UNIVERISTY OF CALIFORNIA, SAN DIEGO

# Structural Studies of Multidrug Transporter Proteins

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Biology

by

Paul Ralph Szewczyk

Committee in charge:

Professor Milton Saier, Chair
Professor Geoffrey Chang
Professor Amro Hamdoun
Professor Randy Hampton
Professor Julian Schroeder

2014

The Dissertation of Paul Szewczyk is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2014

# Table of Contents


Signature Page ..... ii
Table of Contents ..... iv
List of Abbreviations ..... vi
List of Figures ..... viii
List of Tables ..... ix
Acknowledgements ..... x
Vita ..... xi
Abstract of the Dissertation ..... xii
Introduction ..... 1
Chapter 1: Structural Determination of NorM from *Vibrio Cholerea* ..... 4
1.1: Introduction ..... 4
1.2: The Structure of NorM_VC ..... 5
1.3: The Cation Binding Site of NorM_VC ..... 7
1.4: The Transport Cycle of NorM_VC ..... 8
1.5: Cloning, Mutagenesis, Protein Expression, and Purification ..... 9
1.6: Crystallization of NorM_VC ..... 11
1.7: Derivatization and Cryoprotection ..... 12
1.8: Data Acquisition and Structure Determination ..... 12
1.9: Model Building and Figures ..... 15
1.10: NorM_VC Substrate Binding Activity by Fluorescence Polarization ..... 15
1.11: NorM_VC Transport Activity by Ethidium Bromide (EtBr) Accumulation ..... 17
Chapter 2: Conformational Flexibility and a Novel Epitope of Mouse P-gp ..... 57
2.1: Introduction ..... 57
2.2: Structure Determination ..... 60
2.3: Validation of Structure ..... 61
2.4: Nanobody-Bound Structure of P-gp ..... 61
2.5: Inhibition of P-gp ATPase by Nb592 ..... 63
2.6: Context of the Novel Structures of P-gp ..... 64
2.7: P-gp Protein Expression and Purification ..... 68
2.8: Elicitation of Nanobodies Against P-gp ..... 69
2.9: Nanobody Expression and Purification ..... 70
2.10: Formation of P-gp-Nanobody Complex ..... 71
2.11: ATPase Inhibition ..... 71
2.12: Crystallization, Data Collection, And Structure Determination ..... 72
Chapter 3: Snapshots of Ligand Entry, Malleable Binding, and Induced Helical Movement in P-glycoprotein ..... 89
3.1: Introduction ..... 89
3.2: Synthesis of QZ-Homotrimers ..... 90
3.3: Measurement of ATPase Activity ..... 92
3.4: Calcein-AM Transport Assay ..... 92

3.5: Sensitization assay ..... 93
3.6: Expression, Purification Reductive Methylation, and Crystallization of P-gp ..... 93
3.7: X-ray Data Collection, Structure Determination and Refinement of P-gp Cyclopeptide Cocrystal Structures ..... 94
3.8: Rational Engineering and Functional Characterization Of Cyclopeptide P-gp Ligands ..... 95
3.9: Structures of P-gp in Complex with Cylcopeptide Ligands ..... 97
3.10: Comparison of Cylcopeptide Ligands in Complex with Pgp ..... 97
3.11: Movement of TM4 Upon Ligand Binding ..... 98
3.12: Ligand Binding Site at the Membrane Interface ..... 99
3.13 Context of the Presented Structures ..... 100
References: ..... 114

## List of Abbreviations

08ID-1: Beamline 08ID-1 of the Canadian Light Source
23-ID-B: Beamline 23-ID-B of the Advanced Photon Source
23-ID-D: Beamline 23-ID-D of the Advanced Photon Source
ABCB1: P-glycoprotein
ABCB1a: Mouse P-glycooprotein
ABC: ATP-binding cassette
ALS: Advanced Light Source
APS: Advanced Photon Source
ATP: Adenosine triphosphate
β-DDM: n-dodecyl−β-D-maltopyranoside
BDM: Bakers dimercurial
β-NG: n-nonyl-β-glucopyranoside)
BL5.0.1: Beamline 5.0.1 of the Advanced Light Source
BL 11-1: Beamline 11-1 of the Stanford Synchrotron Radiation Laboratory
Calcein-AM: Calcein acetoxymethyl ester
CDRs: Complementarity Determining Regions
CLS: Canadian Light Source
$Cs^+$: Cesium
D: Daltons
Desalting-NG: Desalting buffer with the addition of 0.23% β-NG
EH: Elbow helix
EtBr: Ethidium Bromide
ECL: Extracellular loop
FA-231: 3$\alpha$-hydroxy-7$\alpha$,12$\alpha$-bis[($\beta$-D-maltopyranosyl)ethyloxy] cholane
FDA: United States Food and Drug Administration
FDPP: Pentafluorophenol diphenylphosphinate
$IC_{50:}$ Half-maximal inhibition/Half maximal inhibitory concentration
IH: Intracellular helix
$K_i$: Inhibition constant
MAD: Multiple anomalous dispersion
MATE: Multidrug and toxic compound extrusion
MDR: Multidrug resistance
MFS: Major facilitator
mlf: Maximum likelihood target with amplitudes
MR: Molecular replacement
Nbs: Nanobodies
NBD: Nucleotide-binding domain
NCS: Non-crystallographic symmetry
NorM_VC: NorM *Vibrio cholerea*
NorM_VP: NorM from *Vibrio parahaemolyticus*
NBD: Nucleotide-binding domain
$OD_{600}$: Optical density at $\lambda$=600nm
PEG250DME: Polyethylene glycol 250 dimethyl ether

P-gp: P-glycoprotein
Pi: Inorganic phosphate
SA: Simulated annealing
QZ-Ala: Cyclopeptide bearing three R-group alanines
QZ-Leu: Cyclopeptide bearing three R-group leucines
QZ-Phe: Cyclopeptide bearing three R-group phenylalanines
QZ-Val: Cyclopeptide bearing three R-group valines
$Rb^+$: Rubidium
RG6: Rhodamine G6
RND: Resistance modulation and cell division
SAR: Structure-activity relationship study
SEC: Size-exclusion Chromatography
SMR: Small multidrug resistance
SSRL: Stanford Synchrotron Radiation Laboratory
TM: Transmembrane/Transmembrane helix
TPP: Tetraphenylphosphonium
TMD: Transmembrane domain
VER: Verapamil
Vi: Sodium orthovanadate

## List of Figures

Figure 1.1: Amino Acid Sequence Alignment of NorM_VC ..... 19
Figure 1.2: Effect of Recombinant NorM_VC on EtBr Uptake ..... 20
Figure 1.3: Substrate Binding Activity of NorM_VC ..... 21
Figure 1.4: Stereo Views of NorM_VC ..... 22
Figure 1.5: Experimental Electron Density Map of NorM_VC ..... 23
Figure 1.6: Stereo Views of Fo-Fc Simulated Annealing Omit Density for TM Helices of NorM_VC ..... 24
Figure 1.7: Stereo Views of Fo-Fc Simulated Annealing Omit Density NorM_VC with Multiple Density Maps Superimposed ..... 32
Figure 1.8: Validation of the NorM_VC Model ..... 35
Figure 1.9: Topology Verification of NorM_VC Structure ..... 37
Figure 1.10: Close Up Views of Mercury Sites ..... 38
Figure 1.11: Electrostatic potential surface representation of NorM_VC ..... 46
Figure 1.12: Stereo view of electrostatic potential surface of NorM_VC ..... 47
Figure 1.13: NorM_VC Internal Cavity ..... 48
Figure 1.14: Conserved Amino Acid Residues of NorM_VC ..... 49
Figure 1.15: The Cation-Binding Site of NorM_VC ..... 50
Figure 1.16: Loss of Cation as a Result of ΔD371N and ΔD371A Mutations ..... 51
Figure 2.1: Experimental Electron Density Map ..... 76
Figure 2.2: Structure of Inward-Facing P-gp ..... 77
Figure 2.3: Close Up Views of Mercury Sites ..... 78
Figure 2.4: X-ray Structure of P-gp in Complex with Nb592 ..... 79
Figure 2.5: Size-exclusion Chromatography (SEC) Plot of P-gp-Nb592 Complex ..... 80
Figure 2.6: Sequence Alignment Between Mouse and Human P-gp ..... 81
Figure 2.7: Nb592 is a Strong Inhibitor of P-gp's ATPase Activity ..... 82
Figure 2.8: Conformational changes by P-gp ..... 83
Figure 2.9: Stereo Views of Fo-Fc Simulated Annealing Omit Map Density for Crystal1 ..... 84
Figure 2.10: Stereo Views of Fo-Fc Simulated Annealing Omit Map Density for Crystal2 ..... 85
Figure 2.11: Stereo Views of Fo-Fc Simulated Annealing Omit Map Density for the P-gp-Nb592 Complex ..... 86
Figure 3.1: Structure and Function of Selenium-labelled Cylcopeptides ..... 102
Figure 3.2: NMR spectra of QZ-Ala, QZ-Leu and QZ-Phe ..... 103
Figure 3.3: Cytotoxicity of the Homotrimeric Cyclopeptide Compounds ..... 106
Figure 3.4: Sensitization of CR1R12 Cells to the Anti-Cancer Drug Colchicine by Homotrimeric Cyclopeptides ..... 107
Figure 3.5: Overview of Mouse P-gp at 3.4Å Resolution ..... 108
Figure 3.6: Overview of P-gp Cyclopeptide Co-Crystal Structures ..... 109
Figure 3.7: Key P-gp Residues Involved in Binding Homotrimeric Cyclopeptides ..... 110
Figure 3.8: Overview of EH2 Ligand Binding Site and Ligand Induced Movement of TM4 ..... 111

## List of Tables

Table 1.1: Affinity of NorM_VC to Its Substrates Determined by Fluorescent Polarization ..... 52
Table 1.2: Experimental Map Phasing Statistics ..... 53
Table 1.3: Model Refinement Statistics ..... 54
Table 1.4: Data Processing of Mutants and Additional Data Sets ..... 55
Table 1.5: Residues in the Internal Cavity ..... 56
Table 2.1: Data Collection, Phasing, and Refinement Statistics ..... 87
Table 2.2: Data Collection Statistics of Single Site Mutations Used for Topological Validation of P-gp Model ..... 88
Table 3.1: Data Collection and Refinement Statistics ..... 112
Table 3.2: Solvent Accessible Surface Area of Cyclopeptide Ligands Buried in the Pocket of P-gp as Calculated by PISA ..... 113

## Acknowledgements

I would like to acknowledge the support and guidance of my graduate committee; Drs. Milton Saier, Geoffrey Chang, Julian Schroeder, Amro Hamdoun, and Randolph Hampton. I would also like to acknowledge the contributions made to this work by my co-workers and lab mates, without whom none of this would have been possible.

Chapter 1 is derived from the following publication, of which I was a major contributing author: "Structure of A Cation-bound Multidrug and Toxic Compound Extrusion (MATE) Family Transporter" Nature, 467:991-994, 2010

Chapter 2 is derived from the publication, of which I was a primary author: "Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide binding domain" Proc Natl Acad Sci, 110:13386-13391, 2013

Chapter 3 is derived from the follow publication, which is currently under submission to ActaD: "Snapshots of Ligand Entry, Malleable Binding, and Induced Helical Movement in P-glycoprotein" in submission.

# Vita

## Education

2001 Bachelor of Science, Virginia Polytechnic and State University

2002-2009 Research Assistant, Department of Molecular Biology, The Scripps Research Institute

2014 Doctor of Philosophy, University of California, San Diego

## Publications

"Structure of the multidrug transporter EmrD from Escherichia coli"

Science, 312:741-744, 2006

"Structure of A Cation-bound Multidrug and Toxic Compound Extrusion (MATE) Family Transporter"

Nature, 467:991-994, 2010

"Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide binding domain"

Proc Natl Acad Sci, 110:13386-13391, 2013

"Snapshots of Ligand Entry, Malleable Binding, and Induced Helical Movement in P-glycoprotein" in submission

## Fields of Study

Major Field: X-ray Crystallography and Structure of Membrane Proteins

Studies in Membrane Protein Structure and Biology

Professors Milton Saier and Geoffrey Chang

# Abstract of the Dissertation

Structural Studies of Multidrug Transporter Proteins

by

Paul Ralph Szewczyk

Doctor of Philosophy in Biology, 2014

Professor Milton Saier, Chair

Multidrug transporters are membrane-imbedded proteins which mediate the traversal of molecules across the cell membrane. Of particular interest herein, the multidrug and toxic compound extrusion (MATE) family of secondary transporters play critical roles in plant metabolite transport, bacterial multidrug resistance (MDR), and mammalian renal function. The outward (extracellular) facing 3.65Å structure reported was the first for this family of transporters, and additionally revealed a site which may serve to bind cations and provide the "power stroke" during the transport cycle. This

work also implemented a novel application of a technique to validate the topology and registration of a moderate resolution structure.

