; Figure 4.3.2.2, D and E).

The fluorescence emission of AcrB⁺ cells increased by ~30 units in the first 10 minutes and then decreased gradually (Figure 4.3.2.2). A similar trend is also seen in AcrB⁺ cells in the presence of compound 4/5, which increased by up to 32 units in the first 10 minutes. The initial fluorescence increase is likely due to incomplete mixing before the first measurement. However, in AcrB⁻ cells, an immediate decrease of fluorescence emission was observed since the assay started, suggesting that doxorubicin gets accumulated in AcrB⁻ cells in an acute manner. In AcrB⁺ cells with compound 3, the fluorescence quenching also started from the beginning of the assay. In the presence of compound 1/2/6, the fluorescence intensity increased to up to 8 units in the first 5 minutes, followed by an obvious fluorescence quenching. The relatively fast fluorescence quenching indicates that compounds 1, 2, 3 and 6 are probably more potent inhibitors of doxorubicin efflux than compounds 4 and 5. This trend was observed in the triplicate measurements.

Since these hits were identified by a whole-cell transport assay, the mechanism of inhibition remains unknown. The identified inhibitors may prevent the functional rotation mechanism of AcrB, in which case they are promising EPIs, or competing with doxorubicin for its substrate-binding pockets of AcrB, in which case they are competitive substrates. It is also possible that these inhibitors bind to AcrA or TolC, either inhibiting their function or disrupting the formation of the AcrAB-TolC complex. These compounds could even interact with the extracellular doxorubicin molecules and alter their fluorescence properties, therefore leading

to fluorescence quenching. Further investigation is needed to understand the mechanism that underpins the inhibitory effect. Although the level of inhibition varies greatly between different compounds, since we have yet to elucidate the mechanism behind the inhibitory effect, it is still worthwhile to further investigate all six identified compounds rather than just focusing on the best two hits.

Compound	Chemical name	Formula	MW (Da)
1	1-(5-isopropyl-2-methyl-phenyl)-3-(4-nitrophenyl)urea	C ₁₇ H ₁₉ N ₃ O ₃	313
2	1-(1,3-benzothiazol-2-yl)-3-(4-chloro-3-nitrophenyl)urea	C ₁₄ H ₉ ClN ₄ O ₃ S	349
3	6-(2-naphthalenylthio)-2,4-quinazolinediamine, ethanol (2:1)	C ₁₈ H ₁₄ N ₄ S.C ₂ H ₆ O (2:1)	341
4	1-(4-chlorophenoxy)-3-[4-(3-chlorophenyl)piperazin-1-yl]propan-2-ol	C ₁₉ H ₂₂ Cl ₂ N ₂ O ₂	381
5	1-(2-chlorophenoxy)-3-[4-(3-methylphenyl)piperazin-1-yl]propan-2-ol	C ₂₀ H ₂₅ ClN ₂ O ₂	361
6	7-isopropyl-1,4a-dimethyl-2,3,4,9,10,10a-hexahydrophenanthren-1-yl)methanamine, benzoic acid (1:1)	C ₂₀ H ₃₁ N.C ₇ H ₆ O ₂ (1:1)	408

Table 4.3.2.1. Chemical names, formulas and MWs of the identified compounds.

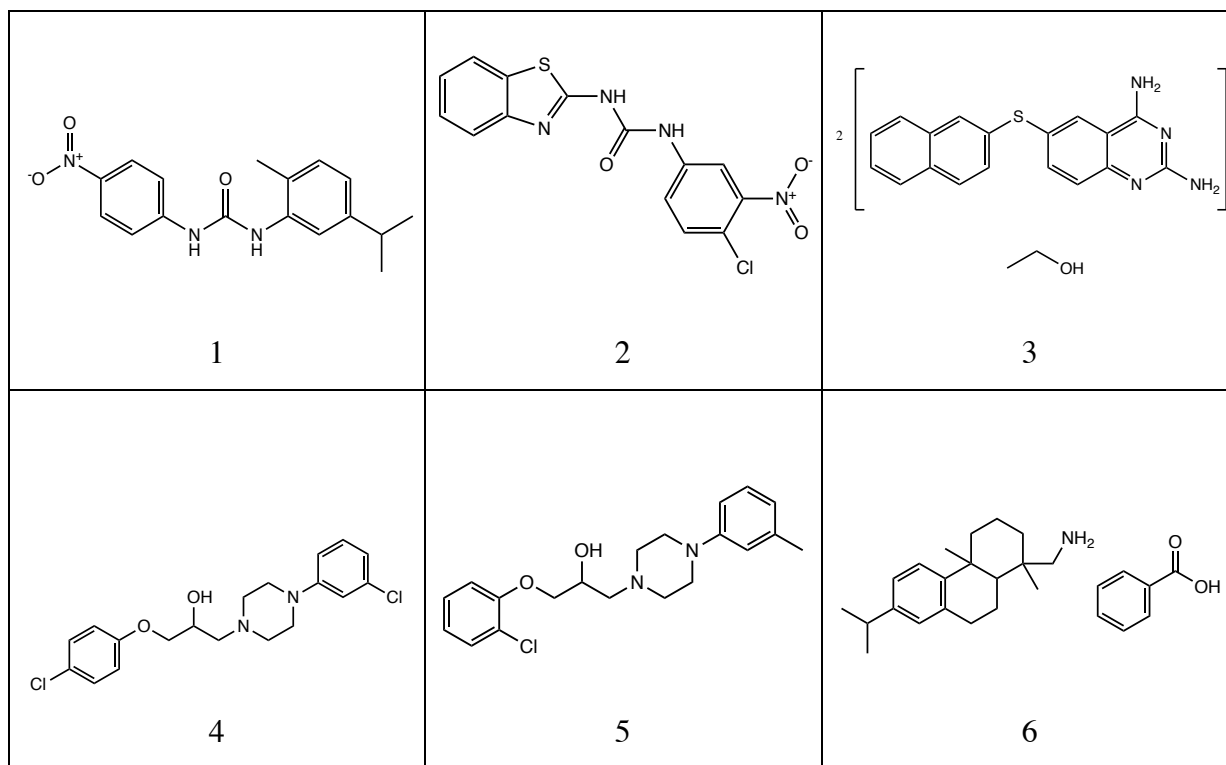


Figure 4.3.2.1. Structures of the identified inhibitors.

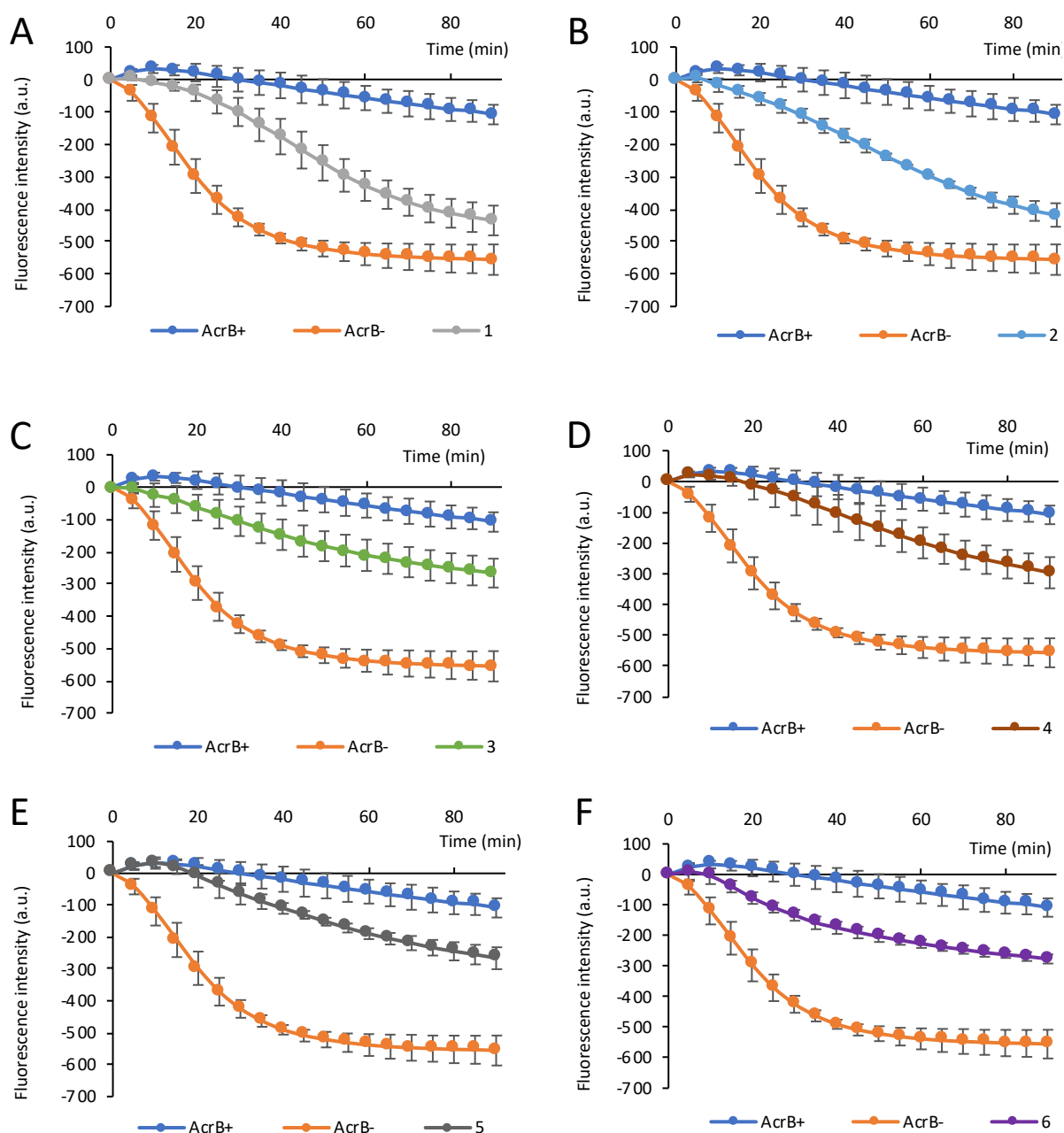


Figure 4.3.2.2. Compounds 1-6 inhibit doxorubicin efflux. Fluorescence changes in AcrB- and AcrB+ *E. coli* cells are shown in orange and blue, respectively. All six compounds show an enhancement on fluorescence quenching, with compounds 1 and 2 showing the highest inhibitory effect on doxorubicin efflux. The final concentration of the test compounds was 100 μ M. The data correspond to the mean values of three independent experiments from different *E. coli* colonies (n=3). Error bars indicate the standard deviation of values recorded. a.u.: arbitrary units.

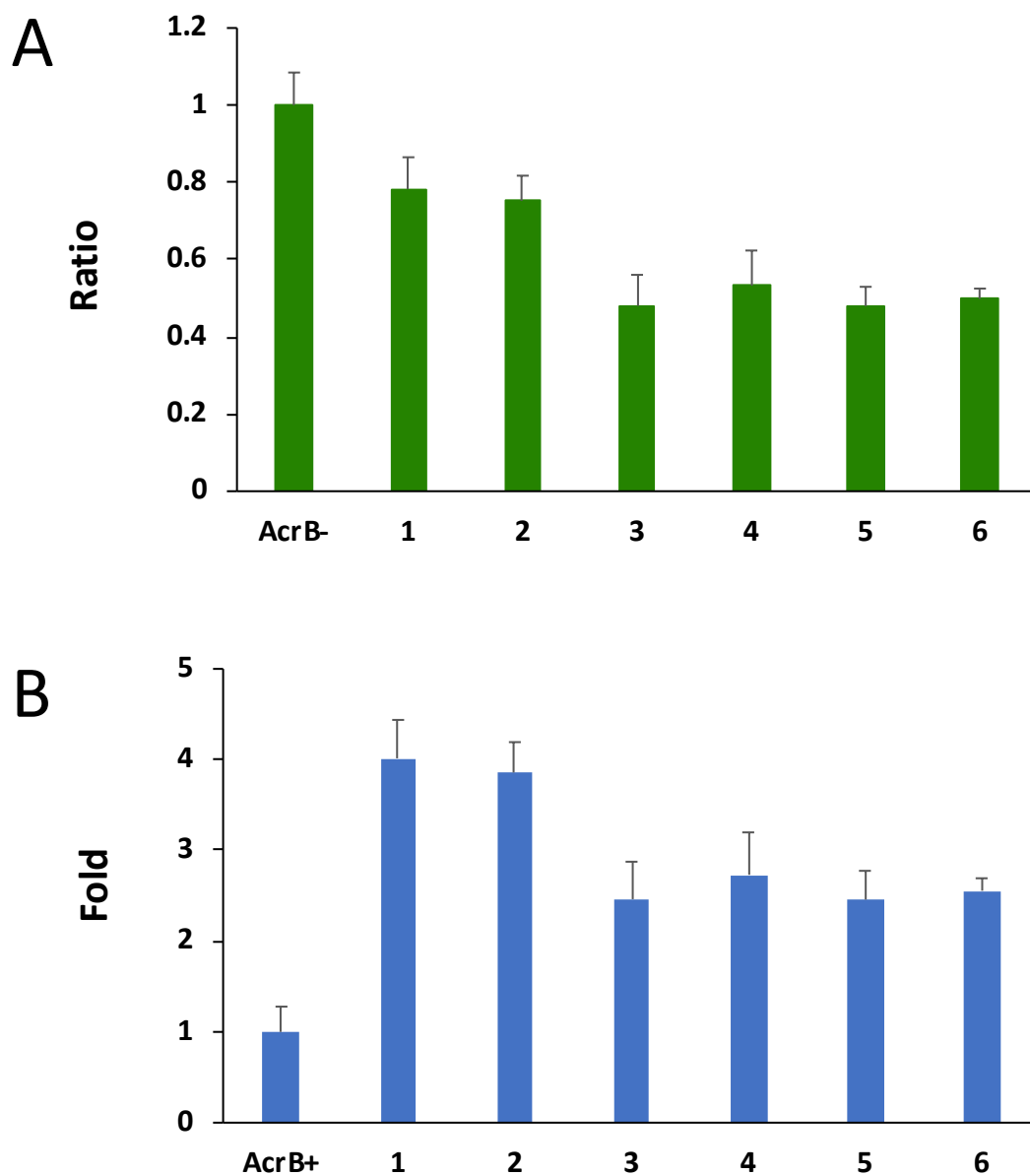


Figure 4.3.2.3. Fluorescence quenching compared with the controls. Positive (A) and negative (B) controls. Fluorescence quenching in the presence of a test compound was divided by the average fluorescence quenching in the positive (*i.e.* AcrB⁻ cells with DMSO) and negative (*i.e.* AcrB⁺ cells with DMSO) controls. The data correspond to mean values of three independent experiments from different *E. coli* colonies (n=3). Error bars indicate the $\pm$ standard deviation of recorded value.

4.3.3. Potentiation of antibiotics

The MIC is the lowest concentration of an antimicrobial agent that prevents visible growth of a microbe (Waitz, 1990). It is a standard approach in clinical studies to test antimicrobial susceptibility. In this study, the MIC test was performed against the wild-type *E. coli* K-12 BW25113 strain to validate the six identified inhibitors. MICs of the clinical antibiotics, intrinsic MICs of the identified inhibitors and MICs of antibiotics in the presence of an inhibitor were tested (Table 4.3.3.1). MICs of antibiotics in the presence of PA β N, a known EPI, was tested as a positive control.

AcrAB-TolC confers resistance to cells against rifampicin, novobiocin (613 Da), chloramphenicol (323 Da), tetracycline (444 Da) and erythromycin, while carbenicillin (378 Da) is not a substrate of this system (Elkins & Nikaido, 2002; Kobayashi *et al.*, 2014). An antibacterial test was performed on these six clinically used antibiotics, which covered AcrB substrates of different MWs and a non-substrate. The MICs of the antibiotics were comparable to their published MIC values (Table 4.3.3.1) (Opperman *et al.*, 2014). Small differences in the MIC values (2-4-fold) were observed for rifampicin, novobiocin, tetracycline and erythromycin, which might be due to the different *E. coli* strains used in the assay. Compounds 1, 2, 4 and 5 alone did not exhibit an antibacterial activity at 100 μ M, suggesting that their intrinsic MIC values > 100 μ M (Table 4.3.3.1). At a final concentration as low as 6.25 μ M, compound 3 demonstrated an inhibitory effect on cell growth (Table 4.3.3.1). This means that compound 3 itself is an antibacterial agent, which might be the reason for the fluorescence quenching observed in the doxorubicin efflux assay (Figure 4.3.2.2C). However, the mechanism that underpins the antibacterial effect of compound 3 remains unknown, therefore limiting its further clinical development. The MIC value of compound 6 was shown to be 100

μM. In the following assay, potentiation effect of compound 6 was tested at a final concentration of 25 μM, which is in accordance with a previous study (Opperman *et al.*, 2014).