The adenosine triphosphate (ATP)-binding cassette (ABC) family of primary transporters are perhaps best exemplified by p-glycoprotein (P-gp), possibly the most studied of all membrane transporters. In humans, P-gp is expressed throughout the body, including the gut lumen, kidneys, blood-brain barrier, and functions as a ubiquitous molecular "gate-keeper" in the recognition and expulsion of substances from the cytoplasm and intermembrane leaflet. As such, P-gp plays well-categorized roles in mediating the bioavailability of pharmaceuticals, and it's overexpression in some forms of cancer has been linked to MDR. Previous to this work, a 3.9Å structure of mouse P-gp had been elucidated. However, a more detailed structural understanding of the means by which P-gp recognizes and expels molecules has remained nebulous. The work described here expands on previous knowledge by generating a topologically validated novel structure of mouse P-gp extending to 3.4Å, the discovery of a novel epitope binding site on the cytoplasmic nucleotide-binding domain (NBD) face of this molecule, and the first structure-activity relationship study (SAR) for this protein utilizing a rationally engineered series of ligands in order to probe the transporter:substrate relationship.

# Introduction

Multidrug transporters are membrane-resident, predominantly alpha-helical proteins which expel their substrates from the inner membrane leaflet or cytoplasm to the extracellular environment. As such, they have been directly implicated in a number of disease states such as antibiotic resistance in bacteria, and drug resistant cancers. Additionally, by the recognition and transport of exogenous substances they directly impact the pharmaceutical impact of medicines and drugs by altering uptake and clearance rates in the intestines, kidneys, and across the blood-brain-barrier.

Broadly speaking, these proteins can be divided into one of five superfamilies based on phylogenetics, and their predicted topologies. These include the ATP-binding cassette (ABC), small multidrug resistance (SMR), the resistance, nodulation, and cell division (RND), the major facilitator (MFS) and the multidrug and toxic-compound extrusion (MATE) families. Transporters of the ABC family couple the energetics of transport to the hydrolysis of ATP at specialized nuceltide binding domains (NBDs) as primary transporters. The other superfamilies function as secondary transporters where the energetics coupled to substrate transport are derived through proton or ionic electrochemical gradients across the cell membrane itself.

Chapter 1 which follows details the elucidation of NorM, a multidrug transporter protein of the MATE family. This was the first protein structure of this family ever determined, and the last of the five families for which no previous structural information

was available. The moderate, 3.65Å resolution of this structure combined with a previously unseen topology necessitated the development of an experimental method to systematically validate the registration and topology of our model. This was accomplished by the introduction, via multiple separate single site substitution mutants, of cysteine residues into the primary sequence of the transporter protein. Mercury is well known to covalently bind to accessible cysteines trough thiol linkage, and is additionally easily identifiable by crystallographic methods in a given diffraction experiment. Therefore, by generating multiple mutants of NorM at known residue positions, crystallizing the mutants, collecting diffraction data, and analysis we were able to pinpoint the exact position of the mercury-labeled, cysteine substitution, and this then allowed us to validate the registration and topology of the overall model.

This technique was applied in Chapter 2 to several novel structures of murine P-glycoprotein (P-gp) in the validation of a new model for this transporter. This chapter also discusses the structurally determined binding sites to P-gp of a cocrystallized nanobody (Nb). Nanobodies are camelid-derived molecules functionally equivalent to antibodies save for the absence of a light chain. Compared to conventional antibodies, they are extremely thermally stable, resistant to extremes in pH, can readily be re-folded into a function molecule, and can readily cross cellular membranes. These properties have led them as a class to being of immense interest as potential therapeutic scaffolds.

Chapter 3 is a further study into the structural dynamics of P-gp substrate recognition and translocation. This work represents the first structure activity relationship study between P-gp, and a series of rationally engineered cyclopeptide ligands while additionally generating the highest current resolution structure of this

protein at 3.4Å. By systematically increasing the size and hydrophobicity of these ligands, and determining their binding sites by cocrystallization with P-gp, the first clear pictures emerged of ligand induced movements by tis transporter protein. A novel binding site was also observed on the exterior surface of P-gp near proposed ligand entry portals facing the inner leaflet of the cell membrane, providing the first structural glimpse of a possible mode of substrate entry into the binding pocket for this transporter.

This work showcases efforts in the determination of an entirely novel protein structure (the MATE family transporter NorM), and it's validation though the application of known crystallographic phenomena in a novel context. This work also demonstrates a novel epitope on P-gp at the highest current resolution attained, insights into the substrate translocation pathway for this transporter, its induced responses to ligand binding, and a novel entry site for substrate on the exterior of the protein.

Chapter 1:

# Structural Determination of NorM from *Vibrio Cholerea*

## 1.1: Introduction

The export of toxins and substrates from the cell is a fundamental life process, and members of the multidrug and toxic compound extrusion (MATE) family represent the last class of multidrug resistance (MDR) transporters to be structurally characterized. MATEs are involved in a variety of important biological functions across all kingdoms of life. In plants, MATE transporters are highly prevalent with 58 paralogues found in *Arabidopsis thaliana(Hvorup et al., 2003)* and secrete a diverse range of secondary metabolites as a defense against herbivores and microbial pathogens(Morita et al., 2009; Omote, Hiasa, Matsumoto, Otsuka, & Moriyama, 2006). Additionally, plant MATEs play an important role in tolerance towards phytotoxic aluminum in acidic soils, a major limitation of crop production in 50% of the world's arable land(S. Wood, 2000). In mammals, they export a structurally diverse array of xenobiotic cations in the liver and kidney, a process that influences the plasma concentrations of many drugs including metformin, a widely prescribed Type 2 diabetes drug, thereby mitigating therapeutic efficacy(Becker et al., 2009; Tsuda et al., 2009). Bacterial MATEs function primarily as xenobiotic efflux pumps and can confer resistance to tigecycline, a new glycylcycline class antibiotic that has been developed to overcome methicillian-resistant and vancomycin-resistant *Staphylococcus aureus*(Kaatz, McAleese, & Seo, 2005; McAleese et al., 2005).

MATE transporters use either $H^+$ or $Na^+$ gradients across the membrane to drive substrate export, although the coupling mechanism is not well understood. All MATE proteins share ~40% protein sequence similarity(Omote et al., 2006) (Fig. 1.1), suggesting an overall conserved structure and transport function. To provide a basis for understanding the function of MATE transporters, we present two X-ray structures of NorM from *Vibrio cholera* (NorM_VC) with and without rubidium ($Rb^+$) using protein purified in n-dodecyl-β-D-maltopyranoside, and crystallized with n-nonyl-β-glucopyranoside (β-NG) and a novel facial amphiphile(Zhang et al., 2007), 3$\alpha$-hydroxy-7$\alpha$,12$\alpha$-bis[($\beta$-D-maltopyranosyl)ethyloxy] cholane (FA-231) as additives by methods described in(Yin, He, Szewczyk, Nguyen, & Chang, 2006). Although nearly identical, the NorM_VC apo-structure was solved to a higher resolution at 3.65Å. The structure reveals a novel topology distinct from all other MDR transporter families(Aller et al., 2009; Y. J. Chen et al., 2007; Murakami, Nakashima, Yamashita, & Yamaguchi, 2002; A. Ward, C. L. Reyes, J. Yu, C. B. Roth, & G. Chang, 2007; Yin et al., 2006). This outward-facing conformation presents a monovalent cation-binding site within the internal cavity of the transporter. NorM_VC retains transport activity in *E. coli* (Fig. 1.2) and purified NorM_VC protein is active and binds several MATE substrates and drugs(Omote et al., 2006) (Figs. 1.33A and 1.3B; Table 1.1).

**1.2: The Structure of NorM_VC**

The structure of NorM_VC spans ~50 Å in the plane of the lipid bilayer and represents an outward-facing conformation arranged as two bundles of six transmembrane helices (TMs 1-6 and TMs 7-12) forming a large internal cavity open to the extracellular space (Fig. 1.4). The two halves are related by an intramolecular 2-fold that is likely a result of gene duplication consistent with its protein sequence. These halves are connected by a cytoplasmic loop (residues 218-232 between TMs 6 and 7), while the initial helix of each half (TM 1 and TM7) are preceded by a helical extension (residues 2-18, and 233-247) from the inner membrane leaflet side. An additional helix (residues 450-461) after TM12 is nestled under the cytoplasmic side of TM11. The model was obtained as described in sections 1.8, and 1.9 (Fig. 1.5, Table 1.2). The model was refined (Table 1.3) and verified by multiple $F_{obs} - F_{calc}$ simulated annealing omit maps (Figs. 1.6-1.7). The topology and structural correctness of the model was validated using a series of 16 mercury-labeled single cysteine mutant crystals (Figs. 1.8A and 1.8B, 1.1, 1.9–1.10; Table 1.4). In this crystal form, two structurally similar NorM_VC molecules (NorM1 and NorM2) constitute the asymmetric unit allowing independent verification of the position and identity for each mercury-labeled cysteine residue (Fig. 1.10).

The topology of NorM_VC is unique amongst all known transporters and, unlike the ATP-Binding Cassette (ABC), Major Facilitator Super Family (MFS), and Small Multidrug Resistance (SMR) drug efflux pumps, no TM-helices cross over one another (Fig. 1.8C and 1.8D). Two non-equivalent portals open within the lipid bilayer allowing the expulsion of substrates directly to the outer membrane leaflet or the extracellular space (Fig. 1.11). These portals are formed by TMs 1 and 8 on one side and TMs 2 and

7 on the other (Figs. 1.11 and 1.12). At the widest point within the lipid bilayer, the portals are ~12 Å wide and can accommodate the passage of its transport substrates (Table 1.1). The volume of the internal cavity embedded within the lipid bilayer is ~4,300 $Å^3$ (Fig. 1.13), approximately four times larger than the substrate binding pockets of the transcription MDR regulators xPXR and BmrR(Watkins et al., 2001; Zheleznova, Markham, Neyfakh, & Brennan, 1999), and nearly two thirds that of ABC exporters Pgp or MsbA(Aller et al., 2009; A. Ward et al., 2007). Residues facing the internal cavity are both relatively conserved and contributed by all TM helices except for TMs 3 and 9, which are located at the periphery in this conformation (Figs. 1.4 and 1.14). Of the 91 residues facing the cavity, 57 are hydrophobic, 14 are aromatic, 20 are polar, and 5 are charged. Nine aromatic residues are clustered in the c-terminal half of the internal cavity (Table 1.5).

**1.3: The Cation Binding Site of NorM_VC**

In the internal cavity, we observed the binding of $Rb^+$ and also cesium ($Cs^+$) ions, which are heavier alkali metal analogs of sodium more easily visualized by x-ray crystallography (Fig. 1.15A-1.15B)(Yin et al., 2006). These cations bind to a site in the C-terminal half of NorM_VC bounded by residues from TMs 7, 8, and 10-12, a fundamentally different topology then found in members of the MFS family (Fig. 1.8C-1.8D). Previous studies have identified three conserved acidic residues that are critical for transport function in NorM from *V. parahaemolyticus* (NorM_VP)(Otsuka et al., 2005), a closely related ortholog of NorM_VC with 76% sequence identity and 86% similarity(Altschul et al., 1997; Altschul et al., 2005). These residues correspond to

D36, E255, and D371 in NorM_VC (D32, E251, and D367 in NorM_VP). In our $Rb^+$-bound NorM_VC structure, residues E255 and D371 face the internal cavity on the C-terminal half of the molecule and are located near the cation (Fig. 1.15A). For NorM_VP, both of these negatively charged residues are important for transport function as certain mutations at these positions decouple substrate/$Na^+$ antiport(Otsuka et al., 2005). Like NorM_VP, mutations at residue 255 in NorM_VC affect protein folding and, as a consequence, do not express well(Otsuka et al., 2005). Substitution mutations at D371 to alanine or asparagine abolish $Rb^+$ and $Cs^+$-binding in the outward-facing conformation (Fig. 1.16). These findings suggest that these and possibly other surrounding residues are crucial for the recognition and binding of cations from the extracellular space. Mutations of D367 in NorM_VP (D371 in NorM_VC) disrupt transport activity, further suggesting that cation binding at this site plays an important role in the transport cycle. In our outward-facing NorM_VC structure, residue D36 (D32 in NorM_VP) is located in the N-terminal half, far from the C-terminal cation binding site. We did not observe the presence of either $Rb^+$ or $Cs^+$ near D36 and speculate that this residue may play an important role in other conformations during the transport cycle.

**1.4: The Transport Cycle of NorM_VC**

We propose that the outward-facing conformation of NorM_VC, like other efflux pumps(Jardetzk.O, 1966), represents a state in the transport cycle that has high-affinity for monovalent cations (Fig. 1.15C; step 1), and a lower-affinity for substrates. Although the internal cavity is accessible to mercury compounds (Figs. 1.8A, 1.8B and

1.9), this conformation is not amenable to co-crystallization or soaking with several well-established drugs or transport substrates (Table 1.1). Upon cation binding, we propose that NorM_VC undergoes an outward- to inward-facing conformational change more favorable to substrate binding (Fig. 1.15C; step 2). Herein, cation is released into the cytoplasm, and the binding of substrate induces a structural change back to the outward-facing conformation(Jardetzk.O, 1966), where the substrate is released into the outer leaflet of the lipid bilayer and/or extracellular space (Fig. 1.15C; step 3).

The structure of NorM_VC is representative of the last MDR transporter family to be elucidated by x-ray crystallography. Unlike MDR transporters from the Resistance, Nodulation, and Cell-Division (RND) family where substrate binding is mediated by an extracellular domain, MATEs along with other MDR transporters from the ABC, MFS, and SMR families have poly-specific drug binding sites composed of TM helices within the lipid bilayer(Aller et al., 2009; Y. J. Chen et al., 2007; A. Ward et al., 2007; Yin et al., 2006). Despite wide divergence in their primary protein sequences, the extraction and expulsion of hydrophobic drugs and substrates directly from the inner to the outer membrane leaflet through portals facing the lipid bilayer (Fig. 1.15C) may be a common theme for these transporter families. The V-shaped conformations among these structures suggest that this may indeed be the case, forming a common molecular basis for the transport of their hydrophobic and amphipathic substrates across the lipid bilayer.

**1.5: Cloning, Mutagenesis, Protein Expression, and Purification**

*NorM* (NCBI accession number AE003852.1 REGION: 1652328..1653713) was cloned from *Vibrio cholera* genomic DNA (ATCC 39315) into the expression vector pET19b (Novagen Inc.) between the *NdeI* and *XhoI* restriction sites of the multiple cloning site. The pET19b vector bears an N-terminus His10-tag followed by an enterokinase cleavage site upstream of the multiple cloning sites. Single cysteine substitution mutants for structure and topology verification were generated from the pET19b-*NorM_VC* construct using the QuickChange™ Site-directed Mutagenesis Kit (Stratagene), following the protocol from the manufacturer. Cloning and mutagenesis primers were likewise designed according to the manufacturer's instructions and synthesized by Integrated DNA Technologies. PET19b-*NorM_VC* plasmid was transformed into, and expressed in *Escherichia coli* BL21 DE3 (Invitrogen Inc.) as 100L batches grown in a bioreactor (BioFlow 5000, New Brunswick Scientific) from an initial 1 L seed culture. Cells were grown to $OD_{600}$=0.550 at 37°C and induced with 0.4 g/L IPTG (BioPioneer) at 30°C for 4 hours. All media used was low-salt LB broth. Cells were pelleted by centrifugation at 15,900 g and ruptured by single passage through a cell disruptor at 60 psi (Model 110L, Microfluidics). NorM_VC was solubilized with 1% n-dodecyl–β-D-maltopyranoside (β-DDM, Anatrace) in 20 mM NaCl, 20 mM Tris-HCl pH 8.0, 30 mM imidazole, and 5% glycerol for four hours. The insoluble fraction was removed by centrifugation at 38,400 g for 45 minutes, and NorM_VC purified from the supernatant by Ni-affinity chromatography. Elutent NorM_VC was concentrated in an Amicon Ultra (MW cutoff = 50 kD) centrifugal device (Millipore), desalted through a HiPrep™ 26/10 column (GE Healthcare) in 20 mM Tris-HCl pH 7.5, 20 mM NaCl,

0.012% β-DDM, and 1 mM EDTA desalting buffer, then re-concentrated to 40 mg/ml. The N-terminal His10-tag was removed by enterokinase (New England Biolabs) digestion in the manufacturer's recommended buffer for 12 hours at 16°C at a protein to enzyme stock ratio of 25:1 (volume:volume). Following digestion, the free His10-tag was removed by Ni-affinity chromatography, the tag-free NorM_VC protein concentrated to 40 mg/ml, then ultracentrifuged at 323,400 g at 4°C for 55 minutes. Protein purity was evaluated by SDS PAGE and the concentration was determined by spectrophotometry at $A_{280}$ using an extinction coefficient $\varepsilon_{280}$ = 56.4 x $10^{-3}$ $M^{-1}cm^{-1}$. Purification of the mutants followed an identical protocol.

**1.6: Crystallization of NorM_VC**

Crystals of NorM_VC were grown from β-DDM solubilized protein and 0.23% n-nonyl-β-glucopyranoside (β-NG) using the following range of conditions: 50mM Tris-HCl pH 7.2-8.6, 87 mM $(NH_4)_2SO_4$, and 16-24% polyethylene glycol 250 dimethyl ether (PEG 250DME; Omega Scientific) in deuterium oxide (Cambridge Isotope Laboratories, Inc.). Crystals were grown by the sitting drop method from protein samples in a concentration range of 6-12 mg/ml at 22°C, with a protein to precipitant ratio of 1:1 (volume:volume) in a 4 μl drop. The crystals appeared after three to four days, reaching peak size within a week. All single site substitution cysteine mutants were crystallized in this manner (Table 1.4). We discovered that the addition of a novel facial amphiphile, 3$\alpha$-hydroxy-7$\alpha$,12$\alpha$-bis[(β-D-maltopyranosyl)ethyloxy] cholane (FA-231)(Zhang et al., 2007), to the crystallization conditions improved the

reproducibility of the crystals lowering background precipitation. Both our highest diffracting crystal, Crystal1 (3.65 Å; *a*=142.8 Å, *b*=240.8 Å, *c*=45.7 Å, α,β,γ = 90$^o$) and the $Rb^+$-bound crystal, Crystal2 (4.2 Å; *a*=159.6 Å, *b*=241.7 Å, *c*=46.2 Å, α,β,γ = 90$^o$) were grown in the presence of 0.02% FA-231 (Table 1.3). In addition, the two cation-site mutants, D371A and D371N, were also crystallized in the presence of the facial amphiphile (Table 1.4). We attribute the higher diffraction quality of Crystal1 mostly to serendipity as there was a wide prep dependence on crystal growth.

**1.7: Derivatization and Cryoprotection**

Derivatization of native and single cysteine mutant crystals for phasing and structure verification was accomplished by soaking with either 5mM mercury (II) acetate ($Hg(OAc)_2$), 5mM mercury (II) cyanide ($Hg(CN)_2$), or 5mM Bakers dimercurial (BDM), in fresh mother liquor made with 24% PEG 250DME and supplemented with 0.009% β-NG, and 0.016% β-DDM as cryoprotectant for periods ranging between 30 minutes to an hour. Rubidium and Cesium derivatization of native and cation-binding site mutant crystals (D371A and D371N) was achieved by soaking with 200 mM RbCl, or 200 mM CsCl, in mother liquor described above. RbCl, CsCl, and 100 mM NaCl (used as a control) soaks all elicit rapid crystal degradation and precluded cation derivatization beyond a minute, likely due to disruption of lattice contacts due to protein conformational changes.

**1.8: Data Acquisition and Structure Determination**

The initial experimental electron density map and structure of NorM_VC was determined by multiple anomalous dispersion (MAD) phasing and phase combination of three data sets: a ΔT130C crystal derivatized by $Hg(OAc)_2$, and two native crystals soaked with $Hg(CN)_2$ and BDM, respectively (Table 1.2). Protein phasing, solvent flattening, non-crystallographic symmetry (NCS) averaging, and phase extension was performed using the programs PHASES, SOLVE/RESOLVE, and CNS v1.2(Brunger, 2007; Brunger et al., 1998; Furey & Swaminathan, 1997; Terwilliger, 2000; Terwilliger & Berendzen, 1999). The experimental electron density maps revealed two molecules in the asymmetric unit (NorM1 and NorM2) (Fig. 1.5). The starting electron density correlation coefficient from the application of NCS averaging between NorM1 and NorM2 was 60% (Table 1.2).

To verify the topology and structure of NorM_VC, we generated 16 unique single site substitution mutants each replacing a native residue with a cysteine. These single cysteine mutants were crystallized and derivatized as described above. Diffraction data was collected at synchrotron sources (Table 1.4) and analyzed by the processing programs HKL2000 and MOSFLM(Leslie, 1992; Otwinowski & Minor, 1997). The isomorphous difference Fourier synthesis for each mutant data set (Table 1.4) was generated with the CNS v1.2 software suite(Brunger, 2007) using the model from Crystal 2 with the $Rb^+$ removed. For each of these calculations, model phases were generated by rigid-body refinement and gradient minimization using this model and the mutant data (Table 1.4). For each mutant, two NCS-related isomorphous difference peaks corresponding to a mercury site were observed for each molecule in

the asymmetric unit (NorM1 and NorM2) confirming the position and identity of each cysteine mutation (Fig. 1.8, 1.9-1.10, and Table 1.4). A composite of all the isomorphous difference Fourier maps of all cysteine mutants (Fig. 1.8 and 1.9) was generated by real-space map transformation of each isomorphous difference map (Fig. 1.10) to the coordinates of the model from Crystal 2. The isomorphous difference Fourier maps showing the presence of $Rb^+$ or $Cs^+$ (Fig. 1.15A-1.15B); and their absence as a consequence of cation-binding mutations D371A and D371N (Fig. 1.16) were calculated in a similar fashion.

Crystallographic refinement using data from Crystals 1 and 2 was accomplished using the maximum likelihood target with amplitudes (mlf) using CNS v1.2(Brunger, 2007). Endogenous and mutant-derived mercury site positions were used to validate the refined model (Fig. 1.8, and 1.9-1.10). A final round of group B-factor refinement and bulk solvent correction produced a single model with $R_{cryst}$ of 31.2% and $R_{free}$ of 34.3% for Crystal 1 and with $R_{cryst}$ of 30.9% and $R_{free}$ of 34.2% for Crystal 2 (Table 1.3). Additional modeling restraints used during crystallographic refinement included hydrogen-binding constraints for all secondary structure elements and harmonic non-crystallographic constraints (150 kcal/mole) between NorM1 and NorM2.

All structures of NorM were verified using a sigma weighted 2Fo-Fc composite simulated annealing (SA) omit map (iteratively omitting 5% of the model) as well as multiple Fo-Fc SA difference maps (Fig. 1.6-1.7). The Fo-Fc maps were calculated using CNS v1.2(Brunger, 2007) systematically omitting 8 consecutive residues throughout the model, generously omitting a neighboring sphere size of 4.0 Å and a

map cushion surrounding the omitted region of 2.0 Å. The Fo-Fc maps were composed into a single continuous map for confirming the residue positions.

### 1.9: Model Building and Figures

All models were built using the programs CHAIN, Coot, and PyMOL(DeLano, 2009; Emsley & Cowtan, 2004; Sack, 1988). Stereo figures were generated using PyMOL and Adobe Photoshop 7.0.

### 1.10: NorM_VC Substrate Binding Activity by Fluorescence Polarization

A fluorescence based approach was used with well known MATE substrates (*10*) to characterize the binding activity of purified detergent solubilized NorM_VC as summarized (Table 1.1). The procedure of these fluorescence polarization experiments was similar to that described previously(Y. J. Chen et al., 2007). The experiments were accomplished using a microplate reader equipped Beckman DTX 880 fluorimeter and *Nunc*™ 384-well plates. Briefly, protein samples were two-fold serially diluted into 20μl of desalting buffer with the addition of 0.23% β-NG (Desalting-NG) to obtain conditions similar to those used for crystallization. Equal amounts of substrate were then added into each well. The reaction mixture was pre-incubated at $25^0$C for 10 min and fluorescence was measured at the same temperature. All experiments were carried out in at least duplicate and experiments with each substrate were repeated at least three times using protein samples from different preparations (Fig. 1.3A, 1.3B and Table 1.1).

Experimental polarization values were calculated according to the equation $P = 1000 \times (I_V - G \times I_H)/(I_V + G \times I_H)$, where $P$ is the polarization value in each well, $I_V$ is the fluorescence intensity polarized parallel to the excitation light, and $I_H$ is the fluorescence intensity polarized perpendicularly. $G$ equals 0.65 and is the intensity ratio of the vertical to horizontal components of emission when the sample is excited with horizontally polarized light, $G = I_{HV}/I_{HH}$. To determine the dissociation constant, $K_D$, the experimental P values vs. NorM_VC protein concentration were plotted and fitted to the equation for single binding site protein-substrate complex formation: $y = P_o + (P_{max} - P_o)([\text{NorM\_VC}]/([\text{NorM\_VC}] - K_D))$, where $P_o$ is the minimum fluorescence polarization value of free substrate, $P_{max}$ is the maximum fluorescence polarization when all fluorescent molecules are bound in protein-substrate complex, and [NorM_VC] is the protein concentration in each well, with $P_o$, $P_{max}$, and $K_D$ treated as variable during the calculations to achieve the least fitting error.

A competitive binding assay was used to determine the affinity of NorM_VC to non-fluorescent substrates (Fig. 1.3C, 1.3D and Table 1.1). Non-fluorescent substrate was two-fold diluted as described above. NorM_VC was pre-incubated with fluorescent substrate for 20 min and added into each well at equal concentrations. The reaction mixtures were further incubated for another 30 min before fluorescence polarization was measured. The dissociation curve of the displaced substrate was plotted as its polarization vs. concentration of the competing substrate. The half of maximal inhibitory concentration ($IC_{50}$) was obtained by fitting the dissociation curve to the equation $y = P_o + (P_{max} - P_o)(1/(1 + 10^{\log([TPP]/IC50)})$, where $P_o$, $P_{max}$, and $IC_{50}$ were treated as variable. And the inhibition constant ($K_i$) was calculated using formula $K_i =$

$IC_{50}/(1$ + [fluorescent substrate]$/K_D)$, where $K_D$ is the dissociation constant for the protein-fluorescent substrate complex. Experimental data were calculated and plotted in Excel. The non-linear fitting of data was performed in Origin 7 (Origin Lab, Inc.)..