MICs of clinical antibiotics in the presence of the identified inhibitors were subsequently tested. Final concentrations were 100 μM for compounds 1, 2, 4, 5, while 25 μM for compound 6 and PAβN. Compound 3 was not tested due to the high intrinsic antibacterial activity. MICs of chloramphenicol and tetracycline in the presence of compound 6 were not tested due to time limitations. PAβN potentiated the bactericidal activity of several known AcrAB-TolC substrates, including rifampicin, novobiocin, chloramphenicol ($\geq$ 32-fold) and erythromycin (4-fold), but not carbenicillin, which is not a substrate (Table 4.3.3.1) (Kobayashi *et al.*, 2014). Similarly, none of the six test inhibitors potentiated the antibacterial activity of carbenicillin. PAβN did not increase susceptibility to tetracycline, which is consistent with a previous study (Opperman *et al.*, 2014). Compounds 1, 2, 4 and 5 at 100 μM did not show an obvious enhancement of bacterial susceptibility to the tested antibiotics, except for a 2-fold decrease in the MIC of either rifampicin, chloramphenicol or tetracycline. This indicates that these compounds might be weak competing substrates specific to doxorubicin, or they induce doxorubicin quenching through other mechanisms such as altering the doxorubicin fluorescent properties. Compound 6 at 25 μM increased susceptibility to rifampicin by 4-fold, but not to novobiocin or erythromycin. Compound 6 is presumably a competitive inhibitor of the AcrAB-TolC system. Rifampicin and erythromycin are both HMMDs that bind to different binding sites in the proximal binding site of AcrB (Figure 4.0.4.4, A and B) (Nakashima *et al.*, 2011). The differentiated potentiation effect of compound 6 between rifampicin and erythromycin suggests that this compound probably inserts into the specific rifampicin-binding site, which is located at the top of the proximal binding pocket of AcrB and is distinct to the erythromycin-binding site (Figure 4.0.4.4, A and B). Novobiocin is predicted to bind to the distal binding

pocket, which might explain why compound 6 shows no effect on novobiocin (Takatsuka *et al.*, 2010). Antibacterial assays in clinically-relevant bacteria with more antibiotics are being undertaken by our medicinal microbiologist collaborator, Prof. Laura Piddock (University of Birmingham).

MIC in the presence of an EPI (µg/mL)								
EPI	MIC (µM)	EPI test conc. (µM)	CARB	RIF	NOV	CAM	TET	ERY
C1	>100	100	8	8	512	8	0.5	32
C2	>100	100	4	8	512	4	0.5	32
C3	6.25	ND	ND	ND	ND	ND	ND	ND
C4	>100	100	4	4	512	8	1	32
C5	>100	100	4	4	512	4	1	32
C6	100	25	4	2	512	ND	ND	64
PAβN	ND	25	8	<0.125	<1	0.25	1	8
No EPI		0	4	8	512	8	1	32

Table 4.3.3.1. Potentiation effect on different antibiotics. Compounds 1, 2, 4 and 5 at a concentration of 100 µM did not show an inhibitory effect on cell growth (intrinsic MIC >100 µM). The intrinsic MIC values of compound 3 and compound 6 were established to be 6.25 µM and 100 µM, respectively. The table shows the MICs of diverse antibiotics against *E. coli* K-12 BW25113 cells in the presence/absence of identified inhibitors. PAβN is the EPI known to inhibit the efflux of various antibiotics. PAβN potentiated the bactericidal activity of several known AcrAB-TolC substrates, including rifampicin, novobiocin, chloramphenicol (≥ 32-fold) and erythromycin (4-fold), but not carbenicillin, which is not a substrate. Compound 6 at 25 µM increased susceptibility to rifampicin by 4-fold, but not to novobiocin or erythromycin. MIC shifts of 2-fold, 4-fold and ≥ 8-fold are coloured blue, green and red, respectively. CARB: carbenicillin; RIF: rifampicin; NOV: novobiocin; CAM: chloramphenicol; TET: tetracycline; ERY: erythromycin. ND: not determined.

4.3.4. AcrB purification

Successful crystallisation of a membrane protein requires highly pure protein (Carpenter *et al.*, 2008). The use of different detergents could lead to different sized micelles and different crystal packings, thus potentially changing the quality of crystals. In order to produce optimal AcrB crystals, detergents DDM, DDS and Cymal6 were used for AcrB purification. For DDM purification, AcrB was extracted and purified using DDM throughout the purification process. For DDS and Cymal6, DDM was used for protein extraction due to the high solubilisation efficiency and stability. DDS and Cymal6 was implemented during the IMAC purification step. An additional washing step was performed in order to remove the residual DDM. Purification without this step lead to crystals with low quality and consistency.

The single band from SDS-PAGE indicates that AcrB protein has been purified to a high level, using DDM as a detergent for extraction, and either DDM (Figure 4.3.4.1A, lane 4), DDS (Figure 4.3.4.1B, lane 5) or Cymal6 (Figure 4.3.4.1C, lane 5) for purification. The protein band is slightly lower than theoretical MW of AcrB due to the masking effect of SDS on the hydrophobic region of the protein.

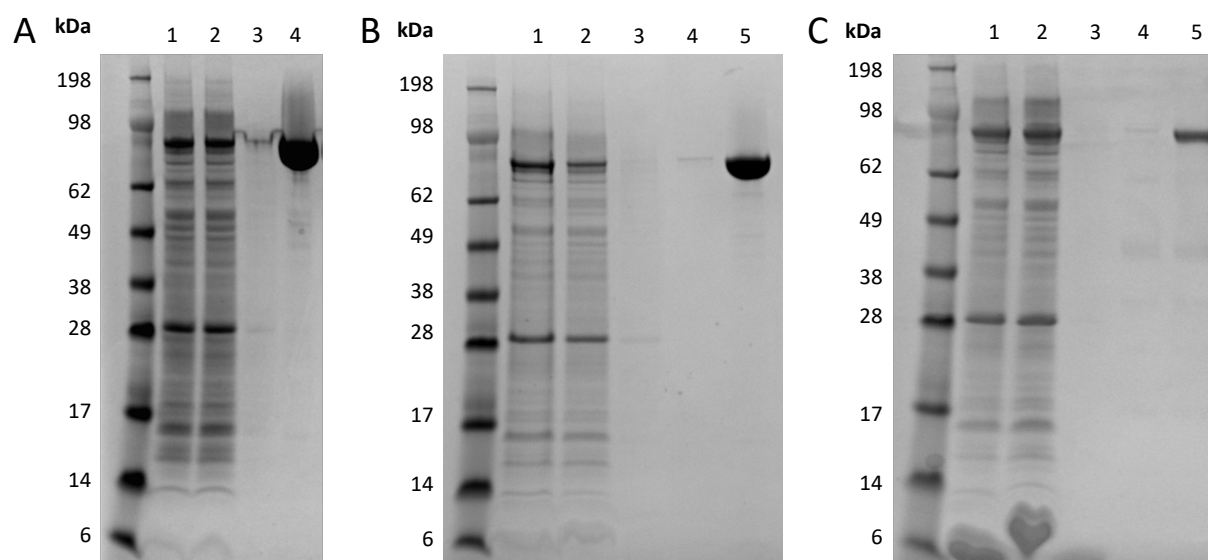


Figure 4.3.4.1. SDS-PAGE for AcrB purification. The SDS-PAGE shows a single band corresponding to the expected MW of AcrB. (A) AcrB purification with DDM. Lane 1: solubilised material, lane 2: flow through of IMAC, lane 3: wash (50 mM imidazole and DDM), lane 4: elution (250 mM imidazole and DDM). (B) AcrB purification with DDM and DDS. Lane 1: solubilised material, lane 2: flow through of IMAC, lane 3: wash A (50 mM imidazole and DDM), lane 4: wash B (50 mM imidazole and DDS), lane 5: elution (250 mM imidazole and DDS). (C) AcrB purification with DDM and Cymal6. Lane 1: solubilised material, lane 2: flow through of IMAC, lane 3: wash A (50 mM imidazole and DDM), lane 4: wash B (50 mM imidazole and Cymal6), lane 5: elution (250 mM imidazole and Cymal6).

4.3.5. AcrB crystallisation and structure

Production of reproducible high-quality crystals is a prerequisite for successful SBDD, which allows high confidence in ligand identification and detailed interpretation of protein-ligand interactions. In this project, extensive crystal optimisation has been attempted in order to produce high-quality AcrB crystals that favour ligand binding. Purified AcrB in DDM, DDS and Cymal6 was screened against the MemGold and MemGold2 (Molecular Dimensions) screening plates at 20 °C and 4 °C. Once hit crystals were identified, crystallisation optimisation was performed against the original crystallisation conditions. The crystal quality was tested by *in-situ* or pin-mode X-ray experiments on the I03, I04, I04-1 and I24 beamlines at the Diamond Light Source.

Properties of detergents used in crystallisation were found to have a major effect on quality, consistency and the space group of the crystals. Initially AcrB-DDM crystals diffracted X-rays to ~ 4 Å and belonged to the space group *R*32. In order to improve crystal quality, Cymal6 was used for protein purification and crystallisation (Pos & Diederichs, 2002), which resulted in an obvious improvement in crystal quality and consistency. AcrB-Cymal6 crystals in reservoir containing 0.05 M lithium sulfate, 0.1 M Tricine (pH7.4) and 7% PEG3000 from handing drop 24-well XRL plates (Figure 4.3.5.1A) diffracted X-rays to ~ 3 Å. However, in the presence of DDM or Cymal6, crystals were formed in a high-symmetry rhombohedral *R*32 space group, which is thought to be physiologically irrelevant. In the physiological condition, the AcrB trimer stays in an asymmetric state with each monomer following a different conformation, the access, binding and extrusion monomers (Murakami *et al.*, 2006; Seeger *et al.*, 2006).

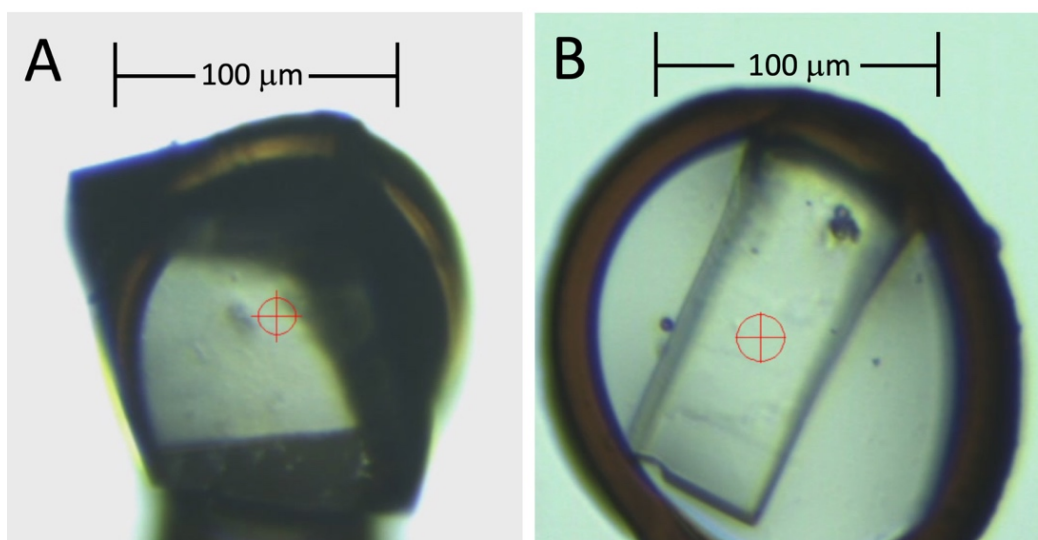


Figure 4.3.5.1. Optimised AcrB crystals. (A) AcrB-Cymal6 crystallised at 20 °C in 0.05 M lithium sulfate, 0.1 M Tricine (pH 7.4) and 7% PEG3000. (B) AcrB-DDS crystallised at 20 °C in 0.1 M ADA (pH 5.5) and 14% (w/v) PEG 6000.

In order to produce asymmetric physiologically-relevant AcrB structures, DDS was used in protein purification and crystallisation (Murakami *et al.*, 2006). At first, both DDM and DDS were used as detergents for purification and crystallisation. Asymmetric AcrB crystals were produced in the presence of DDM and DDS, but only 10% of them diffracted to an optimal resolution of 4-5 Å. This was most likely due to the non-specific interactions between the two types of detergent molecules; thus, the detergent micelles were formed in an unstable and inconsistent state. Finally, DDS was used as the sole detergent for purification and crystallisation. AcrB-DDS crystals were formed in the asymmetric monoclinic $P12_1$ or $C2$ space group. DDS as a crystallisation detergent has significantly improved crystal quality and consistency. Crystals were grown in 0.1% DDS, 0.1 M ADA (pH 5.5), and 14% (w/v) PEG 6000 belonged to the $P12_1$ space group, with 50% of the crystals diffracting X-rays to ~ 2.7 Å (Figure 4.3.5.1B, Figure 4.3.5.2 and Table 4.3.5.1). Three monomers of AcrB follow three different conformations, access, binding and extrusion, which have been shown to be

physiological states in drug export (Figure 4.3.5.3) (Murakami *et al.*, 2006; Seeger *et al.*, 2006). Compared to the previously obtained symmetric AcrB structures in the *R*32 space group, the asymmetric structure significantly increases the possibility of ligand binding due to the presence of both proximal and distal binding pockets. Close-up views of the proximal and distal binding pocket indicate that these two binding pockets are empty, which means that this structure is available for ligand binding (Figure 4.3.5.4). Detergent and lipid densities could be observed at numerous locations proximal to the transmembrane domain of AcrB (Figure 4.3.5.5).

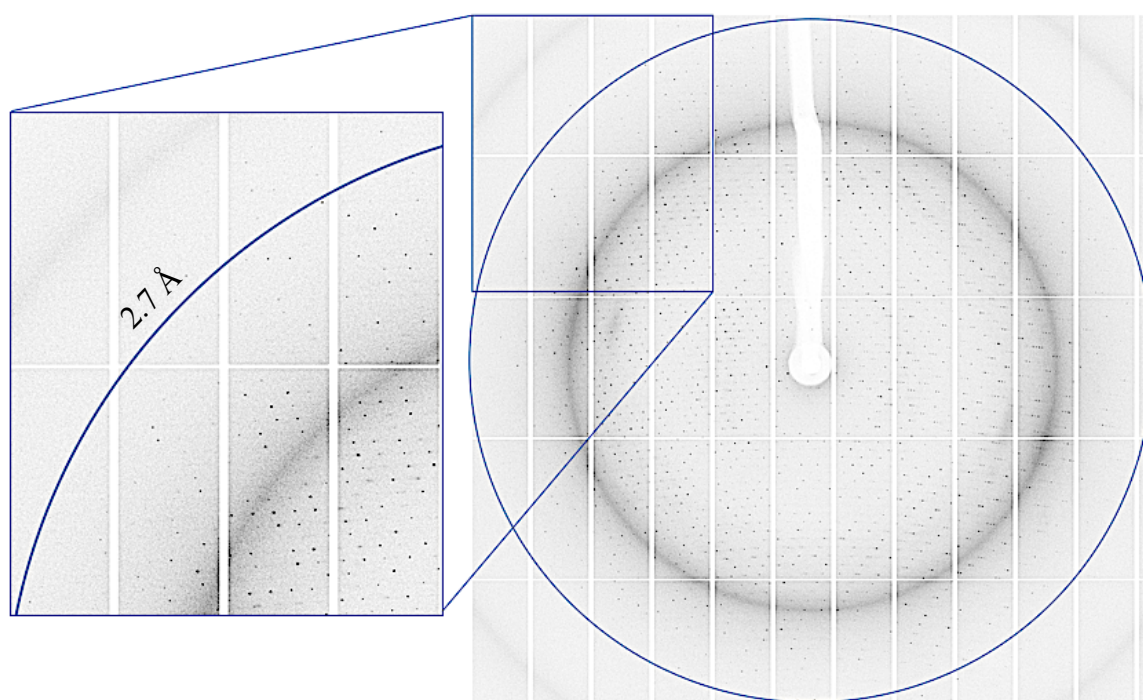


Figure 4.3.5.2. X-ray diffraction pattern for AcrB at 2.57 Å. The blue circle shows the diffraction limit at 2.7 Å and the inset indicates the diffraction spots at high resolution.