**1.11: NorM_VC Transport Activity by Ethidium Bromide (EtBr) Accumulation**

The accumulation assay used was described elsewhere(X. Z. Li, Poole, & Nikaido, 2003) with some modifications. Host strain *E. coli* BL21 DE3 Δ*acrAB*, a generous gift from Prof. M. Saier at UC San Diego, was used to reduce the background efflux due to endogenous MDR transporters. Briefly, these cells were transformed with pET19b or pET19b-*NorM_VC*, grown to mid-log phase ($OD_{600}$=0.7) at 200 rpm 37°C, and induced with 0.4 g/L IPTG at 90 rpm 30°C overnight. The cell cultures were washed with 50 mM Tris HCl pH 7.0, resuspended to $OD_{600}$=0.7, and incubated with 300 μM EtBr for 60 min at 25°C in triplicate aliquots. Cells were harvested by centrifugation and the supernatant aspired. Cell pellets were washed three times with the above buffer, resuspended in 1 ml of 3% trichloracetic acid, and vortexed vigorously. The cell debris was separated by centrifugation at 18,000 rcf for 10 min. Fluorescence of the EtBr in the supernatant was measured in 50 μL samples in 384-well format plates using a Beckman DTX 880 fluorimeter with excitation at 485 nm and emission at 595 nm. The assay was conducted in triplicate and the data averaged (Fig. 1.2).

Chapter 1 is derived from the following publication, of which I was a major contributing author: “Structure of A Cation-bound Multidrug and Toxic Compound Extrusion (MATE) Family Transporter” Nature, 467:991-994, 2010

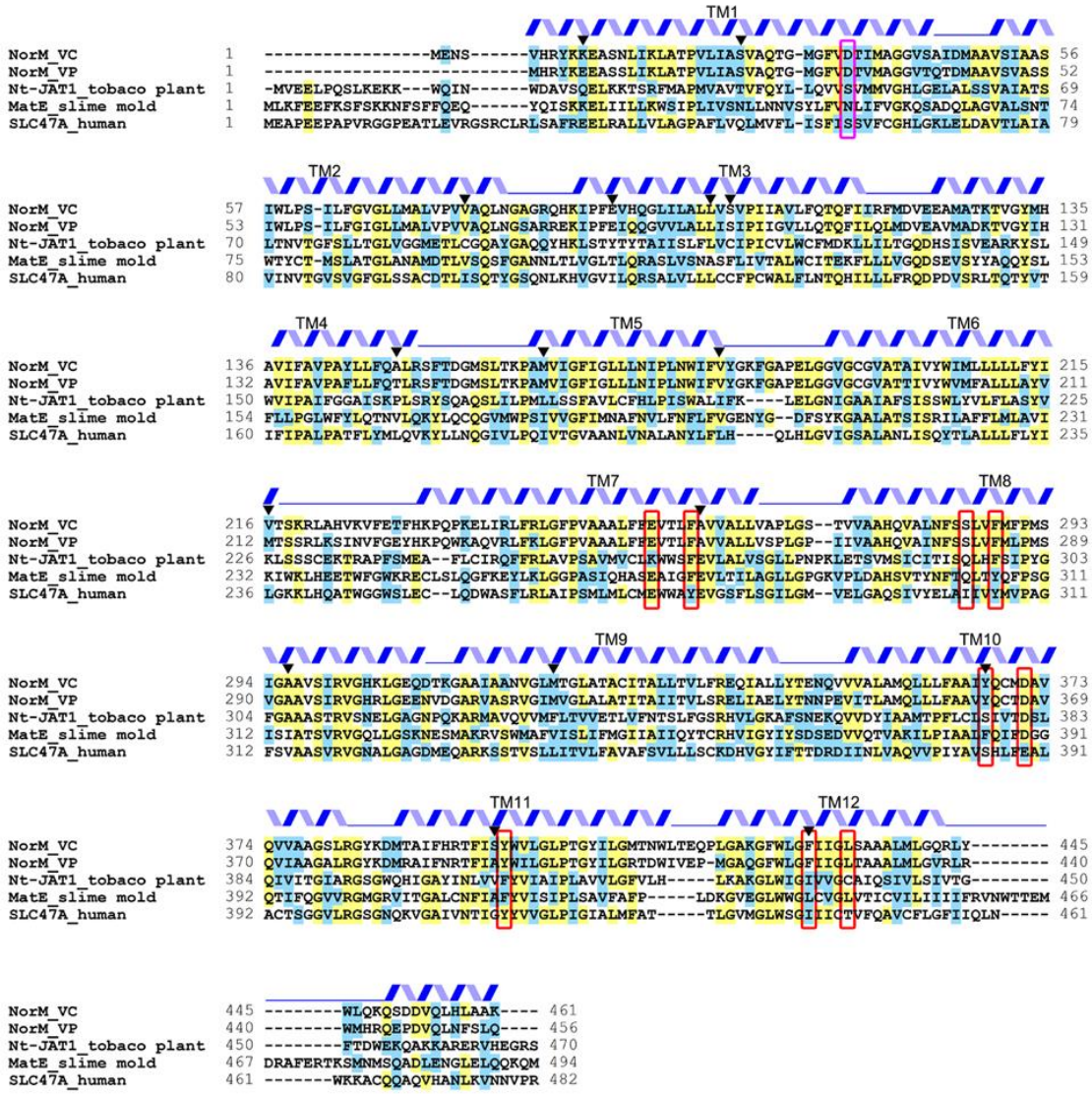



Figure 1.1: Amino acid sequence alignment of NorM_VC. Similar residues are shaded blue, identical residues are shaded yellow. Threshold for shading is 60%. NCBI reference sequences used are: AAF94694.1 (NorM *V. cholera*), BAA31456.1 (NorM *V. parahaemolyticus*), CAQ51477.1 (Nt-JAT1 *Nicotiana tabacum*), XP_642174.1 (MatE *Dictyostelium discoideum AX4*), and NP_060712.2 (SCL47A homo sapiens). The extra C-term sequences of both slime mold and human MATE are not comparable to the other orthologs and omitted for clarity. Secondary structure of NorM_VC is shown as a cartoon above the sequence alignment. Residues that were mutated to cysteines for structure and topology verifications are indicated by black arrowheads. Every TM helix in our structure has at least one mercury-labeling cysteine corresponding mutant. Residues that constitute the cation-binding site (Fig. 1.13 and Table 1.5) and are highlighted with red rectangles. Residue D36 (highlighted with a pink rectangle) corresponds to D32 in NorM_VP.

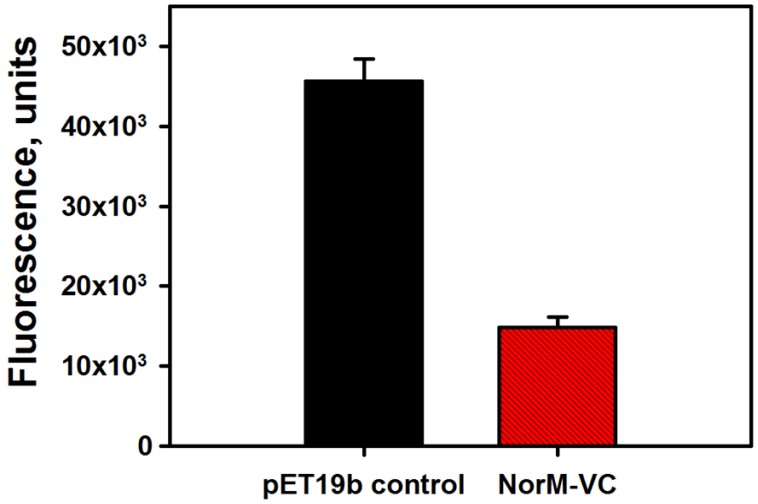



Figure 1.2: Effect of recombinant NorM_VC on EtBr uptake. Cells were transformed with pET19b vector alone as a control (black column) or with pET19b-*NorM_VC* (red column). Fluorescence intensity values reflect EtBr accumulation between cell samples of equal mass.

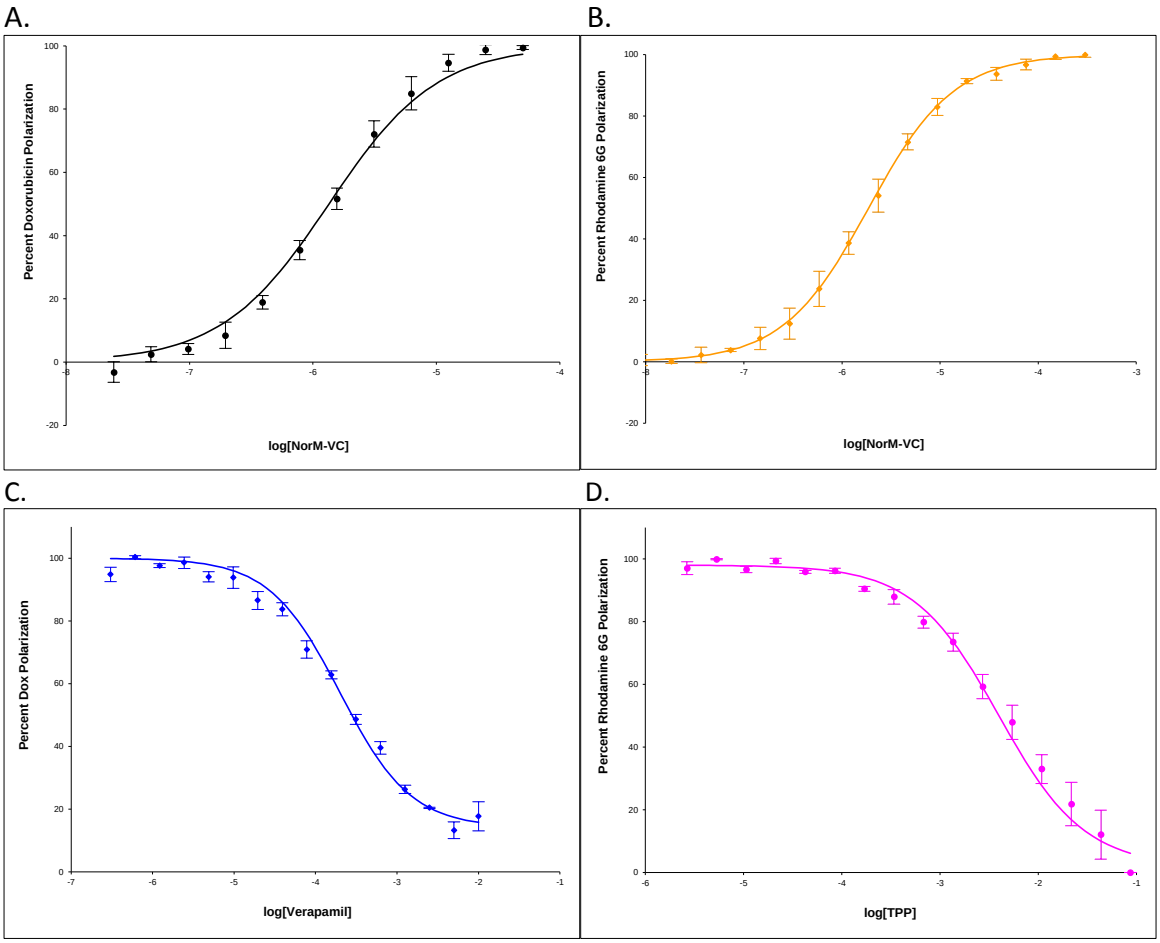



Figure 1.3: Substrate binding activity of NorM_VC. (A) NorM_VC binds to doxorubicin (●) and (B) to rhodamine 6G (◆), as determined by fluorescence polarization. Data is plotted as percent substrate polarization versus the logarithim of NorM_VC concentration (molar). Curves were fit by the Michealis-Menton equation. (C) Competitive binding of verapamil in the presence of 1 μM doxorubicin (◆), and (D) TPP in the presence of 1 μM rhodamine G6 (●) to NorM_VC. Data is plotted as percent polarization of the measured fluorescent substrate versus the logarithm of competitive substrate concentration (molar). The fitted curves model single-site competition. The diminishing polarization indicates dissociation of the fluorescent compound from NorM_VC due to its displacement by the competing substrate. For verapamil competitive binding to NorM_VC against doxorubicin, 10 μM doxorubicin and 10 μM NorM_VC were used. For tetraphenylphosphonium (TPP) competitive binding against rhodamine G6 (RG6), 1 μM RG6 and 20 μM NorM_VC were used. Data points represent the mean of three independent runs using distinct protein populations with at least two replicates per run. Error bars shows the standard deviation among runs. Dissociation and inhibition constants, excitation, and emission wavelengths are presented in Table 1.1.

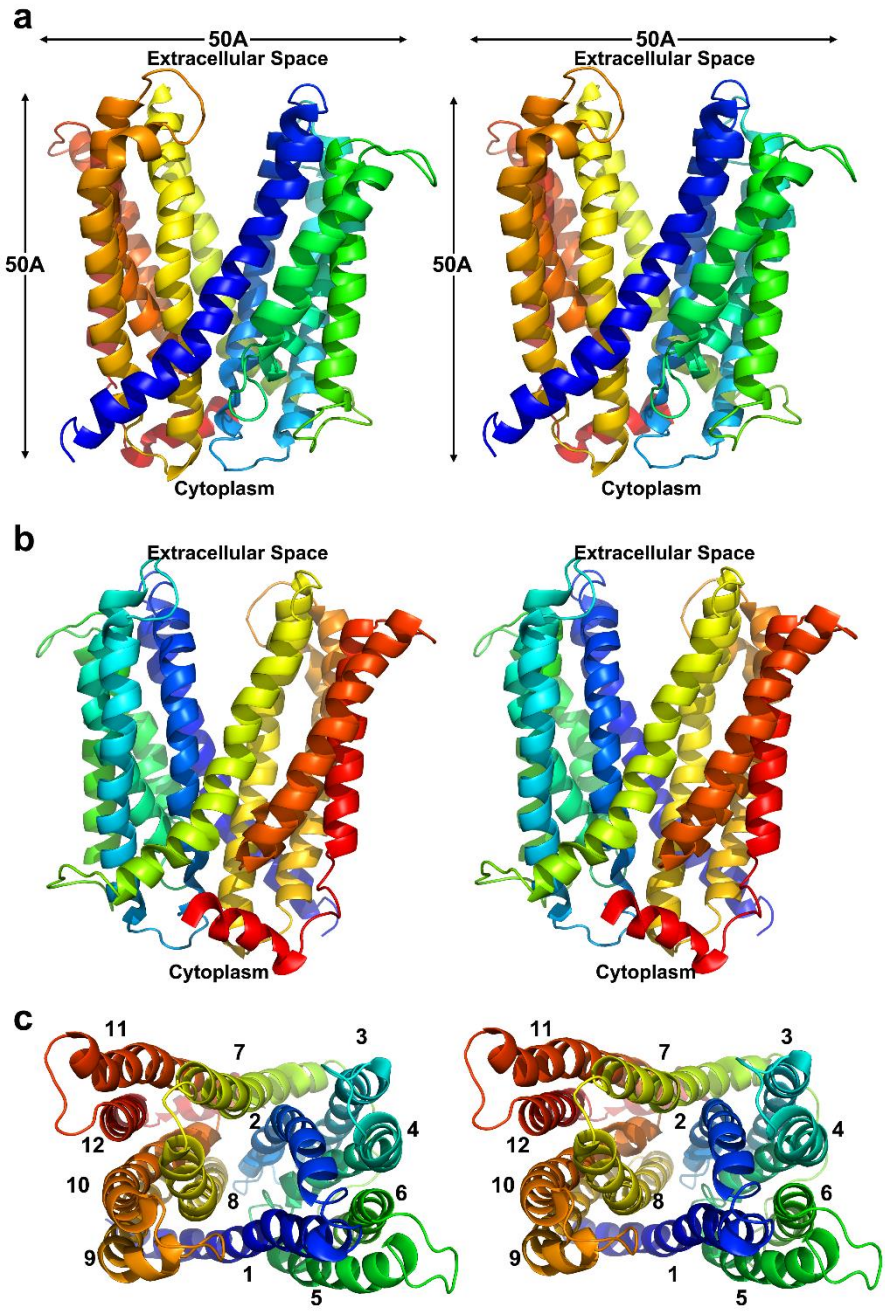
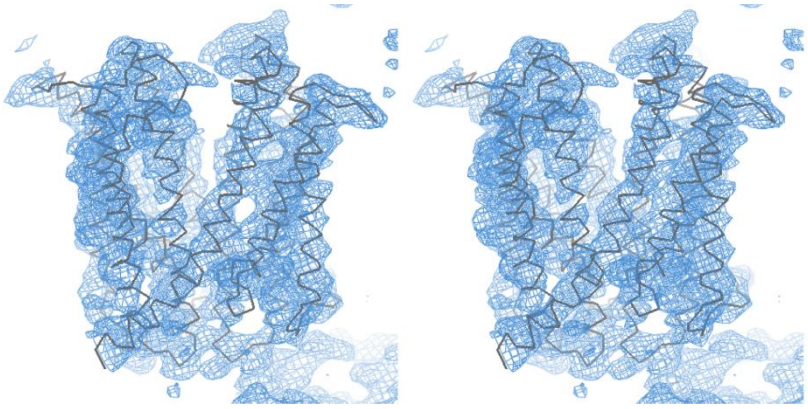



Figure 1.4: Stereo-views of NorM_VC. a, Front and b, back views of NorM_VC, the extracellular and cytoplasmic sides of the molecule are indicated. c, Extracellular view of NorM_VC with TM helices 1-12 marked. The internal cavity is open to the extracellular space and is occluded on the cytoplasmic side. The molecule is colored using a rainbow gradient from the N- (blue) to C-terminus (red).

A



B.

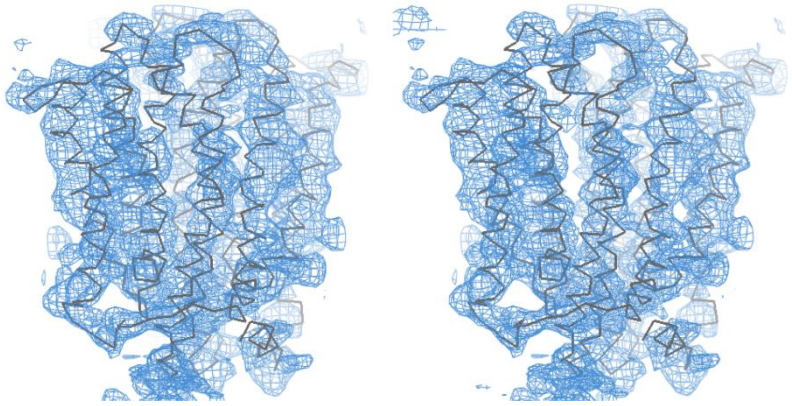


Figure 1.5: Experimental electron density map of NorM_VC. (A-B) Stereo views of electron density (blue mesh) contoured at 1 σ derived from NCS averaging of the dataset at 4.5Å. NorM_VC monomer is presented in gray ribbon. Additional electron density surrounding this molecule is from symmetry-related copies in the unit cell. The map was generated using the program PHASES(Furey & Swaminathan, 1997) as described.

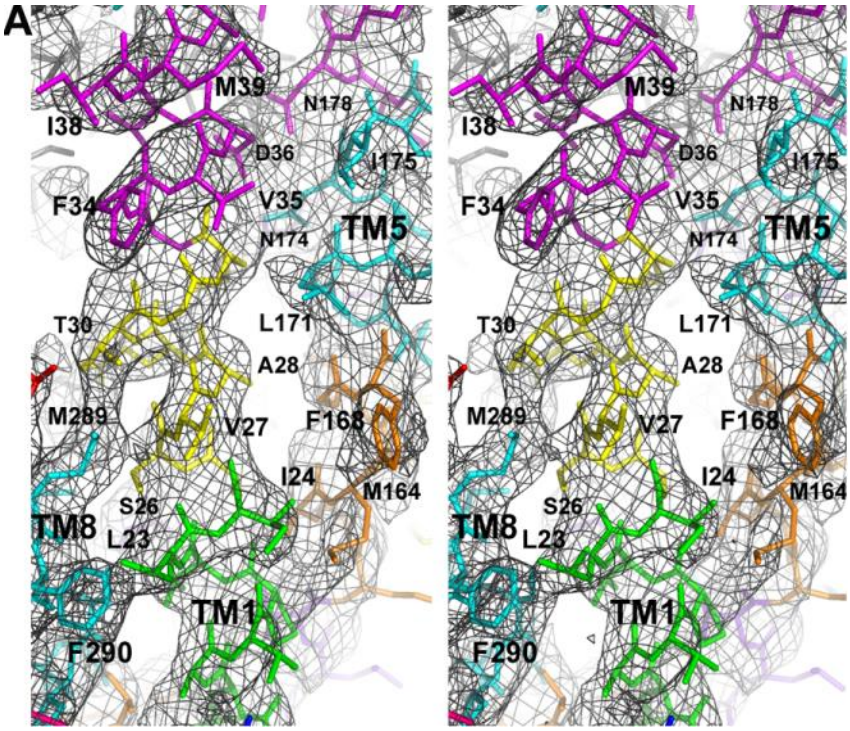



Figure 1.6: (A-J) Stereo views of Fo-Fc simulated annealing omit density for TM helices of NorM_VC (Crystal 1) with multiple density maps superimposed. (K) Stereo view of the C-terminal helix following TM12. (L-O) Stereo views of select loop regions. The Fo-Fc density maps (grey) were generated as described in the text. Difference density for an entire NorM_VC molecule was achieved by superimposing individual Fo-Fc maps calculated using sequential segments of eight residues that were iteratively omitted from the model calculated with a neighboring sphere size of 4Å and a map cushion of 2Å. The windows of residues omitted from the model are shown in different colors. Maps B-D and F-K were contoured at 2.0σ while A, E, and L-O were at 1.5 σ.

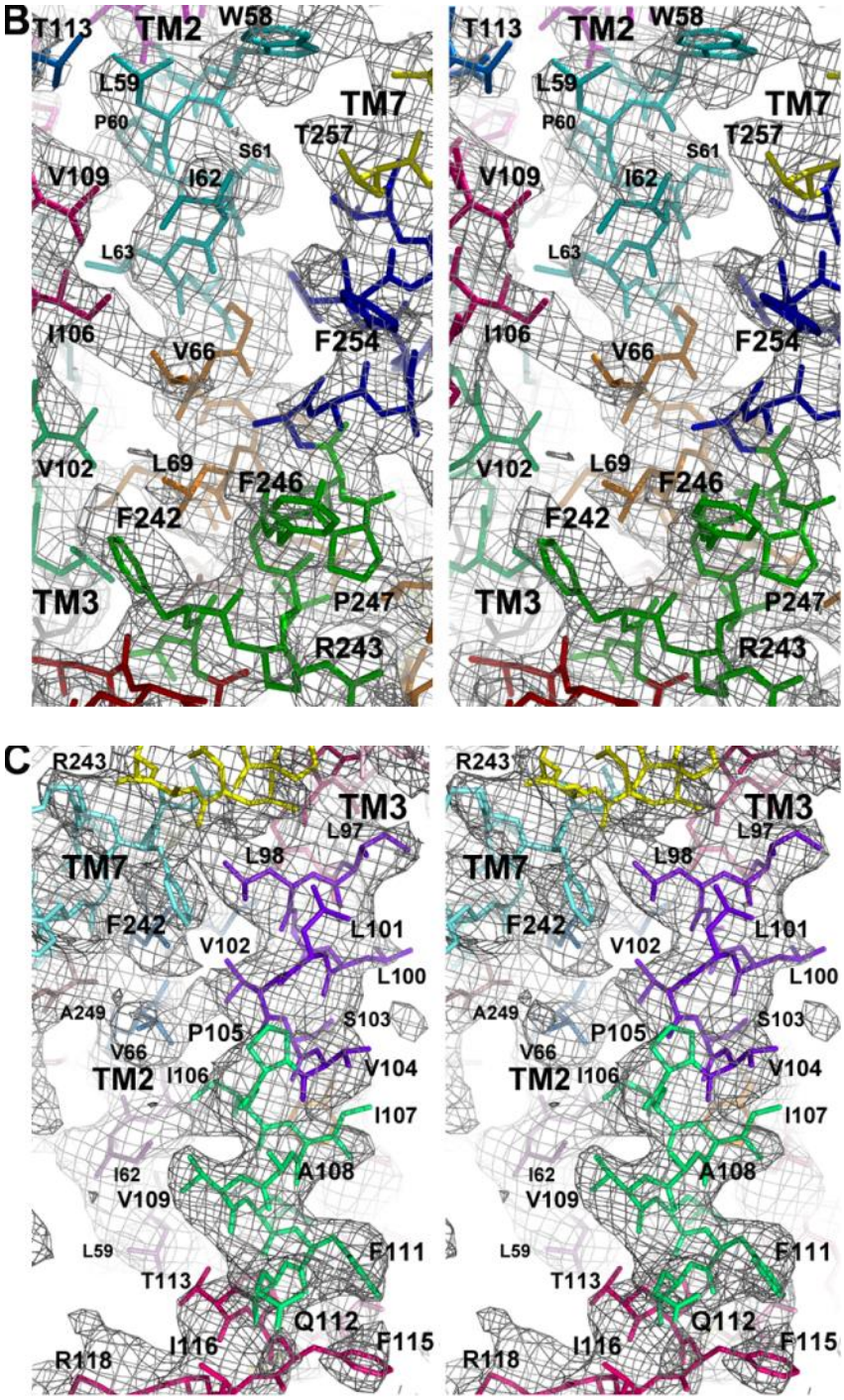



Figure 1.6: (A-J) Stereo views of Fo-Fc simulated annealing omit density for TM helices of NorM_VC (Crystal 1) with multiple density maps superimposed, continued.