Data collection	
Beamline	DLS-I03
Wavelength (Å)	0.9763
Resolution (Å)	46.26 - 2.57 (2.61 - 2.57)
Space group	<i>P</i> 12 ₁ 1
Cell dimensions	
a,b,c (Å)	130.03, 132.07, 296.97
α, β, γ (°)	90, 94.26, 90
NO. of reflections	730760 (35762)
NO. of unique reflections	315656 (15743)
Completeness (%)	99.2 (99.1)
Redundancy	2.3 (2.3)
R _{meas} (%)	7.5 (1.034)
Mean (I/ σ (I))	10.1 (1.3)
Mn(I) half-set correlation CC (1/2) (%)	1 (0.5)
Refinement	
R _{work} / R _{free} (%)	24.20/28.58
Average B-factors (Å ²)	79.3
NO. of non-hydrogen atoms	47801
RMS deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.73
Ramachandran	
Favoured (%)	96.65
Allowed (%)	2.37
Outliers (%)	0.98

Table 4.3.5.1. Data collection and refinement statistics for AcrB crystals. The R-free was calculated by excluding 5% of data from refinement. Values in parentheses are for the high-resolution shells.

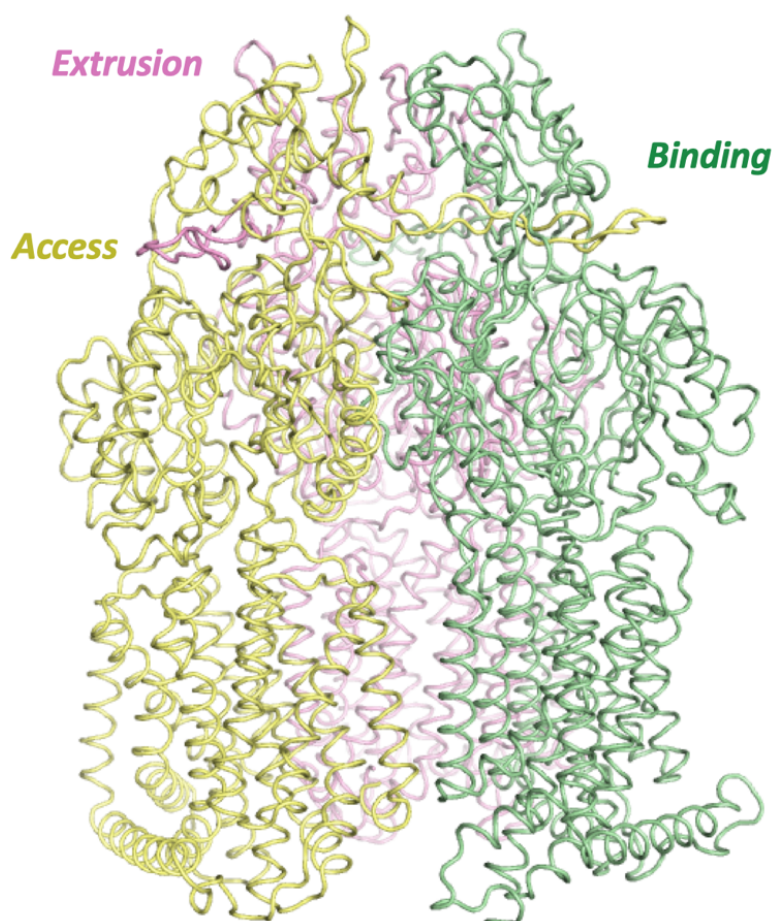


Figure 4.3.5.3. Structure of the apo asymmetric AcrB. In this structure, each monomer follows a different conformation. The access, binding and extrusion monomers are coloured in yellow, green and pink, respectively.

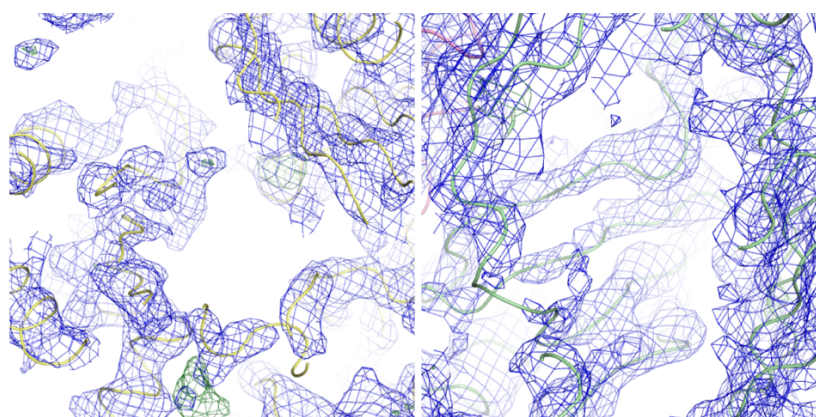


Figure 4.3.5.4. Electron density of the AcrB binding pockets. The proximal and distal binding pockets are shown in the left and right panels, respectively. The blue mesh indicates the 2Fo-Fc electron density maps contoured at 1 σ (carve=1.6 Å). The positive Fo-Fc difference maps contoured at 3 σ (carve=1.6 Å) are shown as green meshes.

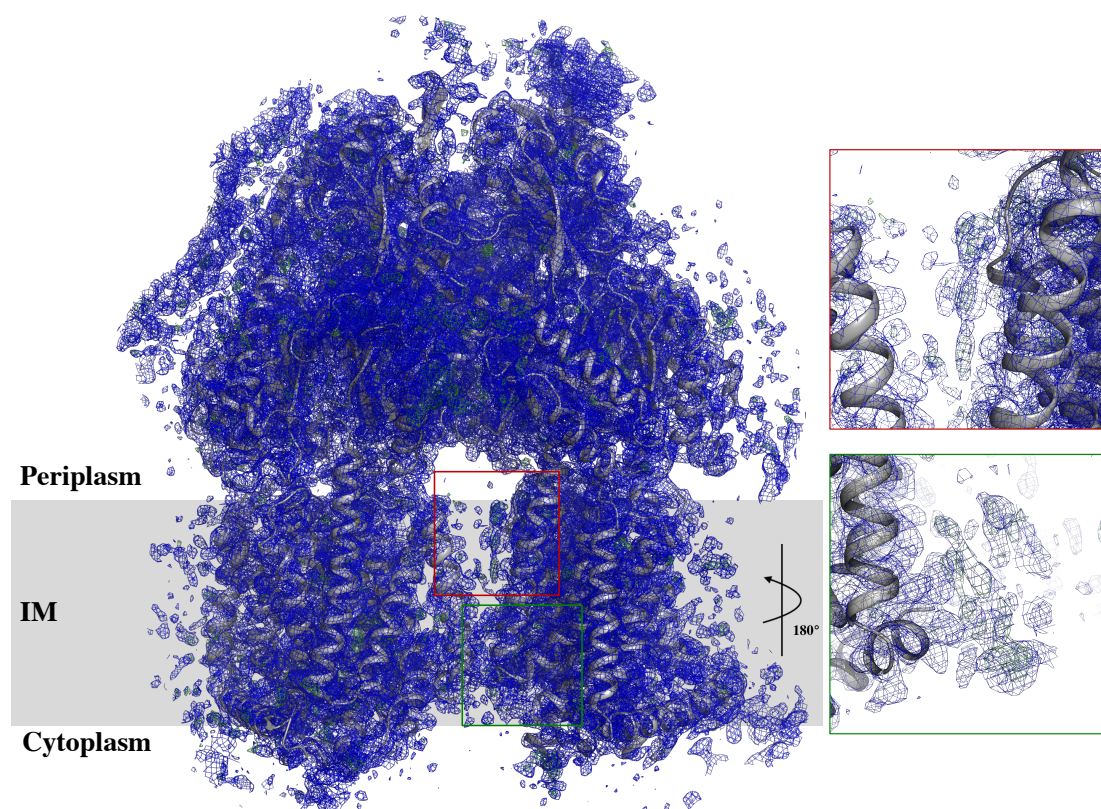


Figure 4.3.5.5. Electron density of the AcrB trimer. The blue and green meshes show the 2Fo-Fc maps (contour=1 σ) and positive Fo-Fc maps (contour=3 σ) (carve=10 Å), respectively. The red inset shows the density for a lipid molecule, while the green inset highlights the density for the head group of DDS molecule (180° rotated). For clarity, a third monomer is omitted.

Both symmetric and asymmetric crystals were tested for soaking and co-crystallisation experiments with the identified inhibitors. However, to date, we still have not obtained an inhibitor-bound structure of AcrB, which could be explained by several reasons. First, both the doxorubicin efflux assay and the MIC tests are whole-cell based, so it remains unknown whether these inhibitors target AcrB specifically or another component of AcrAB-TolC system. Second, the identified inhibitors might be weak binders of AcrB and are therefore lost during crystallisation. Although the determined AcrB structure follows the asymmetric structure, it is possible that the structure does not favour ligand binding due to the composition of the crystallisation buffer and presence of detergent molecules.

4.3.6. Application of XChem for AcrB

Fragment-based drug discovery has been established for over two decades by industrial and academic researchers (Rees *et al.*, 2004). In fragment screening, thousands of low-MW fragments (200-300 Da) are tested by biophysical techniques (*e.g.* SPR) against the target protein. The identified fragments may have low affinity and/or activity, but they provide useful chemical information and usually enable straightforward follow-up synthesis and extension. Fragment screening can be an ideal starting point for drug discovery. X-ray crystallography provides more direct, atomic-level structural information of protein-ligand interactions (Kuntz, 1992). The third-generation synchrotrons are equipped with cutting-edge detectors, robotics and computational instruments, which allows the rapid production of protein crystal structures.

At the Diamond Light Source, scientists from the I04-1 beamline, in partnership with the Structural Genomics Consortium (SGC) at Oxford University, have set up the XChem fragment screening platform, which integrates fragment screening with high-throughput crystal handling, automated X-ray data collection and data analysis (Collins *et al.*, 2018). The XChem team provides several fragment libraries totalling over 2000 compounds available in DMSO. These compounds either allow for rapid follow-up chemistry (*e.g.* DSi-Poised Library, (Cox *et al.*, 2016)) or for 3D properties (*e.g.* Leeds-3D library, (Aimon *et al.*, 2018)). The XChem laboratory is equipped with an acoustic liquid handling system Echo for efficient fragment soaking and the Crystal Shifter Device for fast crystal harvesting (Collins *et al.*, 2018). High-throughput data collection (up to 700 datasets per day) can be done in an automated, unattended mode on the I04-1 beamline at the Diamond Light Source. Data can be integrated automatically by the XChemExplorer and analysed using PanDDA (Krojer *et al.*, 2017; Pearce *et al.*, 2017). XChemExplorer is an SQLite-based graphical data-managing tool that implements programs typically used during data processing, map calculation, ligand identification and refinement to

support large-scale simultaneous analysis of protein-ligand complexes (Krojer *et al.*, 2017). PanDDA is a density modification algorithm that automatically eliminates strong electron density with no statistical significance by subtracting the ground-state density provided by dozens (>30) of apo crystals. Therefore, regions that are unique to each dataset (changed-state) can be objectively identified from statistical analysis of density distributions. PanDDA dramatically improves signal-to-noise and allows subtle changes from ground-state to be modelled (Pearce *et al.*, 2017). The development of XChemExplorer and PanDDA have simplified the challenges and significantly increased the speed and sensitivity in hit detection (Krojer *et al.*, 2017; Pearce *et al.*, 2017).

XChem fragment screening has become a well-established, powerful approach in early stage drug (lead) discovery (Collins *et al.*, 2018). Using the XChem fragment screening platform, it is possible to soak 1000 crystals with different fragments and obtain structures of the protein in complex with hit fragments within a week. Over the years, the XChem fragment screening platform has witnessed great success in lead identification of many soluble protein targets which proves the potency and efficiency of this system. For an XChem project to be successful, it is important to produce reproducible high-quality crystals. Many soluble proteins can diffract consistently to ~ 1.5 Å, which permits high confidence in ligand identification and detailed interpretation of protein-ligand interactions. However, the situation is always more challenging for membrane proteins (Carpenter *et al.*, 2008).

Membrane proteins are often fragile and unstable, and their surface is relatively hydrophobic. Detergents are always needed to extract membrane proteins from the cell membrane to maintain protein stability and solubility by covering the hydrophobic region of the protein. Due to the presence of detergent micelles, membrane protein crystals usually have a high solvent content. Consequently, membrane protein crystals are often unstable, anisotropic, variable in

crystal quality and usually diffract to low resolutions. This has been reflected by the high sensitivity of AcrB crystals during the initial XChem test and soaking experiments. For a membrane protein project, a large number of crystals have to be pre-screened before a high-quality dataset can be collected.

AcrB is the first example of a membrane protein examined on the XChem fragment screening platform and multiple challenges have been encountered when applying this system for a membrane protein. After extensive crystal optimisation, asymmetric AcrB crystals have been produced that diffract constantly to ~ 3 Å (section 4.3.5). In terms of resolution and consistency, AcrB crystals are not comparable to most of the soluble protein projects tested on the XChem platform. However, the significant improvement in AcrB crystallisation has allowed for an initial trial of AcrB using the XChem fragment screening platform. Extensive work has been done to establish the optimal DMSO concentration and cryoprotectant tolerated by AcrB crystals as well as finding suitable compound soaking and crystal harvesting times. For membrane protein crystals, the detergent concentration must be higher than its critical micelle concentration (CMC), the lowest concentration for detergent micelles to be formed. A detergent concentration lower than its CMC leads to crystal melting. 300 nL of glycerol stock (50%) was initially added to the crystal drops to reach a final concentration of 30%. However, this significantly reduced the concentration of DDS and the crystals melt upon glycerol addition. Subsequently, 50% glycerol stock was prepared by diluting 100% glycerol with a crystallisation reservoir containing 0.1% DDS, and the crystals maintained their morphology. Additionally, crystal quality dropped significantly once the incubation time exceeded one hour. Therefore, crystal soaking and harvesting was limited to a maximum of 60 minutes to maintain crystal quality.

To date, AcrB-DDS crystals have been soaked with the six identified inhibitors and ~300 fragments from the DSi-Poised Library (Cox *et al.*, 2016). The DSi-Poised Library contains 769 poised fragments (Table 4.2.1.1). Each of these fragments contains at least one functional group, which allows rapid synthesis and extension for more potent inhibitors. XChem soaking experiment with the six identified inhibitors were performed in triplicate which resulted in 18 datasets. 169 datasets were collected from the AcrB crystals soaked with the DSi-Poised fragments. 60 datasets were collected from different apo AcrB crystals to generate an averaged ground-state map (Pearce *et al.*, 2017). ~131 crystals were lost due to low resolution, which reflected the variable qualities of membrane protein crystals. Results from the XChem fragment screening platform are summarised in Table 4.3.6.1. The obtained 247 datasets were imported to XChemExplorer and PanDDA for evaluation and analysis, but an error was reported at a very early stage of the PanDDA run. Four possible explanations have been proposed, which include (1) relatively low resolution, (2) variable locations of detergent molecules, (3) the relatively large size of AcrB and (4) the crystal packing that results in two molecules of the AcrB trimer in the asymmetric unit. The problem has been reported to the developer for further investigation. At the same time, manual ligand inspection is ongoing, but nothing was observed in the binding pocket of AcrB. The fragments tested so far may not bind to AcrB or bind to AcrB with low affinity. The short soaking time might have limited ligand binding. Certain components in the crystallisation buffer might have also hindered ligand binding to AcrB. Additionally, it is possible that manual inspection could not identify weak binders, in which case it would be worthwhile to analyse the data again using PanDDA once an updated version is released.

	Soaking concentration	NO. of datasets collected
Apo crystal	0	60
Six inhibitors	500 mM	18
DSi-posed library	25 mM	169

Table 4.3.6.1. Results from the XChem fragment screening.