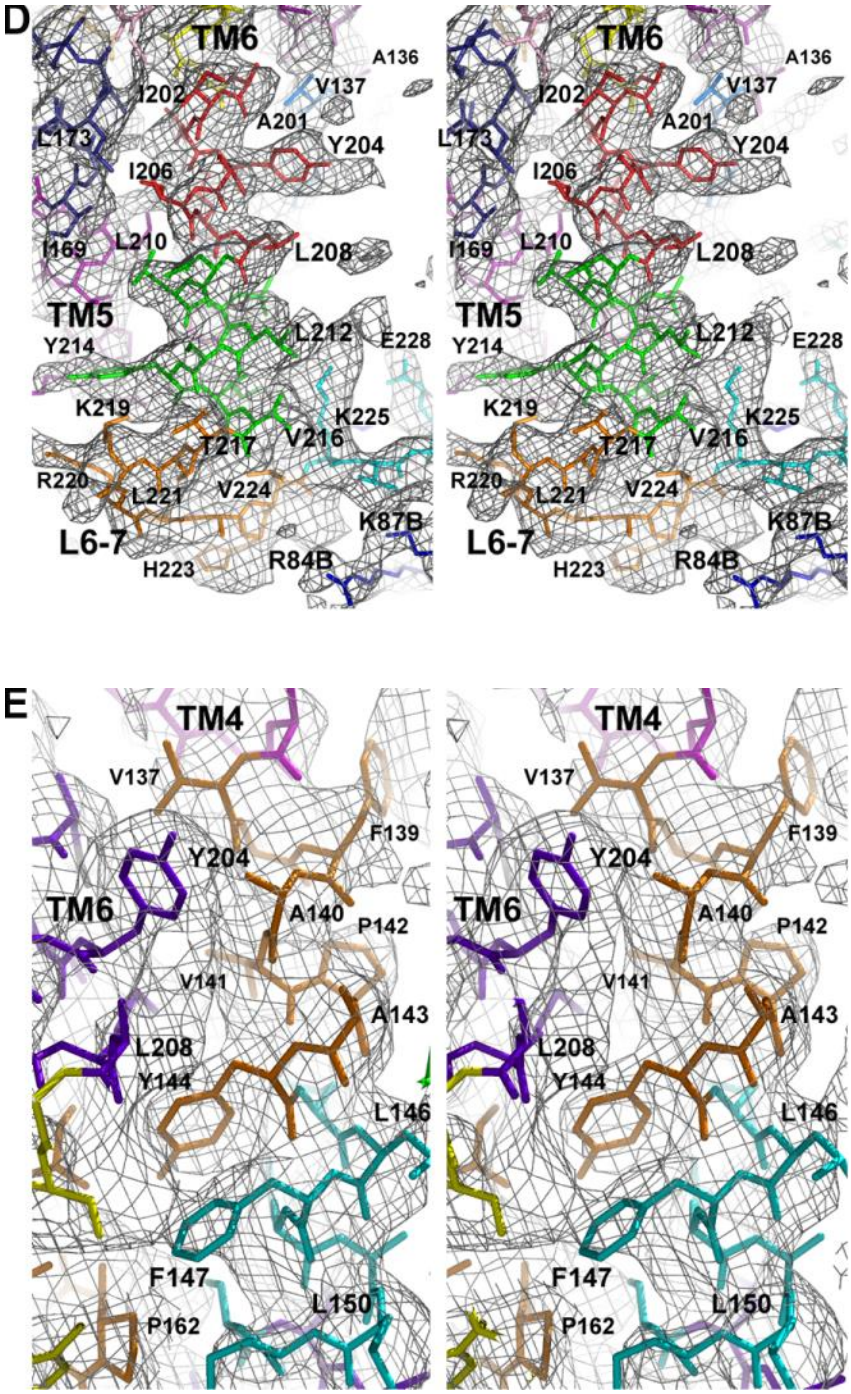



Figure 1.6: (A-J) Stereo views of Fo-Fc simulated annealing omit density for TM helices of NorM_VC (Crystal 1) with multiple density maps superimposed, continued.

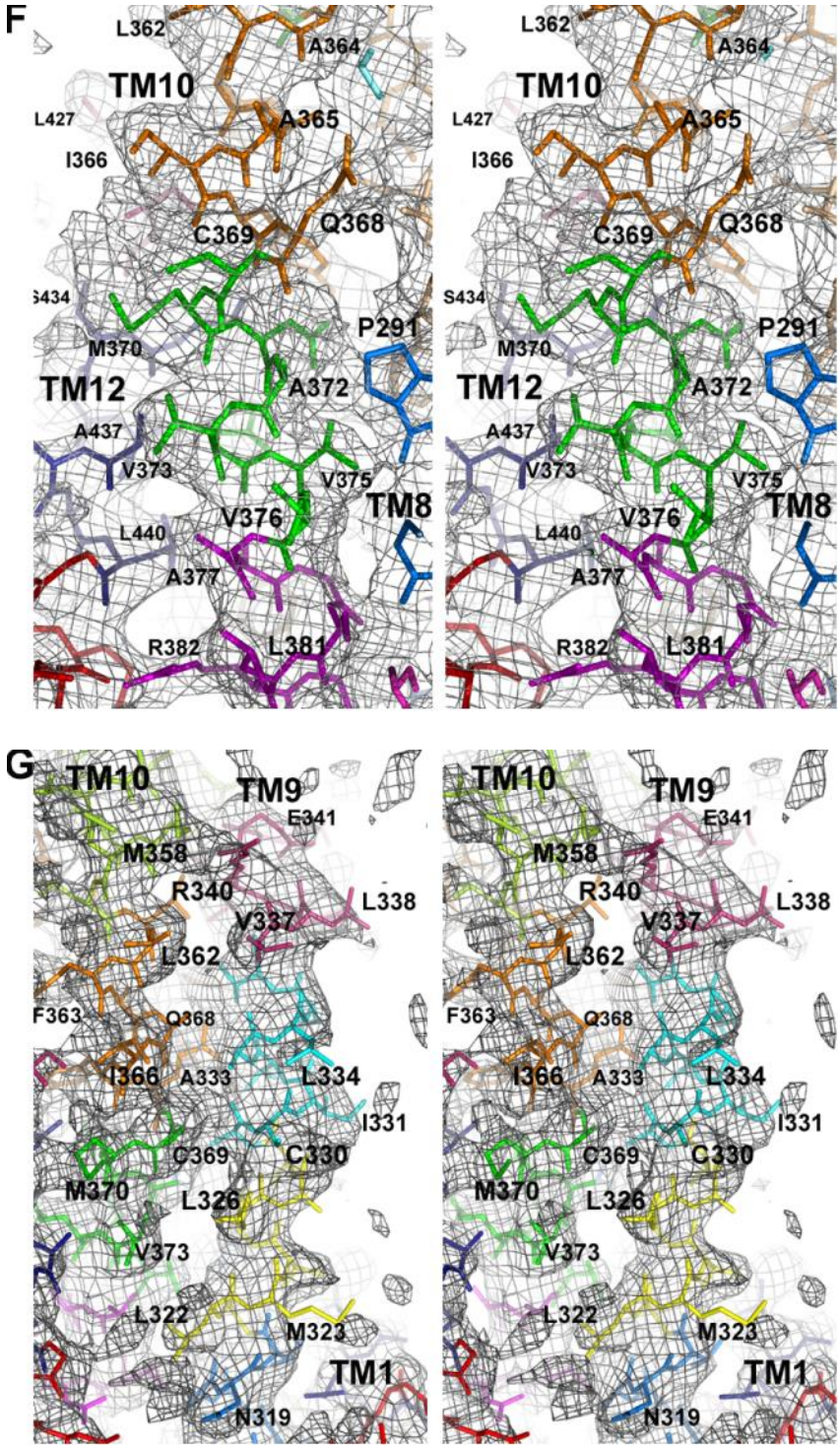



Figure 1.6: (A-J) Stereo views of Fo-Fc simulated annealing omit density for TM helices of NorM_VC (Crystal 1) with multiple density maps superimposed, continued.

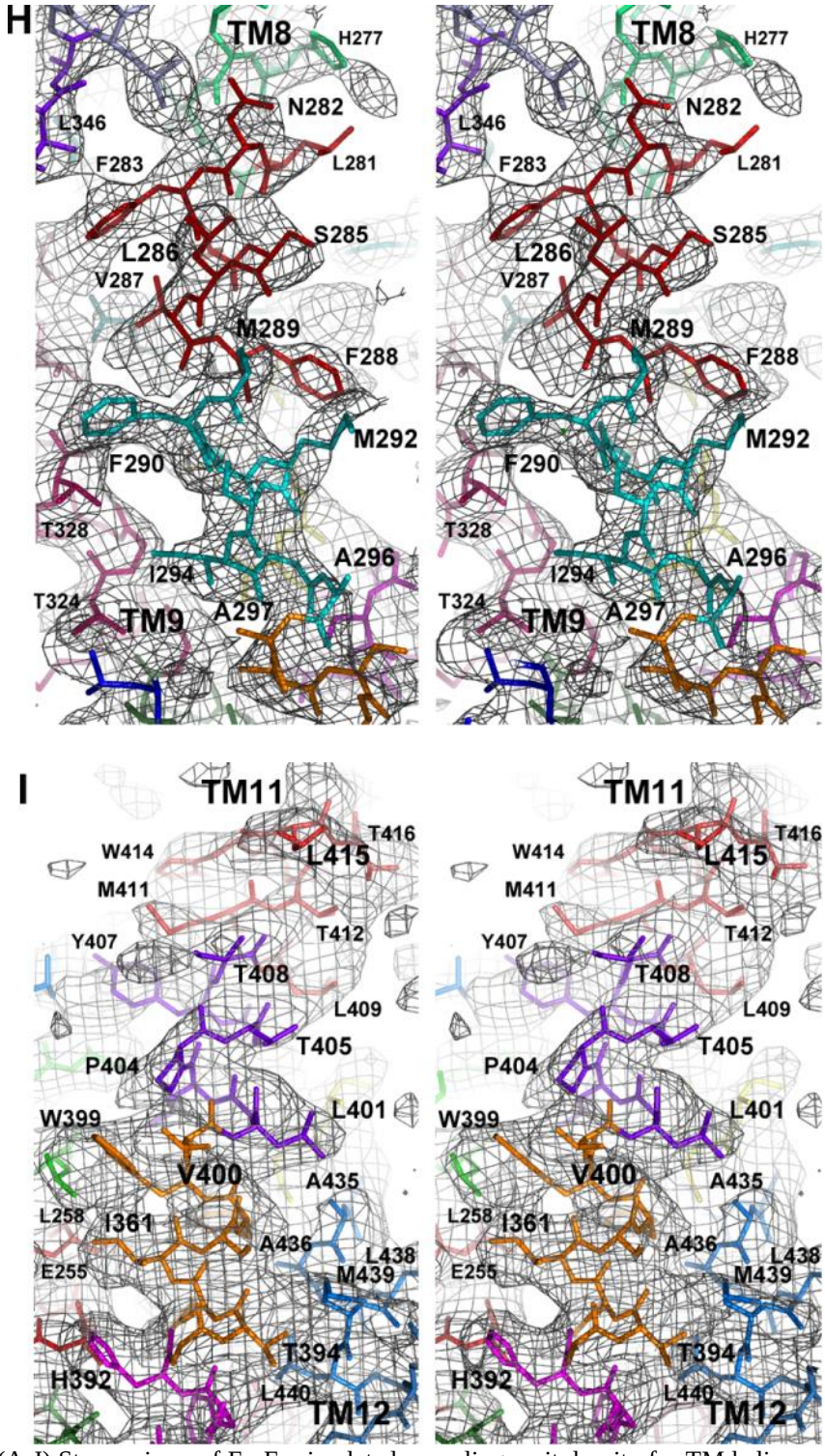



Figure 1.6: (A-J) Stereo views of Fo-Fc simulated annealing omit density for TM helices of NorM_VC (Crystal 1) with multiple density maps superimposed, continued.

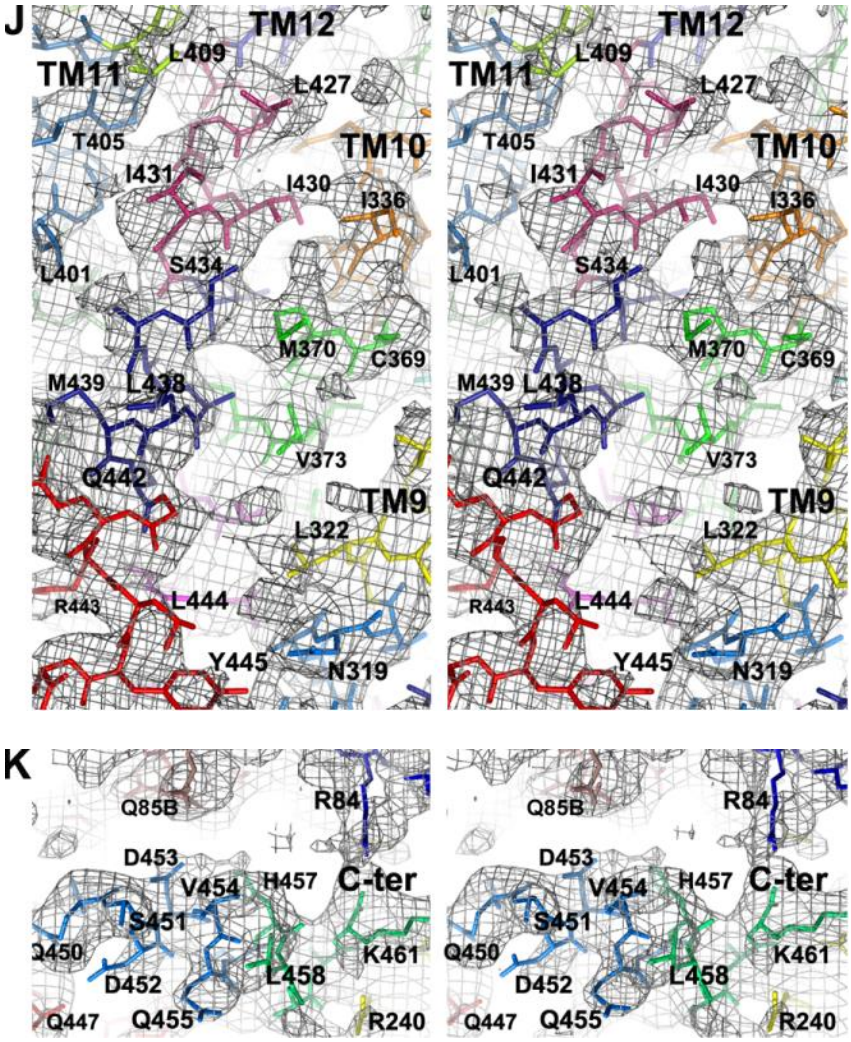



Figure 1.6: (A-J) Stereo views of Fo-Fc simulated annealing omit density for TM helices of NorM_VC (Crystal 1) with multiple density maps superimposed, continued. (K) Stereo view of the C-terminal helix following TM12.

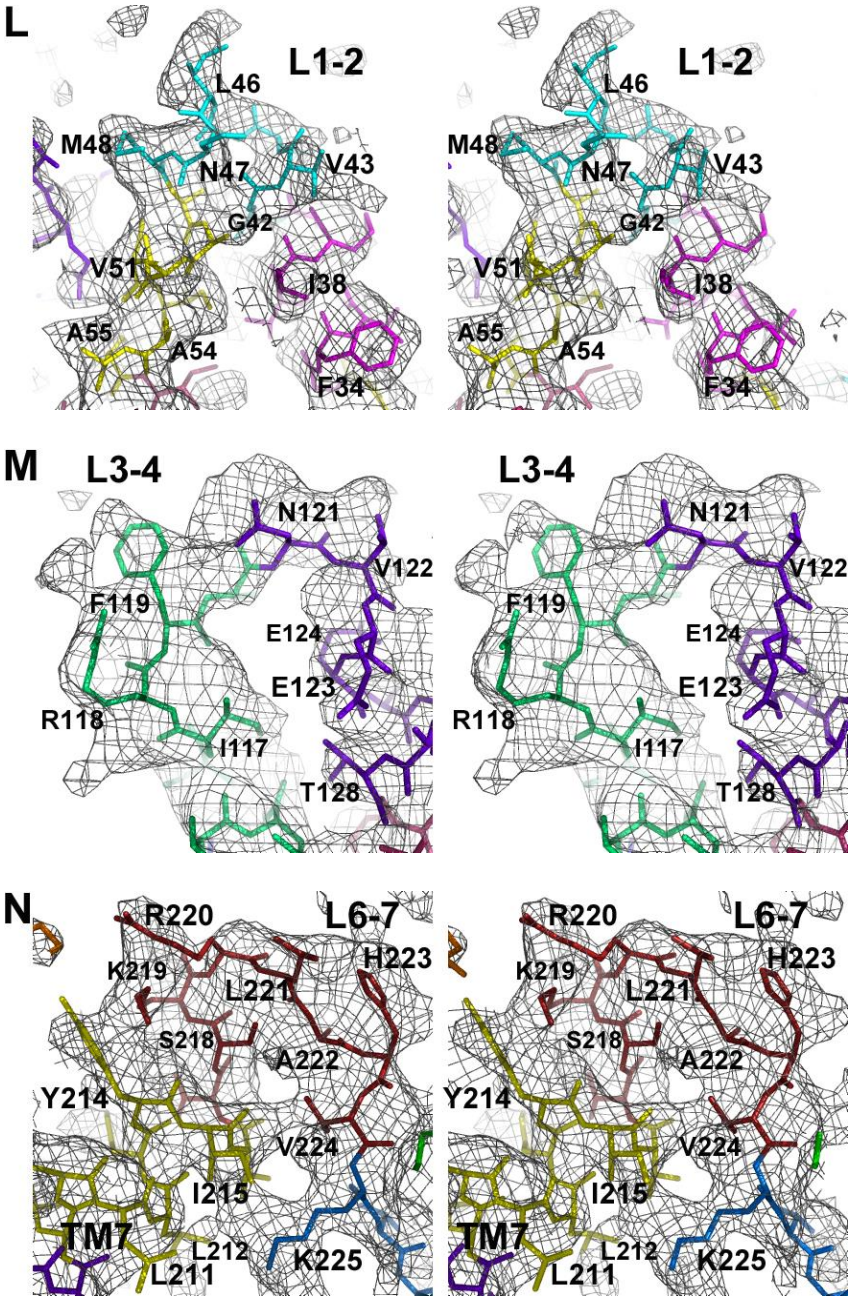



Figure 1.6: (L-O) Stereo views of select loop regions, continued..

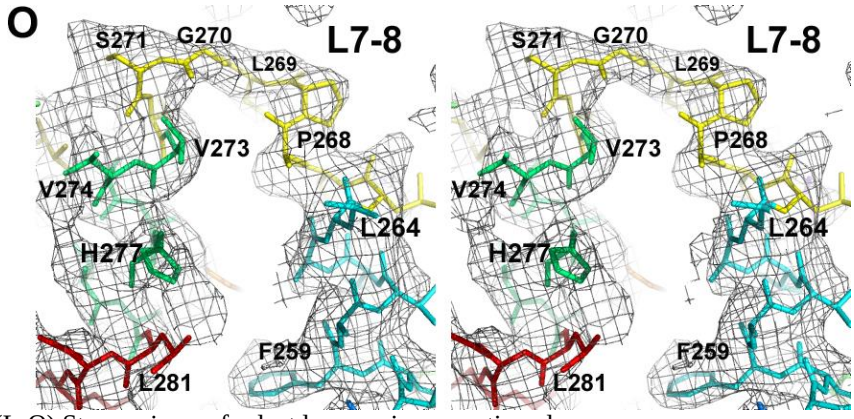



Figure 1.6: (L-O) Stereo views of select loop regions, continued.

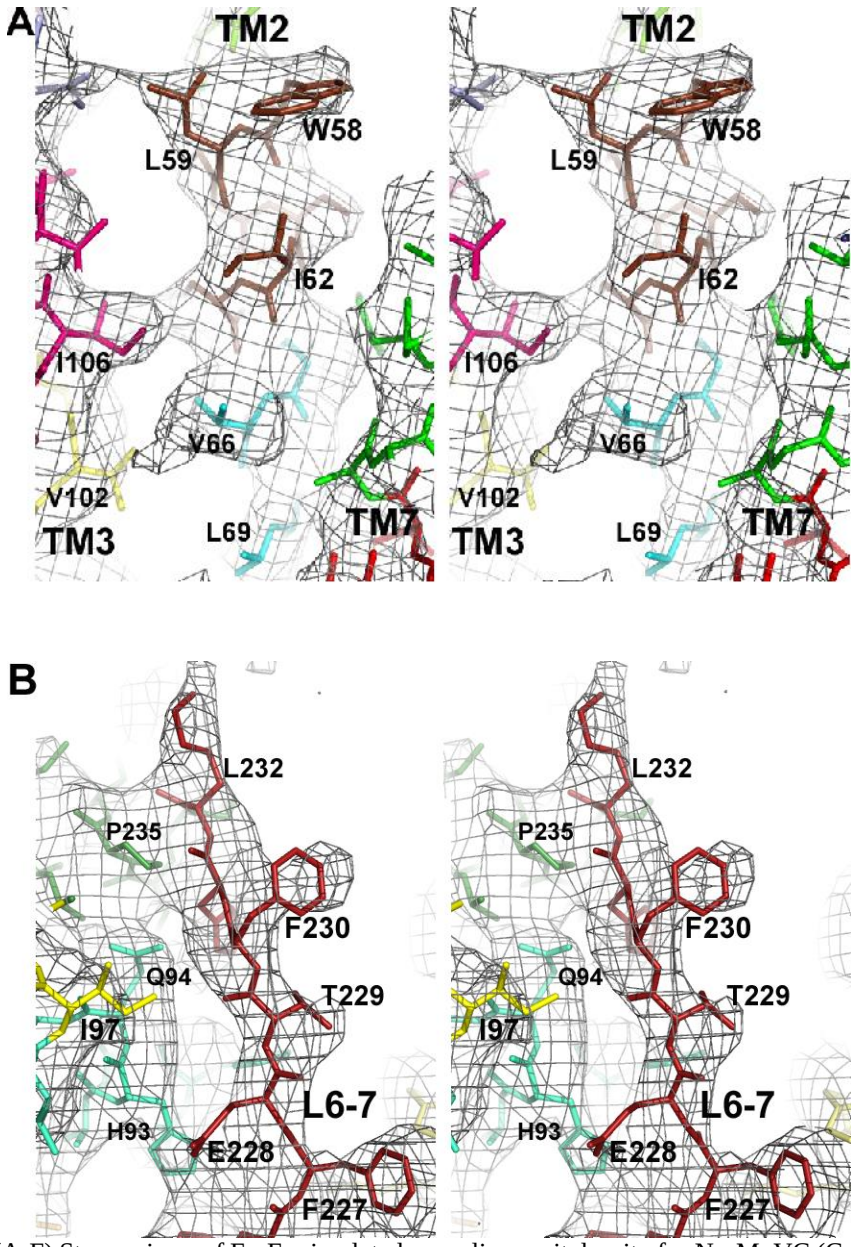



Figure 1.7: (A-F) Stereo views of Fo-Fc simulated annealing omit density for NorM_VC (Crystal 2) with multiple density maps superimposed. The Fo-Fc density maps (grey) were generated in the same manner as Fig. 1.6. The windows of residues omitted from the model are shown in different colors. Maps (A-F) were contoured at 2.0σ.

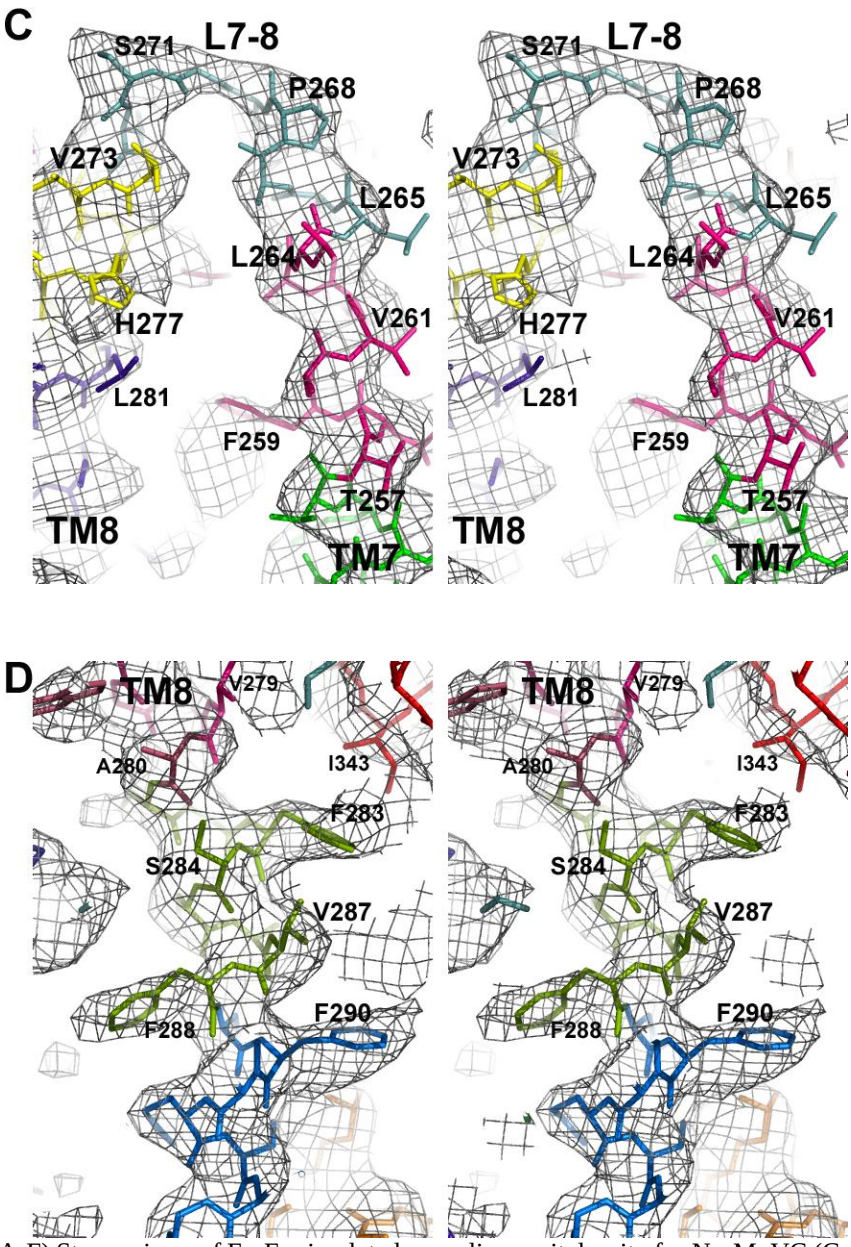



Figure 1.7: (A-F) Stereo views of Fo-Fc simulated annealing omit density for NorM_VC (Crystal 2) with multiple density maps superimposed, continued. The Fo-Fc density maps (grey) were generated in the same manner as Fig. 1.6. The windows of residues omitted from the model are shown in different colors. Maps (A-F) were contoured at 2.0σ.