4.4. Discussion

In this work, we describe the preliminary identification and characterisation of a compound as a potential EPI of the AcrAB-TolC efflux pump, which is the major reason for MDR in *E. coli* and other *Enterobacteriaceae* (Piddock, 2006). Six promising candidates were identified through a fluorescence-based doxorubicin transport assay against a small molecule library (Figure 4.3.2.2). Subsequent MIC testing demonstrated that compound 6 increased susceptibility of *E. coli* cells to rifampicin by 4-fold while it did not potentiate carbenicillin, which is not a substrate of AcrAB-TolC (Table 4.3.3.1). Compound 6 is probably a competitive substrate that competes with doxorubicin and rifampicin for the substrate binding sites. Previous studies have shown that NMP, PA β N and MBX2319 potentiated several antibiotics (Lomovskaya *et al.*, 2001; Kern *et al.*, 2006; Opperman *et al.*, 2014). However, our data show that compound 6 is more specific to rifampicin (Table 4.3.3.1). This has raised the potential of developing an EPI that can be specifically used in combination therapy with rifampicin, hereby reducing the acquired resistance (Poole, 2005; Piddock, 2006). Additionally, compound 6 appears to be a novel EPI as it contains a hexahydrophenanthren scaffold (Figure 4.3.2.1), which is structurally distinct to any of the previously described EPIs, including PA β N, NMP, ABI-PP and MBX2319 (Figure 4.1.2.1). Structural comparison of compound 6 and other EPIs reveals that EPIs of AcrAB-TolC usually contain at least two hydrophobic rings, which probably interact with hydrophobic residues in the binding pocket of AcrB (Nakashima *et al.*, 2013). As described by Lomovskaya *et al.*, an EPI of AcrAB-TolC has to satisfy the following criteria: (1) must enhance the antibacterial activity of antibiotics that are transported by AcrAB-TolC, (2) must not potentiate antibiotics that are not transported, (3) must not show potentiation activity against a strain lacking a functional AcrAB-TolC pump, (4) must increase the level of

accumulation of AcrAB-TolC substrates, and (5) must not affect the energy source of the pump (*i.e.* proton motive force) (Lomovskaya *et al.*, 2001). Future work will test the potentiation activity of compound 6 on doxorubicin and other antibiotics against *E. coli* cells with/without the AcrAB-TolC efflux pump. Providing the experimental set up is available, we want to test whether compound 6 can disrupt the proton gradient across the *E. coli* inner membrane.

Co-crystallisation and soaking experiments with the six identified inhibitors were performed in order to obtain an inhibitor-bound AcrB structure. At the same time, we managed to produce high-quality AcrB crystals that diffract consistently to ~ 3 Å. In this structure, each monomer follows a different conformation and both the proximal and distal binding pockets are empty (Figure 4.3.5.3 and Figure 4.3.5.4). This means that this crystal form is physiologically relevant and favours ligand binding. In order to develop inhibitors that specifically inhibit AcrB, we used the XChem fragment screening platform to soak AcrB crystals with the six identified inhibitors and libraries of different fragments. Although until now no hit has been identified, our work using the XChem fragment screening platform has provided valuable information for future inhibitor development of AcrB. AcrB is the first example of a membrane protein examined on this platform, which has provided precious experience for application of the XChem fragment screening platform for membrane proteins.

Providing further funding can be available, future work will examine binding of libraries of different compounds to AcrB using either SPR or the thermal stability assay (TSA) (Alexandrov *et al.*, 2008; Mowla *et al.*, 2018). In SPR experiments, AcrB can be immobilised to a sensor surface. When analyte containing the compound flows past and binds to the protein, the refractive index will change owing to compound accumulation at the sensor surface. A real-time measurement of the refractive index change generates a response versus time plot. Association and dissociation constants between different compounds and AcrB can be

determined by SPR. In TSA of membrane proteins, a fluorescent sensor N-[4-(7-diethylamino-4-methyl-3-coumarinyl) phenyl] maleimide (CPM) is commonly used (Mowla *et al.*, 2018). Cysteines are typically buried in the interior of a folded protein as they are hydrophobic. When a protein denatures at higher temperature, thiol groups of cysteine become available and a fluorescent signal can be read from reacted CPM. Subsequently, the transition unfolding temperature (T_m) can be recorded. Changes in transition melting temperature (ΔT_m) can be used to estimate the binding affinity between AcrB and different compounds. Once AcrB binders are obtained, doxorubicin efflux assay and MIC tests can be performed to validate the clinical potentials of the identified compounds. EPIs will be co-crystallised with AcrB. Once a complex structure is obtained, residues that interact with the ligands will be mutated and AcrB mutants will be expressed and purified. SPR/TSA, efflux assay and MIC tests will be performed to validate the importance of these residues. Hit compounds with higher affinity and activity can be tested on microbes overexpressing other multidrug transporters such as MexB and MexY. The target spectrum and specificity of these ligands can therefore be determined.

Although potential EPIs of AcrAB-TolC can be identified by whole-cell assays, their mechanisms of action remain unknown. Extensive crystal soaking has been done in order to develop specific AcrB inhibitors, whereas nothing promising has been identified to date. The lack of positive results has raised the question of whether AcrB is a good target for EPI development. Crystallisation and biochemical assays of AcrB are challenging, thus hindering large-scale inhibitor screening and further optimisation. The MFP AcrA is a periplasmic soluble protein that is essential for the functional assembly of AcrAB-TolC. Therefore, AcrA is probably a more feasible target for inhibitor development since it is generally easier to tackle a soluble protein. TolC is also an attractive target since it is recruited by a range of different multidrug transporters (*e.g.* MacAB, MexAB and EmrAB), inhibitors of which are likely to have a broad target spectrum (Koronakis *et al.*, 2004).

5. Conclusion and outlook

To date, available structural and biochemical data have revealed molecular details governing substrate recognition and transport of different transporters (Du *et al.*, 2018). However, given the large number of membrane transporters and their broad spectrum of substrates, it is likely that our knowledge is still limited to the more representative members. The work from this thesis has enabled us to describe several survival strategies utilised by bacteria for nutrient acquisition and drug efflux. In chapter 2, we have determined the structure of TauA in complex with different ligands and revealed the structural mechanism governing substrate specificity of TauA. In chapter 3, we have presented the structure of peptide-bound YejA, suggesting YejA is likely to be an oligopeptide-binding protein. Conjugation of the peptide can potentially improve drug discovery profiles of conventional antibiotics. A second survival strategy developed by bacterial cells is the employment of multidrug transporters that confer antibiotic resistance. In chapter 4, we have presented the application of AcrB to the XChem fragment screening platform. AcrB represents the first example of a membrane protein tested using this platform, providing a novel strategy of inhibitor development for membrane protein targets. The lack of positive data in chapter 4 has shown that whole-cell assay is probably not a suitable approach for inhibitor screening, especially for a tripartite system like AcrAB-TolC, since it is difficult to establish the mechanism of inhibition. We propose to look into AcrA as a target for future inhibitor development against the AcrAB-TolC efflux pump.

6. Reference

- Abramson, J., Smirnova, I., Kasho, V., Verner, G., *et al.* (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science*. 301 (5633), 610–615.
- Adelman, K., Yuzenkova, J., La Porta, A., Zenkin, N., *et al.* (2004) Molecular mechanism of transcription inhibition by peptide antibiotic microcin J25. *Molecular Cell*. 14 (6), 753–762.
- Agarwal, V., Metlytskaya, A., Severinov, K.V. & Nair, S.K. (2011) Structural basis for Microcin C7 inactivation by the MccE acetyltransferase. *Journal of Biological Chemistry*. 286(24), 21295–21303.
- Agarwal, V., Tikhonov, A., Metlitskaya, A., Severinov, K., *et al.* (2012) Structure and function of a serine carboxypeptidase adapted for degradation of the protein synthesis antibiotic microcin C7. *Proceedings of the National Academy of Sciences*. 109 (12), 4425–4430.
- Aimon, A., Karageorgis, G., Masters, J., Dow, M., *et al.* (2018) Realisation of small molecule libraries based on frameworks distantly related to natural products. *Organic and Biomolecular Chemistry* 16 (17), 3160–3167.
- Alberts, B., Bray, D., Hopkin, K., Johnson, A.D., *et al.* (2013) Essential cell biology. *Garland Science*.
- Alexandrov, A.I., Mileni, M., Chien, E.Y.T., Hanson, M.A., *et al.* (2008) Microscale fluorescent thermal stability assay for membrane proteins. *Structure*. 16 (3), 351–359.
- Allegretti, M., Klusch, N., Mills, D.J., Vonck, J., *et al.* (2015) Horizontal membrane-intrinsic α -helices in the stator a-subunit of an F-type ATP synthase. *Nature*. 521, 237.
- Ames, Giovanna Ferro-Luzzi, *et al.* (1992) Traffic ATPases: a superfamily of transport proteins operating from *Escherichia coli* to humans. *Advances in Enzymology and Related Areas of Molecular Biology*. 65, 1–47.
- Andersen, C., Koronakis, E., Bokma, E., Eswaran, J., *et al.* (2002) Transition to the open state of the TolC periplasmic tunnel entrance. *Proceedings of the National Academy of Sciences*. 99 (17), 11103–11108.
- Antal, M., Bordeau, V., Douchin, V. & Felden, B. (2005) A small bacterial RNA regulates a putative ABC transporter. *Journal of Biological Chemistry*. 280 (9), 7901–7908.
- Askoura, M., Mattawa, W., Abujamel, T. & Taher, I. (2011) Efflux pump inhibitors (EPIs) as new antimicrobial agents against *Pseudomonas aeruginosa*. *Libyan Journal of Medicine*. 6 (1), 5870.
- Baquero, F., Bouanchaud, D., Martinez-Perez, M.C. & Fernandez, C. (1978) Microcin plasmids: a group of extrachromosomal elements coding for low-molecular-weight antibiotics in *Escherichia coli*. *Journal of Bacteriology*. 135 (2), 342 LP – 347.

- Baranova, N. & Nikaido, H. (2002) The BaeSR two-component regulatory system activates transcription of the yegMNOB (mdtABCD) transporter gene cluster in *Escherichia coli* and increases its resistance to novobiocin and deoxycholate. *Journal of Bacteriology*. 184 (15), 4168–4176.
- Baucheron, S., Imberechts, H., Chaslus-Dancla, E. & Cloeckaert, A. (2002) The AcrB multidrug transporter plays a major role in high-level fluoroquinolone resistance in *Salmonella enterica* serovar Typhimurium phage type DT204. *Microbial Drug Resistance*. 8 (4), 281–289.
- Bay, D.C., Rommens, K.L. & Turner, R.J. (2008) Small multidrug resistance proteins: a multidrug transporter family that continues to grow. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 1778 (9), 1814–1838.
- Beale, J., Lee, S.Y., Iwata, S. & Beis, K. (2010) Structure of the aliphatic sulfonate-binding protein SsuA from *Escherichia coli*. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*. 66 (4), 391–396.
- Becker, M.L., Visser, L.E., van Schaik, R.H.N., Hofman, A., *et al.* (2009) Genetic variation in the multidrug and toxin extrusion 1 transporter protein influences the glucose-lowering effect of metformin in patients with diabetes: a preliminary study. *Diabetes*. 58 (3), 745–749.
- Beis, K. (2015) Structural basis for the mechanism of ABC transporters. *Biochemical Society Transactions*. 43 (5), 889–893.
- Berntsson, R.P.-A., Doeven, M.K., Fusetti, F., Duurkens, R.H., *et al.* (2009) The structural basis for peptide selection by the transport receptor OppA. *The EMBO Journal*. 28 (9), 1332–1340.
- Berntsson, R.P.-A., Smits, S.H.J., Schmitt, L., Slotboom, D.-J., *et al.* (2010) A structural classification of substrate-binding proteins. *FEBS Letters*. 584 (12), 2606–2617.
- Bieler, S., Silva, F., Soto, C. & Belin, D. (2006) Bactericidal activity of both secreted and nonsecreted microcin E492 requires the mannose permease. *Journal of Bacteriology*. 188 (20), 7049 LP – 7061.
- Bjorkmans, A.J., Binnieon, R.A., Zhangb, H., Colell, L.B., *et al.* (1994) The ribose-binding protein in bacterial transport and chemotaxis. *The Journal of Biological Chemistry*. 30206–30211.
- Bohnert, J.A. & Kern, W. V (2005) Selected arylpiperazines are capable of reversing multidrug resistance in *Escherichia coli* overexpressing RND efflux pumps. *Antimicrobial Agents and Chemotherapy*. 49 (2), 849–852.
- Borths, E.L., Locher, K.P., Lee, A.T. & Rees, D.C. (2002) The structure of *Escherichia coli* BtuF and binding to its cognate ATP binding cassette transporter. *Proceedings of the national academy of sciences*. 99 (26), 16642–16647.
- Bountra, K., Hagelueken, G., Choudhury, H.G., Corradi, V., *et al.* (2017) Structural basis for antibacterial peptide self-immunity by the bacterial ABC transporter McjD. *The EMBO Journal*. 36 (20), 3062–3079.

- Browning, C., Martin, E., Loch, C., Wurtz, J.-M., *et al.* (2007) Critical role of desolvation in the binding of 20-hydroxyecdysone to the ecdysone receptor. *Journal of Biological Chemistry*. 282 (45), 32924–32934.
- Buchanan, S.K., Smith, B.S., Venkatramani, L., Xia, D., *et al.* (1999) Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*. *Nature Structural and Molecular Biology*. 6 (1), 56.
- Carpenter, E.P., Beis, K., Cameron, A.D. & Iwata, S. (2008) Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology*. 18 (5), 581–586.
- Cha, H., Müller, R.T. & Pos, K.M. (2014) Switch-loop flexibility affects transport of large drugs by the promiscuous AcrB multidrug efflux transporter. *Antimicrobial Agents and Chemotherapy*. 58 (8), 4767–4772.
- Chacón, K.N., Mealman, T.D., McEvoy, M.M. & Blackburn, N.J. (2014) Tracking metal ions through a Cu/Ag efflux pump assigns the functional roles of the periplasmic proteins. *Proceedings of the National Academy of Sciences*. 111 (43), 15373–15378.
- Chen, Y.-J., Pornillos, O., Lieu, S., Ma, C., *et al.* (2007) X-ray structure of EmrE supports dual topology model. *Proceedings of the National Academy of Sciences*. 104 (48), 18999–19004.
- Chiuchiolo, M.J., Delgado, M.A., Farías, R.N. & Salomón, R.A. (2001) Growth-phase-dependent expression of the cyclopeptide antibiotic microcin J25. *Journal of Bacteriology*. 183 (5), 1755 LP – 1764.
- Choudhury, H.G., Tong, Z., Mathavan, I., Li, Y., *et al.* (2014) Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. *Proceedings of the National Academy of Sciences*. 111 (25), 9145–9150.
- Coban, A.Y., Cayci, Y.T., Erturan, Z. & Durupinar, B. (2009) Effects of efflux pump inhibitors Phenyl-Arginine- β -Naphthylamide and 1-(1-Naphthylmethyl)-Piperazine on the antimicrobial susceptibility of *Pseudomonas aeruginosa* Isolates from cystic fibrosis patients. *Journal of Chemotherapy*. 21 (5), 592–594.
- Collins, P.M., Douangamath, A., Talon, R., Dias, A., *et al.* (2018) Achieving a good crystal system for crystallographic X-ray fragment screening. *Methods in Enzymology*. 610, 251–264.
- Colomer-Lluch, M., Jofre, J. & Muniesa, M. (2011) Antibiotic resistance genes in the bacteriophage DNA fraction of environmental samples. *PLoS One*. 6 (3), 1–11.
- Connell, N., Han, Z., Moreno, F. & Kolter, R. (1987) An *E. coli* promoter induced by the cessation of growth. *Molecular Microbiology*. 1 (2), 195–201.
- Connell, S.R., Tracz, D.M., Nierhaus, K.H. & Taylor, D.E. (2003) Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrobial Agents and Chemotherapy*. 47 (12), 3675–3681.
- Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*. 16 (22), 10881–10890.

Cotter, P.D., Ross, R.P. & Hill, C. (2013) Bacteriocins—a viable alternative to antibiotics? *Nature Reviews Microbiology*. 11 (2), 95.

Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D: Biological Crystallography*. 62 (9), 1002–1011.

Cox, O.B., Krojer, T., Collins, P., Monteiro, O., *et al.* (2016) A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chemical Sciences* 7 (3), 2322–2330.

Cuneo, M.J., Changela, A., Warren, J.J., Beese, L.S., *et al.* (2006) The crystal structure of a thermophilic glucose binding protein reveals adaptations that interconvert mono and disaccharide binding sites. *Journal of Molecular Biology*. 362 (2), 259–270.

Dadashpour, S., Tuylu Kucukkilinc, T., Unsal Tan, O., Ozadali, K., *et al.* (2015) Design, synthesis and *in vitro* study of 5,6-diaryl-1,2,4-triazine-3-ylthioacetate derivatives as COX-2 and β -amyloid aggregation inhibitors. *Archiv der Pharmazie*. 348 (3), 179–187.

Dassama, L.M.K., Kenney, G.E., Ro, S.Y., Zielazinski, E.L., *et al.* (2016) Methanobactin transport machinery. *Proceedings of the National Academy of Sciences*. 113 (46), 13027–13032.

Davidson, A.L., Shuman, H.A. & Nikaido, H. (1992) Mechanism of maltose transport in *Escherichia coli*: transmembrane signaling by periplasmic binding proteins. *Proceedings of the National Academy of Sciences*. 89 (6), 2360–2364.

Dawson, R.J.P. & Locher, K.P. (2006) Structure of a bacterial multidrug ABC transporter. *Nature*. 443, 180.

Dawson, R.J.P. & Locher, K.P. (2007) Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. *FEBS Letters*. 581 (5), 935–938.

Dean, M. & Annilo, T. (2005) Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. *Annual Review of Genomics and Human Genetics*. 6 (1), 123–142.

Delgado, M.A., Solbiati, J.O., Chiuchiolo, M.J., Farias, R.N., *et al.* (1999) *Escherichia coli* outer membrane protein TolC is involved in production of the peptide antibiotic microcin J25. *Journal of Bacteriology*. 181 (6), 1968–1970.

Deng, D., Xu, C., Sun, P., Wu, J., *et al.* (2014) Crystal structure of the human glucose transporter GLUT1. *Nature*. 510, 121.

Destoumieux-Garzón, D., Peduzzi, J., Thomas, X., Djediat, C., *et al.* (2006) Parasitism of iron-siderophore receptors of *Escherichia coli* by the siderophore-peptide microcin E492m and its unmodified counterpart. *BioMetals*. 19 (2), 181–191.

Detmers, F.J.M., Lanfermeijer, F.C., Abele, R., Jack, R.W., *et al.* (2000) Combinatorial peptide libraries reveal the ligand-binding mechanism of the oligopeptide receptor OppA of *Lactococcus lactis*. *Proceedings of the National Academy of Sciences*. 97 (23), 12487–12492.

- 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.
- Diez, J., Diederichs, K., Greller, G., Horlacher, R., *et al.* (2001) The crystal structure of a liganded trehalose/maltose-binding protein from the hyperthermophilic archaeon *Thermococcus litoralis* at 1.85 Å. *Journal of Molecular Biology*. 305 (4), 905–915.
- Dinh, T., Paulsen, I.T. & Saier, M.H. (1994) A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of gram-negative bacteria. *Journal of Bacteriology*. 176 (13), 3825–3831.
- DiSpirito, A.A., Zahn, J.A., Graham, D.W., Kim, H.J., *et al.* (1998) Copper-binding compounds from *Methylosinus trichosporium OB3b*. *Journal of Bacteriology*. 180 (14), 3606–3613.
- Doerrler, W.T., Gibbons, H.S. & Raetz, C.R.H. (2004) MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. *Journal of Biological Chemistry*. 279 (43), 45102–45109.
- Doeven, M.K., Abele, R., Tampé, R. & Poolman, B. (2004) The binding specificity of OppA determines the selectivity of the oligopeptide ATP-binding cassette transporter. *Journal of Biological Chemistry*. 279 (31), 32301–32307.
- Dong, J., Yang, G. & Mchaourab, H.S. (2005) Structural basis of energy transduction in the transport cycle of MsbA. *Science*. 308 (5724), 1023–1028.
- Du, D., Wang-Kan, X., Neuberger, A., van Veen, H.W., *et al.* (2018) Multidrug efflux pumps: structure, function and regulation. *Nature Reviews Microbiology*. 16 (9), 523–539.
- Du, D., Wang, Z., James, N.R., Voss, J.E., *et al.* (2014) Structure of the AcrAB-TolC multidrug efflux pump. *Nature*. 509 (7501), 512–515.
- Duquesne, S., Destoumieux-Garzón, D., Peduzzi, J. & Rebuffat, S. (2007) Microcins, gene-encoded antibacterial peptides from enterobacteria. *Natural Product Reports*. 24 (4), 708–734.
- Eichhorn, E., van der Ploeg, J.R., Kertesz, M.A. & Leisinger, T. (1997) Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *The Journal of Biological Chemistry*. 272 (37), 23031–23036.
- Eichhorn, E., Van Der Ploeg, J.R. & Leisinger, T. (1999) Characterization of a two-component alkanesulfonate monooxygenase from *Escherichia coli*. *Journal of Biological Chemistry*. 274 (38), 26639–26646.
- Eichhorn, E., Van Der Ploeg, J.R. & Leisinger, T. (2000) Deletion analysis of the *Escherichia coli* taurine and alkanesulfonate transport systems. *Journal of Bacteriology*. 182 (10), 2687–2695.

- Elkins, C. A. & Nikaido, H. (2002) Substrate specificity of the RND-type multidrug efflux pumps AcrB and AcrD of *Escherichia coli* is determined predominately by two large periplasmic loops. *Journal of Bacteriology*. 184 (23), 6490–6498.
- Emsley, P. & Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*. 60 (12 I), 2126–2132.
- Ernst, R., Kueppers, P., Klein, C.M., Schwarzmüller, T., *et al.* (2008) A mutation of the H-loop selectively affects rhodamine transport by the yeast multidrug ABC transporter Pdr5. *Proceedings of the National Academy of Sciences*. 105 (13), 5069–5074.
- Fischer, M., Zhang, Q.Y., Hubbard, R.E. & Thomas, G.H. (2010) Caught in a TRAP: substrate-binding proteins in secondary transport. *Trends in Microbiology*. 18 (10), 471–478.
- Fitzpatrick, A.W.P., Llabrés, S., Neuberger, A., Blaza, J.N., *et al.* (2017) Structure of the MacAB–TolC ABC-type tripartite multidrug efflux pump. *Nature Microbiology*. 2, 17070.
- Fleishman, S.J., Harrington, S.E., Enosh, A., Halperin, D., *et al.* (2006) Quasi-symmetry in the cryo-EM structure of EmrE provides the key to modeling its transmembrane domain. *Journal of Molecular Biology*. 364 (1), 54–67.
- Foadi, J., Aller, P., Alguel, Y., Cameron, A., *et al.* (2013) Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*. 69 (8), 1617–1632.
- Fomenko, D., Veselovskii, A. & Khmel, I. (2001) Regulation of microcin C51 operon expression: the role of global regulators of transcription. *Research in Microbiology*. 152 (5), 469–479.
- Fomenko, D.E., Metlitskaya, A.Z., Péduzzi, J., Goulard, C., *et al.* (2003) Microcin C51 plasmid genes: possible source of horizontal gene transfer. *Antimicrobial Agents and Chemotherapy*. 47 (9), 2868–2874.
- Franklund, C. V., Baron, S.F. & Hylemon, P.B. (1993) Characterization of the *baiH* gene encoding a bile acid-inducible NADH: flavin oxidoreductase from *Eubacterium sp.* strain VPI 12708. *Journal of Bacteriology*. 175 (10), 3002–3012.
- Fredrickson, A.G. & Stephanopoulos, G. (1981) Microbial competition. *Science*. 213 (4511), 972–979.
- Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. (1999) Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *Journal of Molecular Biology*. 286 (1), 279–290.
- Furukawa, H., Tsay, J.T., Jackowski, S., Takamura, Y., *et al.* (1993) Thiolactomycin resistance in *Escherichia coli* is associated with the multidrug resistance efflux pump encoded by *emrAB*. *Journal of Bacteriology*. 175 (12), 3723–3729.
- Furuya, E.Y. & Lowy, F.D. (2006) Antimicrobial-resistant bacteria in the community setting. *Nature Reviews Microbiology*. 4, 36.

Gaggero, C., Moreno, F. & Laviña, M. (1993) Genetic analysis of microcin H47 antibiotic system. *Journal of Bacteriology*. 175 (17), 5420 LP – 5427.

Garcia-Bustos, J. & Tomasz, A (1990) A biological price of antibiotic resistance: major changes in the peptidoglycan structure of penicillin-resistant pneumococci. *Proceedings of the National Academy of Sciences*. 87 (14), 5415–5419.

Garcia-Bustos, J.F., Pezzi, N. & Asensio, C. (1984) Microcin 7: purification and properties. *Biochemical and Biophysical Research Communications*. 119 (2), 779–785.

Garcia-Bustos, J.F., Pezzi, N. & Mendez, E. (1985) Structure and mode of action of microcin 7, an antibacterial peptide produced by *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*. 27 (5), 791 LP – 797.

Garrido, M.C., Herrero, M., Kolter, R. & Moreno, F. (1988) The export of the DNA replication inhibitor microcin B17 provides immunity for the host cell. *The EMBO Journal*. 7 (6), 1853–1862.

Gerber, S., Comellas-Bigler, M., Goetz, B.A. & Locher, K.P. (2008) Structural basis of trans-inhibition in a molybdate/tungstate ABC transporter. *Science*. 321 (5886), 246–250.

Giamarellou, H. & Poulakou, G. (2009) Multidrug-resistant gram-negative infections. *Drugs*. 69 (14), 1879–1901.

Gilson, L., Mahanty, H.K. & Kolter, R. (1990) Genetic analysis of an MDR-like export system: the secretion of colicin V. *The EMBO Journal*. 9 (12), 3875–3884.

Gonzales, T. & Robert-Baudouy, J. (1996) Bacterial aminopeptidases: Properties and functions. *FEMS Microbiology Reviews*. 18 (4), 319–344.

Gonzalez-Pastor, J.E., San Millán, J.L., Castilla, M.A., Moreno, F., *et al.* (1995) Structure and organization of plasmid genes required to produce the translation inhibitor microcin C7. *Journal of Bacteriology*. 177 (24), 7131–7140.

González-Pastor, J.E., San Millán, J.L. & Moreno, F. (1994) The smallest known gene. *Nature*. 369, 281.

Gouridis, G., Schuurman-Wolters, G.K., Ploetz, E., Husada, F., *et al.* (2015) Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. *Nature Structural and Molecular Biology*. 22 (1), 57–64.

Grover, S., Apushkin, M.A. & Fishman, G.A. (2006) Topical dorzolamide for the treatment of cystoid macular edema in patients with retinitis pigmentosa. *American Journal of Ophthalmology*. 141 (5), 850–858.

Guan, L. & Kaback, H.R. (2004) Binding affinity of lactose permease is not altered by the H⁺ electrochemical gradient. *Proceedings of the National Academy of Sciences*. 101 (33), 12148 LP – 12152.

- Guijarro, J.I., González-Pastor, J.E., Baleux, F., San Millán, J.L., *et al.* (1995) Chemical structure and translation inhibition studies of the antibiotic microcin C7. *Journal of Biological Chemistry*. 270 (40), 23520–23532.
- Hannula, M. & Hänninen, M.-L. (2008) Effect of putative efflux pump inhibitors and inducers on the antimicrobial susceptibility of *Campylobacter jejuni* and *Campylobacter coli*. *Journal of Medical Microbiology*. 57 (7), 851–855.
- Hasdemir, U.O., Chevalier, J., Nordmann, P. & Pagès, J.-M. (2004) Detection and prevalence of active drug efflux mechanism in various multidrug-resistant *Klebsiella pneumoniae* strains from Turkey. *Journal of Clinical Microbiology*. 42 (6), 2701 LP – 2706.
- Hassan, K.A., Jackson, S.M., Penesyan, A., Patching, S.G., *et al.* (2013) Transcriptomic and biochemical analyses identify a family of chlorhexidine efflux proteins. *Proceedings of the National Academy of Sciences*. 110 (50), 20254–20259.
- Hassan, K.A., Liu, Q., Henderson, P.J.F. & Paulsen, I.T. (2015) Homologs of the *Acinetobacter baumannii* AceI transporter represent a new family of bacterial multidrug efflux systems. *mBio*. 6 (1).
- Hawkey, P.M. (1998) The origins and molecular basis of antibiotic resistance. *British Medical Journal*. 317 (7159), 657.
- He, X., Szewczyk, P., Karyakin, A., Evin, M., *et al.* (2010) Structure of a cation-bound multidrug and toxic compound extrusion transporter. *Nature*. 467, 991.
- Heng, J., Zhao, Y., Liu, M., Liu, Y., *et al.* (2015) Substrate-bound structure of the *E. coli* multidrug resistance transporter MdfA. *Cell Research*. 25, 1060.
- Ho, H., Miu, A., Alexander, M.K., Garcia, N.K., *et al.* (2018) Structural basis for dual-mode inhibition of the ABC transporter MsbA. *Nature*. 557 (7704), 196–201.
- Hobbs, E.C., Yin, X., Paul, B.J., Astarita, J.L., *et al.* (2012) Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance. *Proceedings of the National Academy of Sciences*. 109 (41), 16696–16701.
- Hollenstein, K., Dawson, R.J.P. & Locher, K.P. (2007) Structure and mechanism of ABC transporter proteins. *Current Opinion in Structural Biology*. 17 (4), 412–418.
- Hollenstein, K., Frei, D.C. & Locher, K.P. (2007) Structure of an ABC transporter in complex with its binding protein. *Nature*. 446, 213.
- Holm, L. & Sander, C. (1995) Dali: a network tool for protein structure comparison. *Trends in Biochemical Sciences*. 20 (June), 478–480.
- Hong, J., Kim, T. & Kim, J. (2014) Tertiary matrices for the analysis of polyethylene glycols using MALDI-TOF MS. *Mass Spectrometry Letters*. 5 (2), 49–51.
- Hu, Y., Rech, S., Gunsalus, R.P. & Rees, D.C. (1997) Crystal structure of the molybdate binding protein ModA. *Nature Structural and Molecular Biology*. 4 (9), 703.

Husain, F., Bikhchandani, M. & Nikaido, H. (2011) Vestibules are part of the substrate path in the multidrug efflux transporter AcrB of *Escherichia coli*. *Journal of Bacteriology*. 193 (20), 5847–5849.

Huxtable, R.J. (1992) Physiological actions of taurine. *Physiological Reviews*. 72 (1), 101–163.

Hvorup, R.N., Goetz, B.A., Niederer, M., Hollenstein, K., *et al.* (2007) Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science*. 317 (5843), 1387–1390.

Inloes, D.S., Michaels, A.S., Robertson, C.R. & Matin, A. (1985) Ethanol production by nitrogen-deficient yeast cells immobilized in a hollow-fiber membrane bioreactor. *Applied Microbiology and Biotechnology*. 23 (2), 85–91.

Jannasch, H.W. & Taylor, C.D. (1984) Deep-sea microbiology. *Annual Reviews in Microbiology*. 38 (1), 487.

Javaux, C., Joris, B. & De Witte, P. (2007) Functional characteristics of TauA binding protein from TauABC *Escherichia coli* system. *Protein Journal*. 26 (4), 231–238.

Jiang, D., Zhao, Y., Wang, X., Fan, J., *et al.* (2013) Structure of the YajR transporter suggests a transport mechanism based on the conserved motif A. *Proceedings of the National Academy of Sciences*. 110 (36), 14664–14669.

Jiang, X., Smirnova, I., Kasho, V., Wu, J., *et al.* (2016) Crystal structure of a LacY–nanobody complex in a periplasmic-open conformation. *Proceedings of the National Academy of Sciences*. 113 (44), 12420 LP – 12425.

Johnston, C., Martin, B., Fichant, G., Polard, P., *et al.* (2014) Bacterial transformation: Distribution, shared mechanisms and divergent control. *Nature Reviews Microbiology*. 12 (3), 181–196.

Jornada, D.H., Dos Santos Fernandes, G.F., Chiba, D.E., De Melo, T.R.F., *et al.* (2016) The prodrug approach: a successful tool for improving drug solubility. *Molecules*. 21 (1).

Kaback, H.R., Dunten, R., Frillingos, S., Venkatesan, P., *et al.* (2007) Site-directed alkylation and the alternating access model for LacY. *Proceedings of the National Academy of Sciences*. 104 (2), 491–494.

Kaback, H.R., Sahin-Tóth, M. & Weinglass, A.B. (2001) The kamikaze approach to membrane transport. *Nature Reviews Molecular Cell Biology*. 2, 610.

Kabsch, W. (2010) Xds. *Acta Crystallographica Section D: Biological Crystallography*. 66 (2), 125–132.

Kadaba, N.S., Kaiser, J.T., Johnson, E., Lee, A., *et al.* (2008) The High-affinity *E. coli* methionine ABC transporter: structure and allosteric regulation. *Science*. 321 (5886), 250–253.

- Kaldor, S.W., Kalish, V.J., Davies, J.F., Shetty, B. V., *et al.* (1997) Viracept (nelfinavir mesylate, AG1343): A potent, orally bioavailable inhibitor of HIV-1 protease. *Journal of Medicinal Chemistry*. 40 (24), 3979–3985.
- Kamio, Y. & Nikaido, H. (1976) Outer membrane of *Salmonella typhimurium*: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry*. 15 (12), 2561–2570.
- Kang, J., Hwang, J.-U., Lee, M., Kim, Y.-Y., *et al.* (2010) PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. *Proceedings of the National Academy of Sciences*. 107 (5), 2355–2360.
- Kapp, K., Schrempf, S., Lemberg, M.K. & Dobberstein, B. (2013) Post-targeting functions of signal peptides. *Madame Curie Bioscience Database*. Landes Bioscience.
- Kato, A., Ohnishi, H., Yamamoto, K., Furuta, E., *et al.* (2000) Transcription of *emrKY* is regulated by the *EvgA-EvgS* two-component system in *Escherichia coli* K-12. *Bioscience, Biotechnology, and Biochemistry*. 64 (6), 1203–1209.
- Kazakov, T., Vondenhoff, G.H., Datsenko, K.A., Novikova, M., *et al.* (2008) *Escherichia coli* peptidase A, B, or N can process translation inhibitor microcin C. *Journal of Bacteriology*. 190 (7), 2607–2610.
- Keller, B.O. & Li, L. (2000) Discerning matrix-cluster peaks in matrix-assisted laser desorption/ionization time-of-flight mass spectra of dilute peptide mixtures. *Journal of the American Society for Mass Spectrometry*. 11 (1), 88–93.
- Kern, W. V., Steinke, P., Schumacher, A., Schuster, S., *et al.* (2006) Effect of 1-(1-naphthylmethyl)-piperazine, a novel putative efflux pump inhibitor, on antimicrobial drug susceptibility in clinical isolates of *Escherichia coli*. *Journal of Antimicrobial Chemotherapy*. 57 (2), 339–343.
- Kertesz, M. a, Leisinger, T. & Cook, A.M. (1993) Proteins induced by sulfate limitation in *Escherichia coli*, *Pseudomonas putida*, or *Staphylococcus aureus*. *Notes*. 175 (4), 1187–1190.
- Kertesz, M.A. (2001) Bacterial transporters for sulfate and organosulfur compounds. *Research in Microbiology*. 152 (3), 279–290.
- Khare, D., Oldham, M.L., Orelle, C., Davidson, A.L., *et al.* (2009) Alternating access in maltose transporter mediated by rigid-body rotations. *Molecular Cell*. 33 (4), 528–536.
- Kim, E., Lee, S., Jeon, A., Choi, J.M., *et al.* (2013) A single-molecule dissection of ligand binding to a protein with intrinsic dynamics. *Nature Chemical Biology*. 9, 313.
- King, J.E. & Quinn, J.P. (1997) The utilization of organosulphonates by soil and freshwater bacteria. *Letters in Applied Microbiology*. 24 (6), 474–478.
- Kjelleberg, S., Albertson, N., Flårdh, K., Holmquist, L., *et al.* (1993) How do non-differentiating bacteria adapt to starvation? *Antonie Van Leeuwenhoek*. 63 (3–4), 333–341.

Klepsch, M.M., Kovermann, M., Löw, C., Balbach, J., *et al.* (2011) *Escherichia coli* peptide binding protein OppA has a preference for positively charged peptides. *Journal of Molecular Biology*. 414 (1), 75–85.

Kobayashi, A., Hirakawa, H., Hirata, T., Nishino, K., *et al.* (2006) Growth phase-dependent expression of drug exporters in *Escherichia coli* and its contribution to drug tolerance. *Journal of Bacteriology*. 188 (16), 5693–5703.

Kobayashi, N., Nishino, K. & Yamaguchi, A. (2001) Novel macrolide-specific ABC-type efflux transporter in *Escherichia coli*. *Journal of Bacteriology*. 183 (19), 5639–5644.

Kobayashi, N., Tamura, N., van Veen, H.W., Yamaguchi, A., *et al.* (2014) β -Lactam selectivity of multidrug transporters AcrB and AcrD resides in the proximal binding pocket. *Journal of Biological Chemistry*. 289 (15), 10680–10690.

Kohler, T., Michea-Hamzehpour, M., Plesiat, P., Kahr, A.L., *et al.* (1997) Differential selection of multidrug efflux systems by quinolones in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*. 41 (11), 2540–2543.

Koide, A. & Hoch, J.A. (1994) Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Molecular Microbiology*. 13 (3), 417–426.

Kolter, R., Siegele, D.A. & Tormo, A. (1993) The stationary phase of the bacterial life cycle. *Annual Reviews in Microbiology*. 47 (1), 855–874.

Korkhov, V.M., Mireku, S. a. & Locher, K.P. (2012) Structure of AMP-PNP-bound vitamin B12 transporter BtuCD–F. *Nature*. 490 (7420), 367–372.

Korkhov, V.M., Mireku, S.A., Veprintsev, D.B. & Locher, K.P. (2014) Structure of AMP-PNP-bound BtuCD and mechanism of ATP-powered vitamin B12 transport by BtuCD–F. *Nature Structural and Molecular Biology*. 21, 1097.

Koronakis, V., Eswaran, J. & Hughes, C. (2004) Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annual Review of Biochemistry*. 73 (1), 467–489.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., *et al.* (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature*. 405 (6789), 914–919.

Krewulak, K.D., Shepherd, C.M. & Vogel, H.J. (2005) Molecular dynamics simulations of the periplasmic ferric-hydroxamate binding protein FhuD. *Biometals*. 18 (4), 375–386.

Krojer, T., Talon, R., Pearce, N., Collins, P., *et al.* (2017) The XChemExplorer graphical workflow tool for routine or large-scale protein-ligand structure determination. *Acta Crystallographica Section D: Structural Biology*. 73, 267–278.

Kumar, H., Kasho, V., Smirnova, I., Finer-Moore, J.S., *et al.* (2014) Structure of sugar-bound LacY. *Proceedings of the National Academy of Sciences*. 111 (5), 1784–1788.

- Kunji, E.R.S., Fang, G., Jeronimus-Stratingh, C.M., Bruins, A.P., *et al.* (1998) Reconstruction of the proteolytic pathway for use of β -casein by *Lactococcus lactis*. *Molecular Microbiology*. 27 (6), 1107–1118.
- Kuntz, I.D. (1992) Structure-based strategies for drug design and discovery. *Science*. 257 (5073), 1078–1082.
- Kurepina, N.E., Basyuk, E.I., Metlitskaya, A.Z., Zaitsev, D.A., *et al.* (1993) Cloning and mapping of the genetic determinants for microcin C51 production and immunity. *Molecular and General Genetics*. 241 (5–6), 700–706.
- Kuroda, T. & Tsuchiya, T. (2009) Multidrug efflux transporters in the MATE family. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. 1794 (5), 763–768.
- Lager, I., Fehr, M., Frommer, W.B. & Lalonde, S. (2003) Development of a fluorescent nanosensor for ribose. *FEBS Letters*. 553 (1–2), 85–89.
- Lagos, R., Baeza, M., Corsini, G., Hetz, C., *et al.* (2001) Structure, organization and characterization of the gene cluster involved in the production of microcin E492, a channel-forming bacteriocin. *Molecular Microbiology*. 42 (1), 229–243.
- Lanfermeijer, F.C., Detmers, F.J.M., Konings, W.N. & Poolman, B. (2000) On the binding mechanism of the peptide receptor of the oligopeptide transport system of *Lactococcus lactis*. *The EMBO Journal*. 19 (14), 3649 LP– 3656.
- Langer, G., Cohen, S.X., Lamzin, V.S. & Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature protocols*. 3 (7), 1171–1179.
- Laviña, M., Pugsley, A.P. & Moreno, F. (1986) Identification, mapping, cloning and characterization of a gene (sbmA) required for microcin B17 action on *Escherichia coli* K12. *Microbiology*. 132 (6), 1685–1693.
- Lazazzera, B.A. (2001) The intracellular function of extracellular signaling peptides. *Peptides*. 22 (10), 1519–1527.
- Lecher, J., Pittelkow, M., Zobel, S., Bursy, J., *et al.* (2009) The crystal structure of UehA in complex with ectoine-a comparison with other TRAP-T binding proteins. *Journal of Molecular Biology*. 389 (1), 58–73.
- Lee, M., Choi, Y., Burla, B., Kim, Y.-Y., *et al.* (2008) The ABC transporter AtABCB14 is a malate importer and modulates stomatal response to CO₂. *Nature Cell Biology*. 10, 1217.
- Lee, Y.H., Dorwart, M.R., Hazlett, K.R.O., Deka, R.K., *et al.* (2002) The crystal structure of Zn(II)-free *Treponema pallidum* Troa, a periplasmic metal-binding protein, reveals a closed conformation. *Journal of Bacteriology*. 184 (8), 2300–2304.
- Leimkühler, S., Wuebbens, M.M. & Rajagopalan, K. V (2001) Characterization of *Escherichia coli* MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor. *Journal of Biological Chemistry*. 276 (37), 34695–34701.

- Levdikov, V.M., Blagova, E. V, Brannigan, J.A., Wright, L., *et al.* (2005) The structure of the oligopeptide-binding protein, AppA, from *Bacillus subtilis* in complex with a nonapeptide. *Journal of Molecular Biology*. 345 (4), 879–892.
- Levy, S.B. (2002) Active efflux, a common mechanism for biocide and antibiotic resistance. *Journal of Applied Microbiology*. 92, 65S–71S.
- Levy, S.B. & Marshall, B. (2004) Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine*. 10 (S122).
- Lewinson, O. & Livnat-Levanon, N. (2017) Mechanism of action of ABC importers: conservation, divergence, and physiological adaptations. *Journal of Molecular Biology*. 429 (5), 606–619.
- Li, L., Park, E., Ling, J., Ingram, J., *et al.* (2016) Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature*. 531, 395.
- Li, X.-Z. & Nikaido, H. (2009) Efflux-mediated drug resistance in bacteria. *Drugs*. 69 (12), 1555–1623.
- Li, X., Ma, D. & Livermore, D.M. (1994) Role of efflux pump(s) in intrinsic resistance of *Pseudomonas aeruginosa*: active efflux as a contributing factor to 3-lactam resistance. *Antimicrobial Agents and Chemotherapy*. 38 (8), 1742–1752.
- Lim, S.P. & Nikaido, H. (2010) Kinetic parameters of efflux of penicillins by the multidrug efflux transporter AcrAB-TolC of *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*. 54 (5), 1800 LP – 1806.
- Lin, D.Y., Huang, S. & Chen, J. (2015) Crystal structures of a polypeptide processing and secretion transporter. *Nature*. 523, 425.
- Liu, H. & Naismith, J.H. (2009) A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expression and Purification*. 63 (2), 102–111.
- Locher, K.P. (2016) Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nature Structural and Molecular Biology*. 23 (6), 487–493.
- Locher, K.P., Lee, A.T. & Rees, D.C. (2002) The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science*. 296 (5570), 1091–1098.
- Locher, K.P., Rees, B., Koebnik, R., Mitschler, A., *et al.* (1998) Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell*. 95 (6), 771–778.
- Lomovskaya, O. & Bostian, K.A. (2006) Practical applications and feasibility of efflux pump inhibitors in the clinic — A vision for applied use. *Biochemical Pharmacology*. 71 (7), 910–918.
- Lomovskaya, O., Warren, M.S., Lee, A., Galazzo, J., *et al.* (2001) Identification and characterization of inhibitors of multidrug resistance efflux pumps in *Pseudomonas aeruginosa*:

novel agents for combination therapy. *Antimicrobial Agents and Chemotherapy*. 45 (1), 105 LP – 116.

Long, F., Su, C.-C., Zimmermann, M.T., Boyken, S.E., *et al.* (2010) Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport. *Nature*. 467 (7314), 484–488.

Ma, D., Cook, D.N., Alberti, M., Pon, N.G., *et al.* (1995) Genes *acrA* and *acrB* encode a stress-induced efflux system of *Escherichia coli*. *Molecular Microbiology*. 16 (1), 45–55.

Mahamoud, A., Chevalier, J., Alibert-Franco, S., Kern, W. V., *et al.* (2007) Antibiotic efflux pumps in Gram-negative bacteria: the inhibitor response strategy. *Journal of Antimicrobial Chemotherapy*. 59 (6), 1223–1229.

Majumdar, D.S., Smirnova, I., Kasho, V., Nir, E., *et al.* (2007) Single-molecule FRET reveals sugar-induced conformational dynamics in LacY. *Proceedings of the National Academy of Sciences*. 104 (31), 12640–12645.

Malléa, M., Chevalier, J., Eyraud, A. & Pagès, J.-M. (2002) Inhibitors of antibiotic efflux pump in resistant *Enterobacter aerogenes* strains. *Biochemical and Biophysical Research Communications*. 293 (5), 1370–1373.

Maragakis, L.L. (2010) Recognition and prevention of multidrug-resistant Gram-negative bacteria in the intensive care unit. *Critical Care Medicine*. 38, S345--S351.

Maruyama, Y., Itoh, T., Kaneko, A., Nishitani, Y., *et al.* (2015) Structure of a bacterial ABC transporter involved in the import of an acidic polysaccharide alginate. *Structure*. 23 (9), 1643–1654.

Mathavan, I., Zirah, S., Mehmood, S., Choudhury, H.G., *et al.* (2014) Structural basis for hijacking siderophore receptors by antimicrobial lasso peptides. *Nature Chemical Biology*. 10, 340.

Mazzariol, A., Tokue, Y., Kanegawa, T.M., Cornaglia, G., *et al.* (2000) High-level fluoroquinolone-resistant clinical isolates of *Escherichia coli* overproduce multidrug efflux protein AcrA. *Antimicrobial Agents and Chemotherapy*. 44 (12), 3441 LP – 3443.

McAleese, F., Petersen, P., Ruzin, A., Dunman, P.M., *et al.* (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.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., *et al.* (2007) Phaser crystallographic software. *Journal of Applied Crystallography*. 40 (4), 658–674.

McMurry, L., Petrucci, R.E. & Levy, S.B. (1980) Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in *Escherichia coli*. *Proceedings of the National Academy of Sciences*. 77 (7), 3974–3977.

Metlitskaya, A., Kazakov, T., Kommer, A., Pavlova, O., *et al.* (2006) Aspartyl-tRNA synthetase is the target of peptide nucleotide antibiotic Microcin C. *Journal of Biological Chemistry*. 281 (26), 18033–18042.

- Metlitskaya, A., Kazakov, T., Vondenhoff, G.H., Novikova, M., *et al.* (2009) Maturation of the translation inhibitor microcin C. *Journal of Bacteriology*. 191 (7), 2380–2387.
- Mi, W., Li, Y., Yoon, S.H., Ernst, R.K., *et al.* (2017) Structural basis of MsbA-mediated lipopolysaccharide transport. *Nature*. 549, 233.
- Mikolosko, J., Bobyk, K., Zgurskaya, H.I. & Ghosh, P. (2006) Conformational flexibility in the multidrug efflux system protein AcrA. *Structure*. 14 (3), 577–587.
- Mine, T., Morita, Y., Kataoka, A., Mizushima, T., *et al.* (1999) Expression in *Escherichia coli* of a new multidrug efflux pump, MexXY, from *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*. 43 (2), 415–417.
- Mondal, J., Friesner, R.A. & Berne, B.J. (2014) Role of desolvation in thermodynamics and kinetics of ligand binding to a kinase. *Journal of Chemical Theory and Computation*. 10 (12), 5696–5705.
- Morita, Y., Kodama, K., Shiota, S., Mine, T., *et al.* (1998) NorM, a putative multidrug efflux protein, of *Vibrio parahaemolyticus* and its homolog in *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*. 42 (7), 1778–1782.
- Morrison, E.A., DeKoster, G.T., Dutta, S., Vafabakhsh, R., *et al.* (2011) Antiparallel EmrE exports drugs by exchanging between asymmetric structures. *Nature*. 481, 45.
- Mowla, R., Wang, Y., Ma, S. & Venter, H. (2018) Kinetic analysis of the inhibition of the drug efflux protein AcrB using surface plasmon resonance. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 1860 (4), 878–886.
- Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T., *et al.* (2006) Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature*. 443 (7108), 173–179.
- Murakami, S., Nakashima, R., Yamashita, E. & Yamaguchi, A. (2002) Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature*. 419 (6907), 587–593.
- Murshudov, G.N., Dodson, E.J., Neil, T.K., Dodson, G.G., *et al.* (1994) The structural basis of sequence-independent peptide binding by OppA protein. *Science*. 264 (5165), 1578–1581.
- Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., *et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*. 67 (4), 355–367.
- Nagakubo, S., Nishino, K., Hirata, T. & Yamaguchi, A. (2002) The putative response regulator BaeR stimulates multidrug resistance of *Escherichia coli* via a novel multidrug exporter system, MdtABC. *Journal of Bacteriology*. 184 (15), 4161–4167.
- Nagano, K. & Nikaido, H. (2009) Kinetic behavior of the major multidrug efflux pump AcrB of *Escherichia coli*. *Proceedings of the National Academy of Sciences*. 106 (14), 5854–5858.

- Nakano, M.M. & Zuber, P. (1998) Anaerobic growth of a 'strict aerobe' (*Bacillus Subtilis*). *Annual Review of Microbiology*. 52 (1), 165–190.
- Nakashima, R., Sakurai, K., Yamasaki, S., Hayashi, K., *et al.* (2013) Structural basis for the inhibition of bacterial multidrug exporters. *Nature*. 500 (7460), 102–106.
- Nakashima, R., Sakurai, K., Yamasaki, S., Nishino, K., *et al.* (2011) Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket. *Nature*. 480 (7378), 565–569.
- Nakayama, K., Ishida, Y., Ohtsuka, M., Kawato, H., *et al.* (2003) MexAB-OprM-specific efflux pump inhibitors in *Pseudomonas aeruginosa*. *Bioorganic and Medicinal Chemistry Letters*. 13 (23), 4201–4204.
- Neiditch, M.B., Federle, M.J., Pompeani, A.J., Kelly, R.C., *et al.* (2006) Ligand-induced asymmetry in histidine sensor kinase complex regulates quorum sensing. *Cell*. 126 (6), 1095–1108.
- Nguyen, S.T., Kwasny, S.M., Ding, X., Cardinale, S.C., *et al.* (2015) Structure–activity relationships of a novel pyranopyridine series of Gram-negative bacterial efflux pump inhibitors. *Bioorganic and Medicinal Chemistry*. 23 (9), 2024–2034.
- Nies, D.H. (2003) Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiology Reviews*. 27 (2–3), 313–339.
- Nikaido, H. (1998) Antibiotic resistance caused by gram-negative multidrug efflux pumps. *Clinical Infectious Diseases*. 27 (Suppl 1), S32–41.
- Nikaido, H. (1988) Bacterial resistance to antibiotics as a function of outer membrane permeability. *Journal of Antimicrobial Chemotherapy*. 17–22.
- Nikaido, H. (1996) Multidrug efflux pumps of gram-negative bacteria. *Journal of Bacteriology*. 178 (20), 5853–5859.
- Nikaido, H. (1989) Outer membrane barrier as a mechanism of antimicrobial resistance. *Antimicrobial Agents and Chemotherapy*. 33 (11), 1831.
- Nikaido, H. (1994) Porins and specific diffusion channels in bacterial outer membranes. *Journal of Biological Chemistry*. 269 (6), 3905–3908.
- Nikaido, H. (2018) RND transporters in the living world. *Research in Microbiology*. 169 (7–8), 363–371.
- Nikaido, H. & Pagès, J.-M. (2012) Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria. *FEMS Microbiology Reviews*. 36 (2), 340–363.
- Nikaido, H. & Vaara, M. (1985) Molecular basis of bacterial outer membrane permeability. *Microbiology and Molecular Biology Reviews*. 67 (1), 593–656.
- Nishino, K. & Yamaguchi, A. (2001) Analysis of a complete library of putative drug transporter genes in *Escherichia coli*. *Journal of Bacteriology*. 183 (20), 5803–5812.

Novikova, M., Kazakov, T., Vondenhoff, G.H., Semenova, E., *et al.* (2010) MccE provides resistance to protein synthesis inhibitor microcin C by acetylating the processed form of the antibiotic. *Journal of Biological Chemistry*. M109.

Novikova, M., Metlitskaya, A., Datsenko, K., Kazakov, T., *et al.* (2007) The *Escherichia coli* Yej transporter is required for the uptake of translation inhibitor microcin C. *Journal of Bacteriology*. 189 (22), 8361–8365.

O'Hara, P.J., Sheppard, P.O., Thøgersen, H., Venezia, D., *et al.* (1993) The ligand-binding domain in metabotropic glutamate receptors is related to bacterial periplasmic binding proteins. *Neuron*. 11 (1), 41–52.

Oh, B.-H., Kang, C.-H., De Bondt, H., Kim, S.-H., *et al.* (1994) The bacterial periplasmic histidine-binding protein. *Journal of Biological Chemistry*. 269 (6), 4135–4143.

Okusu, H., Ma, D. & Nikaido, H. (1996) AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. *Journal of Bacteriology*. 178 (1), 306–308.

Oldham, M.L. & Chen, J. (2011) Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science*. 332 (6034), 1202–1205.

Oldham, M.L., Khare, D., Quijcho, F.A., Davidson, A.L., *et al.* (2007) Crystal structure of a catalytic intermediate of the maltose transporter. *Nature*. 450 (7169), 515–521.

Omote, H., Hiasa, M., Matsumoto, T., Otsuka, M., *et al.* (2006) The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends in Pharmacological Sciences*. 27 (11), 587–593.

Opperman, T.J., Kwasny, S.M., Kim, H.S., Nguyen, S.T., *et al.* (2014) Characterization of a novel pyranopyridine inhibitor of the AcrAB efflux pump of *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*. 58 (2), 722–733.

Origin (2002) Origin, Version 7. OriginLab Corporation, Northampton, MA, USA.

Otsuka, M., Matsumoto, T., Morimoto, R., Arioka, S., *et al.* (2005) A human transporter protein that mediates the final excretion step for toxic organic cations. *Proceedings of the National Academy of Sciences*. 102 (50), 17923–17928.

Otto, M. (2010) Looking toward basic science for potential drug discovery targets against community-associated MRSA. *Medicinal research reviews*. 30 (1), 1–22.

Pagès, J.-M. & Amaral, L. (2009) Mechanisms of drug efflux and strategies to combat them: Challenging the efflux pump of Gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. 1794 (5), 826–833.

Pagès, J.-M., Masi, M. & Barbe, J. (2005) Inhibitors of efflux pumps in Gram-negative bacteria. *Trends in Molecular Medicine*. 11 (8), 382–389.

- Pannek, S., Higgins, P.G., Steinke, P., Jonas, D., *et al.* (2006) Multidrug efflux inhibition in *Acinetobacter baumannii*: comparison between 1-(1-naphthylmethyl)-piperazine and phenyl-arginine- β -naphthylamide. *Journal of Antimicrobial Chemotherapy*. 57 (5), 970–974.
- Pao, S.S., Paulsen, I.T. & Saier, M.H. (1998) Major facilitator superfamily. *Microbiology and Molecular Biology Reviews*. 62 (1), 1–34.
- Patzner, S.I., Baquero, M.R., Bravo, D., Moreno, F., *et al.* (2003) The colicin G, H and X determinants encode microcins M and H47, which might utilize the catecholate siderophore receptors FepA, Cir, Fiu and IroN. *Microbiology*. 149 (9), 2557–2570.
- Paulsen, I.T., Nguyen, L., Sliwinski, M.K., Rabus, R., *et al.* (2000) Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *Journal of Molecular Biology*. 301 (1), 75–100.
- Paulsen, I.T., Park, J.H., Choi, P.S. & Jr., M.H.S. (1997) A family of Gram-negative bacterial outer membrane factors that function in the export of proteins, carbohydrates, drugs and heavy metals from Gram-negative bacteria. *FEMS Microbiology Letters*. 156, 1–8.
- Pautsch, A. & Schulz, G.E. (1998) Structure of the outer membrane protein A transmembrane domain. *Nature Structural and Molecular Biology*. 5 (11), 1013.
- Pearce, N.M., Krojer, T., Bradley, A.R., Collins, P., *et al.* (2017) A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nature Communications*. 8, 15123.
- Perez, C., Gerber, S., Boilevin, J., Bucher, M., *et al.* (2015) Structure and mechanism of an active lipid-linked oligosaccharide flippase. *Nature*. 524, 433.
- Piddock, L.J. V (2006) Multidrug-resistance efflux pumps - Not just for resistance. *Nature Reviews Microbiology*. 4 (8), 629–636.
- Pinkett, H.W., Lee, A.T., Lum, P., Locher, K.P., *et al.* (2007) An inward-facing conformation of a putative metal-chelate-type ABC transporter. *Science*. 315 (5810), 373–377.
- Plesiat, P. & Nikaido, H. (1992) Outer membranes of Gram-negative bacteria are permeable to steroid probes. *Molecular Microbiology*. 6 (10), 1323–1333.
- Ploeg, J.R. van der, Iwanicka-Nowicka, R., Bykowski, T., Hryniewicz, M.M., *et al.* (1999) The *Escherichia coli* ssuEADCB gene cluster is required for the utilization of sulfur from aliphatic sulfonates and is regulated by the transcriptional activator Cbl. *Journal of Biological Chemistry*. 274 (41), 29358–29365.
- Ploeg, J.R. van der, Weiss, M.A., Saller, E., Nashimoto, H., *et al.* (1996) Identification of sulfate starvation-regulated genes in *Escherichia coli*: a gene cluster involved in the utilization of taurine as a sulfur source. *Journal of Bacteriology*. 178 (18), 5438–5446.
- Pons, A. M., Delalande, F., Duarte, M., Benoit, S., *et al.* (2004) Genetic analysis and complete primary structure of microcin L. *Antimicrobial Agents and Chemotherapy*. 48 (2), 505 LP – 513.

- Poole, K. (2005) Efflux-mediated antimicrobial resistance. *Journal of Antimicrobial Chemotherapy*. 56 (1), 20–51.
- Poole, K., Gotoh, N., Tsujimoto, H., Zhao, Q., *et al.* (1996) Overexpression of the mexC-mexD-oprJ efflux operon in nfxB-type multidrug-resistant strains of *Pseudomonas aeruginosa*. *Molecular Microbiology*. 21 (4), 713–725.
- Poole, K., Krebes, K., McNally, C. & Neshat, S. (1993) Multiple antibiotic resistance in *Pseudomonas aeruginosa*: evidence for involvement of an efflux operon. *Journal of Bacteriology*. 175 (22), 7363–7372.
- Pos, K.M. (2009) Drug transport mechanism of the AcrB efflux pump. *Biochimica et Biophysica Acta (BBA)*. 1794 (5), 782–793.
- Pos, K.M. & Diederichs, K. (2002) Purification, crystallization and preliminary diffraction studies of AcrB, an inner-membrane multi-drug efflux protein. *Acta Crystallographica Section D*. 58 (10 Part 2), 1865–1867.
- Postgate, J.R. (1976) Death in macrobes and microbes. *Symposium of the Society for General Microbiology*. 1976 pp. 1–18.
- Pugsley, A.P., Moreno, F. & De Lorenzo, V. (1986) Microcin-E492-insensitive Mutants of *Escherichia coli* K12. *Microbiology*. 132 (12), 3253–3259.
- Quiocho, F.A., Spurlino, J.C. & Rodseth, L.E. (1997) Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure*. 5 (8), 997–1015.
- Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., *et al.* (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences*. 106 (6), 1760–1765.
- Ramachandran, G.N., Ramakrishnan, C. & Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*. 7 (1), 95–99.
- Reader, J.S., Ordoukhanian, P.T., Kim, J.-G., de Crécy-Lagard, V., *et al.* (2005) Major biocontrol of plant tumors targets tRNA synthetase. *Science*. 309 (5740), 1533.
- Rebuffat, S. (2013) Chapter 20 - Microcins. *Handbook of Biologically Active Peptides*. Academic Press. 129–137.
- Reddy, K.V.R., Yedery, R.D. & Aranha, C. (2004) Antimicrobial peptides: premises and promises. *International Journal of Antimicrobial Agents*. 24 (6), 536–547.
- Reddy, V.S., Tsu, B. V, Ahmed, M.S., Li, C., *et al.* (2016) The Transporter Classification Database (TCDB): recent advances.
- Rees, D.C., Congreve Miles, Murray, C.W. & Carr, R. (2004) Fragment-based lead discovery. *Nature Reviews Drug Discovery*. 3, 660.

Rees, D.C., Johnson, E. & Lewinson, O. (2009) ABC transporters : The power to change. *Nature Review Molecular Cell Biology*. 10 (3), 218–227.

Reeve, C.A., Bockman, A.T. & Matin, A. (1984) Role of protein degradation in the survival of carbon-starved *Escherichia coli* and *Salmonella typhimurium*. *Journal of Bacteriology*. 157 (3), 758–763.

Rice, A.J., Park, A. & Pinkett, H.W. (2014) Diversity in ABC transporters: Type I, II and III importers. *Critical Reviews in Biochemistry and Molecular Biology*. 49 (5), 426–437.

Rice, L.B. (2009) The clinical consequences of antimicrobial resistance. *Current Opinion in Microbiology*. 12 (5), 476–481.

Robert, X. & Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*. 42 (W1), 320–324.

Rodionov, D.A., Hebbeln, P., Eudes, A., ter Beek, J., *et al.* (2009) A novel class of modular transporters for vitamins in prokaryotes. *Journal of Bacteriology*. 191 (1), 42–51.

Rodionov, D.A., Hebbeln, P., Gelfand, M.S. & Eitinger, T. (2006) Comparative and functional genomic analysis of prokaryotic nickel and cobalt uptake transporters: evidence for a novel group of ATP-binding cassette transporters. *Journal of Bacteriology*. 188 (1), 317–327.

Rosenberg, E.Y., Ma, D. & Nikaido, H. (2000) AcrD of *Escherichia coli* Is an aminoglycoside efflux pump. *Journal of Bacteriology*. 182 (6), 1754–1756.

Rosengren, K.J., Clark, R.J., Daly, N.L., Göransson, U. *et al.* (2003) Microcin J25 has a threaded sidechain-to-backbone ring structure and not a head-to-tail cyclized backbone. *Journal of the American Chemical Society*. 125 (41), 12464–12474.

Roszak, D.B. & Colwell, R.R. (1987) Survival strategies of bacteria in the natural environment. *Microbiological Reviews*. 51 (3), 365–379.

Roush, R.F., Nolan, E.M., Löhr, F. & Walsh, C.T. (2008) Maturation of an *Escherichia coli* ribosomal peptide antibiotic by ATP-consuming N- P bond formation in microcin C7. *Journal of the American Chemical Society*. 130 (11), 3603–3609.

Rubin, S.M., Lee, S.-Y., Ruiz, E.J., Pines, A., *et al.* (2002) Detection and characterization of xenon-binding sites in proteins by ¹²⁹Xe NMR spectroscopy. *Journal of Molecular Biology*. 322 (2), 425–440.

Ruiz, N., Kahne, D. & Silhavy, T.J. (2009) Transport of lipopolysaccharide across the cell envelope: the long road of discovery. *Nature Reviews Microbiology*. 7, 677.

Sack, J.S., Saper, M.A. & Quijcho, F.A. (1989) Periplasmic binding protein structure and function: Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. *Journal of Molecular Biology*. 206 (1), 171–191.

Sack, J.S., Trakhanov, S.D., Tsigannik, I.H. & Quijcho, F.A. (1989) Structure of the L-leucine-binding protein refined at 2.4 Å resolution and comparison with the Leu/Ile/Val-binding protein structure. *Journal of Molecular Biology*. 206 (1), 193–207.

Saint, N., Lou, K.-L., Widmer, C., Luckey, M., *et al.* (1996) Structural and functional characterization of OmpF porin mutants selected for larger pore size II. Functional characterization. *Journal of Biological Chemistry*. 271 (34), 20676–20680.

Salomón, R.A. & Farias, R.N. (1993) The FhuA protein is involved in microcin 25 uptake. *Journal of Bacteriology*. 175 (23), 7741–7742.

Salomón, R.A. & Farias, R.N. (1995) The peptide antibiotic microcin 25 is imported through the TonB pathway and the SbmA protein. *Journal of Bacteriology*. 177 (11), 3323–3325.

Saurin, W., Hofnung, M. & Dassa, E. (1999) Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *Journal of Molecular Evolution*. 48 (1), 22–41.

Saurin, W., Köster, W. & Dassa, E. (1994) Bacterial binding protein-dependent permeases: characterization of distinctive signatures for functionally related integral cytoplasmic membrane proteins. *Molecular Microbiology*. 12 (6), 993–1004.

Scheepers, G.H., Lycklama a Nijeholt, J.A. & Poolman, B. (2016) An updated structural classification of substrate-binding proteins. *FEBS Letters*. 590 (23), 4393–4401.

Schirmer, T. (1998) General and specific porins from bacterial outer membranes. *Journal of Structural Biology*. 121 (2), 101–109.

Schramm, E., Mende, J., Braun, V. & Kamp, R.M. (1987) Nucleotide sequence of the colicin B activity gene *cba*: consensus pentapeptide among TonB-dependent colicins and receptors. *Journal of Bacteriology*. 169 (7), 3350–3357.

Schumacher, M.A., Choi, K.Y., Zalkin, H. & Brennan, R.G. (1994) Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. *Science*. 266 (5186), 763 LP – 770.

Seeger, M.A., Schiefner, A., Eicher, T., Verrey, F., *et al.* (2006) Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science*. 313 (5791), 1295–1298.

Sennhauser, G., Amstutz, P., Briand, C., Storchenegger, O., *et al.* (2007) Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS Biology*. 5 (1), 0106–0113.

Severinov, K., Semenova, E. & Kazakov, T. (2011) Class I microcins: their structures, activities, and mechanisms of resistance. *Prokaryotic Antimicrobial Peptides*. Springer. 289–308.

Shaw, K.J., Rather, P.N., Hare, R.S. & Miller, G.H. (1993) Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. *Microbiological Reviews*. 57 (1), 138–163.

- Sharma, O., Zakharov, S.D., Zhalnina, M. V, Yamashita, E., *et al.* (2013) Chapter 17 - Colicins. *Handbook of Biologically Active Peptides*. Academic Press. 93–100.
- Shiomi, N., Fukuda, H., Fukuda, Y., Murata, K., *et al.* (1991) Nucleotide sequence and characterization of a gene conferring resistance to ethionine in yeast *Saccharomyces cerevisiae*. *Journal of Fermentation and Bioengineering*. 71 (4), 211–215.
- Sjuts, H., Vargiu, A. V, Kwasny, S.M., Nguyen, S.T., *et al.* (2016) Molecular basis for inhibition of AcrB multidrug efflux pump by novel and powerful pyranopyridine derivatives. *Proceedings of the National Academy of Sciences*. 113 (13), 3509–3514.
- Smirnova, I., Kasho, V., Choe, J.-Y., Altenbach, C., *et al.* (2007) Sugar binding induces an outward facing conformation of LacY. *Proceedings of the National Academy of Sciences*. 104 (42), 16504–16509.
- Smirnova, I., Kasho, V., Sugihara, J. & Kaback, H.R. (2009) Probing of the rates of alternating access in LacY with Trp fluorescence. *Proceedings of the National Academy of Sciences*. 106 (51), 21561–21566.
- Solbiati, J.O., Ciaccio, M., Farías, R.N., González-Pastor, J.E., *et al.* (1999) Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25. *Journal of Bacteriology*. 181 (8), 2659 LP – 2662.
- Solbiati, J.O., Ciaccio, M., Farías, R.N. & Salomón, R.A. (1996) Genetic analysis of plasmid determinants for microcin J25 production and immunity. *Journal of Bacteriology*. 178 (12), 3661 LP – 3663.
- Soriano, E. V, Rajashankar, K.R., Hanes, J.W., Bale, S., *et al.* (2008) Structural similarities between thiamin-binding protein and thiaminase-I suggest a common ancestor. *Biochemistry*. 47 (5), 1346–1357.
- Spratt, B.G. (1994) Resistance to antibiotics mediated by target alterations. *Science*. 264 (5157), 388–393.
- Stergiopoulos, I., Zwiers, L.-H. & De Waard, M.A. (2003) The ABC transporter MgAtr4 is a virulence factor of *Mycosphaerella graminicola* that affects colonization of substomatal cavities in wheat leaves. *Molecular Plant-Microbe Interactions*. 16 (8), 689–698.
- Stillwell, W. (2013) *An introduction to biological membranes: from bilayers to rafts*. Newnes.
- Sulavik, M.C., Houseweart, C., Cramer, C., Jiwani, N., *et al.* (2001) Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug efflux pump genes. *Antimicrobial Agents and Chemotherapy*. 45 (4), 1126–1136.
- Sun, J., Bankston, J.R., Payandeh, J., Hinds, T.R., *et al.* (2014) Crystal structure of the plant dual-affinity nitrate transporter NRT1.1. *Nature*. 507, 73.
- Sun, Y., Olson, R., Horning, M., Armstrong, N., *et al.* (2002) Mechanism of glutamate receptor desensitization. *Nature*. 417 (6886), 245–253.

- Symmons, M.F., Bokma, E., Koronakis, E., Hughes, C., *et al.* (2009) The assembled structure of a complete tripartite bacterial multidrug efflux pump. *Proceedings of the National Academy of Sciences*. 106 (17), 7173–7178.
- Tacar, O., Sriamornsak, P. & Dass, C.R. (2012) Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. *Journal of Pharmacy and Pharmacology*. 65 (2), 157–170.
- Takatsuka, Y., Chen, C. & Nikaido, H. (2010) Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of *Escherichia coli*. *Proceedings of the National Academy of Sciences*. 107 (15), 6559–6565.
- Tame, Murshudov, G.N., Dodson, E.J., Neil, T.K., *et al.* (1994) The structural basis of sequence-independent peptide binding by OppA protein. *Science*. 264 (5165), 1578 LP – 1581.
- Tame, J.R.H., Sleigh, S.H., Wilkinson, A.J. & Ladbury, J.E. (1996) The role of water in sequence-independent ligand binding by an oligopeptide transporter protein. *Nature Structural Biology*. 3 (12), 998–1001.
- Tanaka, Y., Hipolito, C.J., Maturana, A.D., Ito, K., *et al.* (2013) Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. *Nature*. 496 (7444), 247–251.
- Taylor, D.E., Garner, R.S. & Allan, B.J. (1983) Characterization of tetracycline resistance plasmids from *Campylobacter jejuni* and *Campylobacter coli*. *Antimicrobial Agents and Chemotherapy*. 24 (6), 930–935.
- Taylor, D.E., Grandis, S.A. DE, Karmali, M.A. & Fleming, P.C. (1981) Transmissible plasmids from *Campylobacter jejuni*. *Antimicrobial Agents and Chemotherapy*. 19, 831–835.
- Taylor, S. V, Kelleher, N.L., Kinsland, C., Chiu, H.-J., *et al.* (1998) Thiamin Biosynthesis in *Escherichia coli*. *Journal of Biological Chemistry*. 273 (26), 16555–16560.
- Tenover, F.C., Williams, S., Gordon, K.P., Nolan, C., *et al.* (1985) Survey of plasmids and resistance factors in *Campylobacter jejuni* and *Campylobacter coli*. *Antimicrobial Agents and Chemotherapy*. 27, 37–41.
- Thanassi, D.G., Cheng, L.W. & Nikaido, H. (1997) Active efflux of bile salts by *Escherichia coli*. *Journal of Bacteriology*. 179 (8), 2512–2518.
- Thomas, C. & Tampé, R. (2018) Multifaceted structures and mechanisms of ABC transport systems in health and disease. *Current Opinion in Structural Biology*. 51, 116–128.
- Tikhonov, A., Kazakov, T., Semenova, E., Serebryakova, M., *et al.* (2010) The mechanism of microcin C resistance provided by the MccF peptidase. *Journal of Biological Chemistry*. 285 (49), 37944–37952.
- Tikhonova, E.B. & Zgurskaya, H.I. (2004) AcrA, AcrB, and TolC of *Escherichia coli* form a stable intermembrane multidrug efflux complex. *Journal of Biological Chemistry*. 279 (31), 32116–32124.

- Tófoli De Araújo, F., Bolanos-Garcia, V.M., Pereira, C.T., Sanches, M., *et al.* (2013) Structural and physiological analyses of the alkanesulphonate-binding protein (SsuA) of the citrus pathogen *Xanthomonas citri*. *PLoS ONE*. 8 (11).
- Törnroth-Horsefield, S., Gourdon, P., Horsefield, R., Brive, L., *et al.* (2007) Crystal structure of AcrB in complex with a single transmembrane subunit reveals another twist. *Structure*. 15 (12), 1663–1673.
- Trujillo, M., Rodríguez, E. & Laviña, M. (2001) ATP synthase is necessary for microcin H47 antibiotic action. *Antimicrobial Agents and Chemotherapy*. 45 (11), 3128–3131.
- Tseng, T.-T.T., Gratwick, K.S., Kollman, J., Park, D., *et al.* (1999) The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *Journal of Molecular Microbiology and Biotechnology*. 1 (1), 107–125.
- Tynkkynen, S., Buist, G., Kunji, E., Kok, J., *et al.* (1993) Genetic and biochemical characterization of the oligopeptide transport system of *Lactococcus lactis*. *Journal of Bacteriology*. 175 (23), 7523–7532.
- van Veen, H.W., Callaghan, R., Soceneantu, L., Sardini, A., *et al.* (1998) A bacterial antibiotic-resistance gene that complements the human multidrug-resistance P-glycoprotein gene. *Nature*. 391, 291.
- Vetting, M.W., De Carvalho, L.P.S., Roderick, S.L. & Blanchard, J.S. (2005) A novel dimeric structure of the RimL N α -acetyltransferase from *Salmonella typhimurium*. *Journal of Biological Chemistry*.
- Vizán, J.L., Hernández-Chico, C., del Castillo, I. & Moreno, F. (1991) The peptide antibiotic microcin B17 induces double-strand cleavage of DNA mediated by *E. coli* DNA gyrase. *The EMBO Journal*. 10 (2), 467–476.
- Vondenhoff, G.H.M., Blanchaert, B., Geboers, S., Kazakov, T., *et al.* (2011) Characterization of peptide chain length and constituency requirements for YejABEF-mediated uptake of microcin C analogues. *Journal of Bacteriology*. 193 (14), 3618–3623.
- Waitz, J.A. (1990) Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically.
- Wang, T., Fu, G., Pan, X., Wu, J., *et al.* (2013) Structure of a bacterial energy-coupling factor transporter. *Nature*. 497, 272.
- Wang, Z., Fan, G., Hryc, C.F., Blaza, J.N., *et al.* (2017) An allosteric transport mechanism for the AcrAB-TolC multidrug efflux pump. *eLife*. 6, 1–19.
- Ward, A., Reyes, C.L., Yu, J., Roth, C.B., *et al.* (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proceedings of the National Academy of Sciences*. 104 (48), 19005–19010.

- White, D.G., Goldman, J.D., Demple, B. & Levy, S.B. (1997) Role of the *acrAB* locus in organic solvent tolerance mediated by expression of *marA*, *soxS*, or *robA* in *Escherichia coli*. *Journal of Bacteriology*. 179 (19), 6122–6126.
- Winnen, B., Hvorup, R.N. & Saier, M.H. (2003) The tripartite tricarboxylate transporter (TTT) family. *Research in Microbiology*. 154 (7), 457–465.
- Winter, G. (2010) Xia2: An expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*. 43 (1), 186–190.
- Woo, J.-S., Zeltina, A., Goetz, B.A. & Locher, K.P. (2012) X-ray structure of the *Yersinia pestis* heme transporter HmuUV. *Nature Structural and Molecular Biology*. 19, 1310.
- Xu, K., Zhang, M., Zhao, Q., Yu, F., *et al.* (2013) Crystal structure of a folate energy-coupling factor transporter from *Lactobacillus brevis*. *Nature*. 497 (7448), 268.
- Yamaguchi, A., Nakashima, R. & Sakurai, K. (2015) Structural basis of RND-type multidrug exporters. *Frontiers in Microbiology*. 6 (APR), 1–19.
- Yang, C.C. & Konisky, J. (1984) Colicin V-treated *Escherichia coli* does not generate membrane potential. *Journal of Bacteriology*. 158 (2), 757 LP – 759.
- Yao, N., Trakhanov, S. & Quijcho, F.A. (1994) Refined 1.89-Å structure of the histidine-binding protein complexed with histidine and its relationship with many other active transport/chemosensory proteins. *Biochemistry*. 33 (16), 4769–4779.
- Yao, XQ., Kimura, N., Murakami, S. & Takada, S. (2013) Drug uptake pathways of multidrug transporter AcrB studied by molecular simulations and site-directed mutagenesis experiments. *Journal of the American Chemical Society*. 135 (20), 7474–7485.
- Yeagle, P.L. (2016) *The Membranes of Cells*. Academic Press.
- Yerushalmi, H., Lebendiker, M. & Schuldiner, S. (1995) EmrE, an *Escherichia coli* 12-kDa multidrug transporter, exchanges toxic cations and H⁺ and is soluble in organic solvents. *Journal of Biological Chemistry*. 270 (12), 6856–6863.
- Yin, Y., He, X., Szewczyk, P., Nguyen, T., *et al.* (2006) Structure of the multidrug transporter EmrD from *Escherichia coli*. *Science*. 312 (5774), 741 LP – 744.
- Yoshida, K., Nakayama, K., Ohtsuka, M., Kuru, N., *et al.* (2007) MexAB-OprM specific efflux pump inhibitors in *Pseudomonas aeruginosa*. Part 7: Highly soluble and in vivo active quaternary ammonium analogue D13-9001, a potential preclinical candidate. *Bioorganic and Medicinal Chemistry*. 15 (22), 7087–7097.
- Yoshihara, E. & Nakae, T. (1989) Identification of porins in the outer membrane of *Pseudomonas aeruginosa* that form small diffusion pores. *Journal of Biological Chemistry*. 264 (11), 6297–6301.

- Yu, E.W., Aires, J.R., McDermott, G. & Nikaido, H. (2005) A periplasmic drug-binding site of the AcrB multidrug efflux pump: A crystallographic and site-directed mutagenesis study. *Journal of Bacteriology*. 187 (19), 6804–6815.
- Yu, E.W., McDermott, G., Zgurskaya, H.I., Nikaido, H., *et al.* (2003) Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump. *Science*. 300 (5621), 976–980.
- Yu, J., Ge, J., Heuveling, J., Schneider, E., *et al.* (2015) Structural basis for substrate specificity of an amino acid ABC transporter. *Proceedings of the National Academy of Sciences*. 112 (16), 5243–5248.
- Zechini, B. & Versace, I. (2009) Inhibitors of multidrug resistant efflux systems in bacteria. *Recent Patents on Anti-Infective Drug Discovery*. 4 (1), 37–50.
- Zgurskaya, H.I. (2009) Multicomponent drug efflux complexes: architecture and mechanism of assembly. *Future Microbiology*. 4, 919–932.
- Zhang, P. (2013) Structure and mechanism of energy-coupling factor transporters. *Trends in Microbiology*. 21 (12), 652–659.
- Zhou, Y., Guan, L., Freites, J.A. & Kaback, H.R. (2008) Opening and closing of the periplasmic gate in lactose permease. *Proceedings of the National Academy of Sciences*. 105 (10), 3774–3778.
- Zou, P. & Mchaourab, H.S. (2009) Alternating access of the putative substrate-binding chamber in the ABC transporter MsbA. *Journal of Molecular Biology*. 393 (3), 574–585.