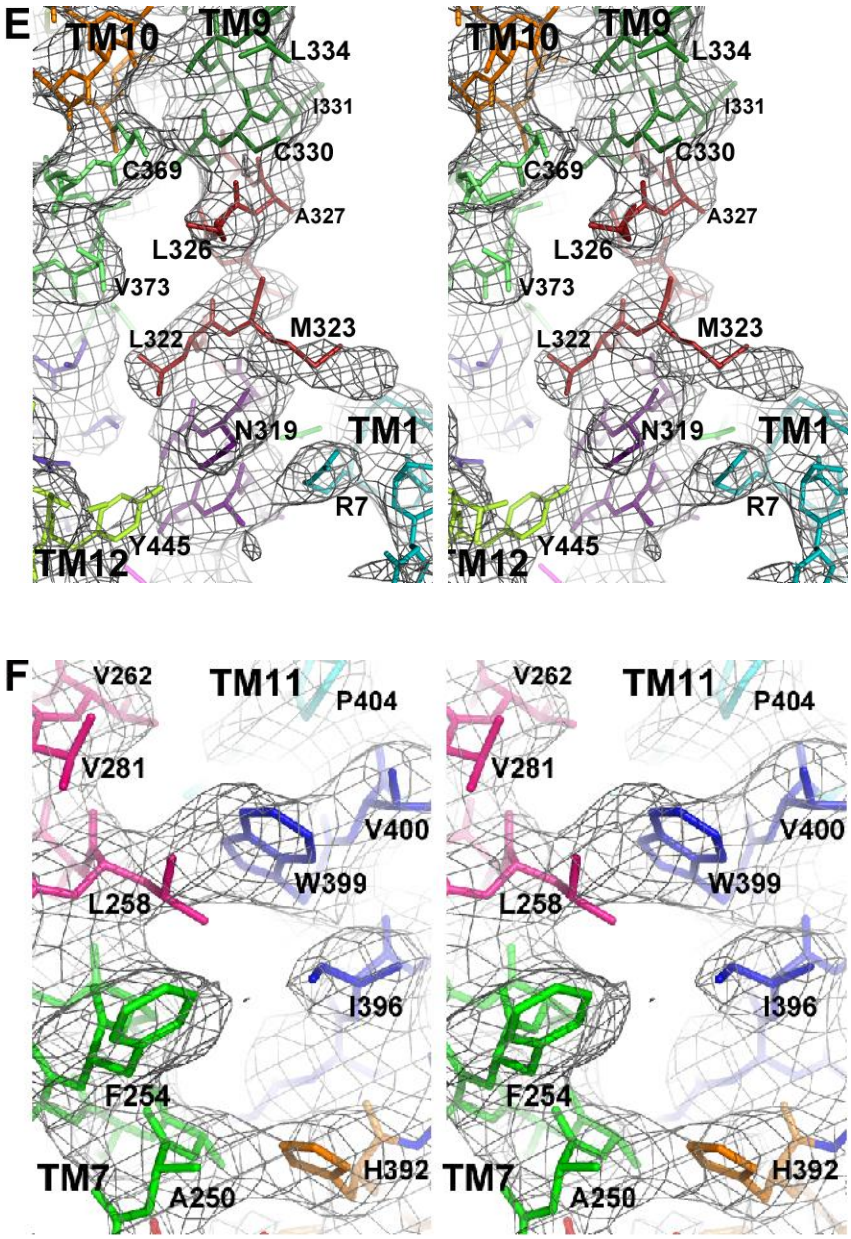



Figure 1.7: (A-F) Stereo views of Fo-Fc simulated annealing omit density for NorM_VC (Crystal 2) with multiple density maps superimposed, continued. The Fo-Fc density maps (grey) were generated in the same manner as Fig. 1.6. The windows of residues omitted from the model are shown in different colors. Maps (A-F) were contoured at 2.0σ

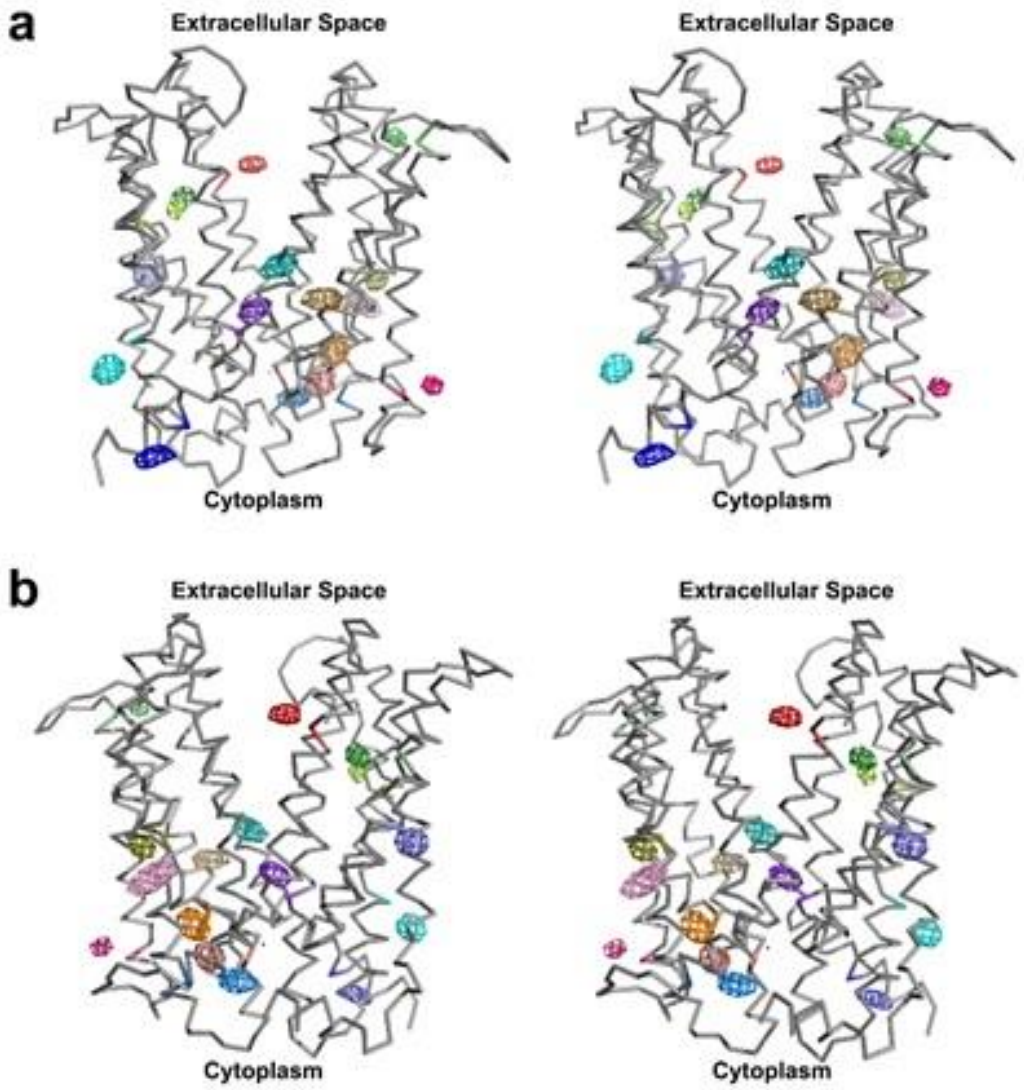



Figure 1.8: Validation of the NorM_VC model, showing the positions of the 16 mercury-labeled cysteine mutants. a, Stereo views from the front and **b**, back of a NorM_VC monomer (gray backbone) in the asymmetric unit with isomorphous difference Fourier peaks, and the corresponding mutated residues colored as: A149C:5.0σ (orange); A260C:3.5σ (red); A296C:4.5σ (purple blue); E91C:4.0σ (marine); F429C:4.0σ (forest); K10C:4.0σ (blue); L101C:4.5σ (pink); M164C: 4.0σ (sand); M323C:5.5 σ(cyan); S26C:3.5σ (deep teal); S103C:4.0σ (olive); S397C:4.5σ (slate); V76C:7.0σ (wheat); V182C:4.5σ (lime); V216C:5.0σ (hot pink); and Y367C:4.0σ (yellow green). Isomorphous difference Fourier maps were generated using data from each mutant crystal and presented together as described.

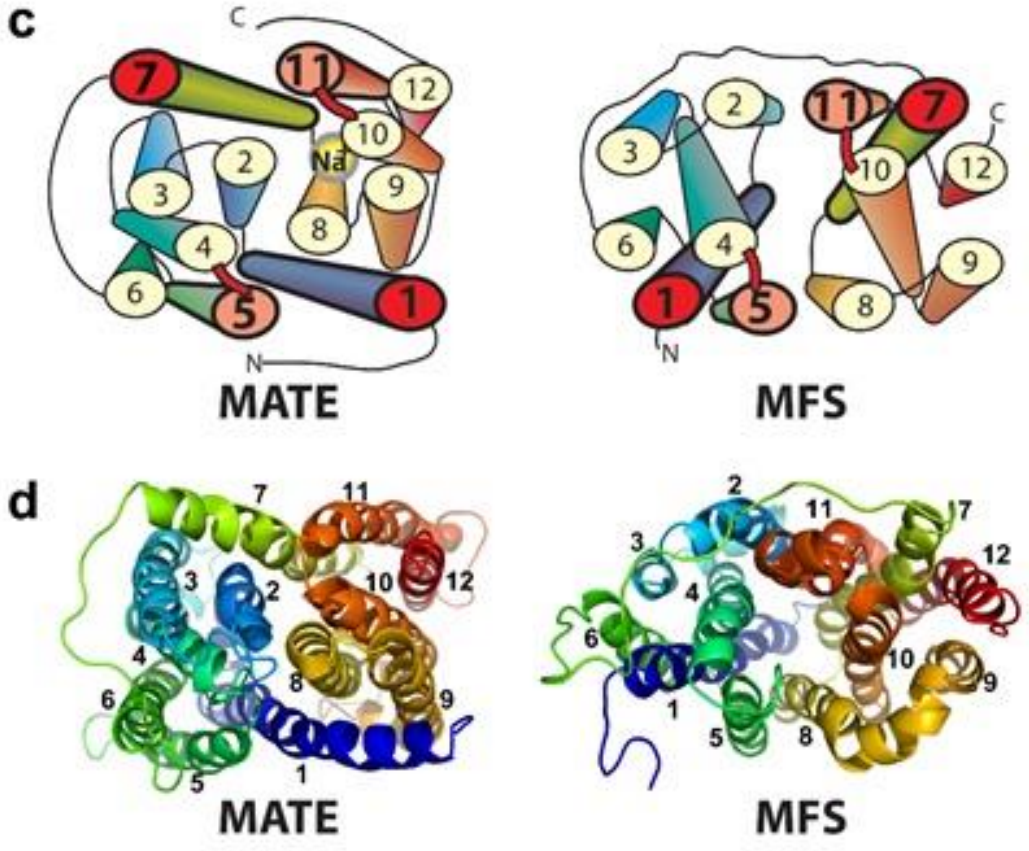



Figure 1.8: Validation of the NorM_VC model, showing the positions of the 16 mercury-labeled cysteine mutants, continued.**c**, Topological comparison between MATE (left) and MFS (right) transporters. The C-terminal helix (residues 450-461) is removed for clarity. The position of the cation ($Na^+$) observed in our MATE structure is indicated. In MATEs, TM1 is between TMs 5 and 8 and TM7 is between TMs 3 and 11. The portals facing the membrane are formed between TMs 1 and 8, and TMs 2 and 7. In contrast, the MFS has TM1 between TMs 5 and 6 and is crossed over by the loop between TMs 4 and 5 (L4-5). Likewise, TM7 is between TMs 11 and 12 and crosses under L10-11. d, X-ray structures of NorM, and EmrD(Yin et al., 2006), as an example of the MFS.

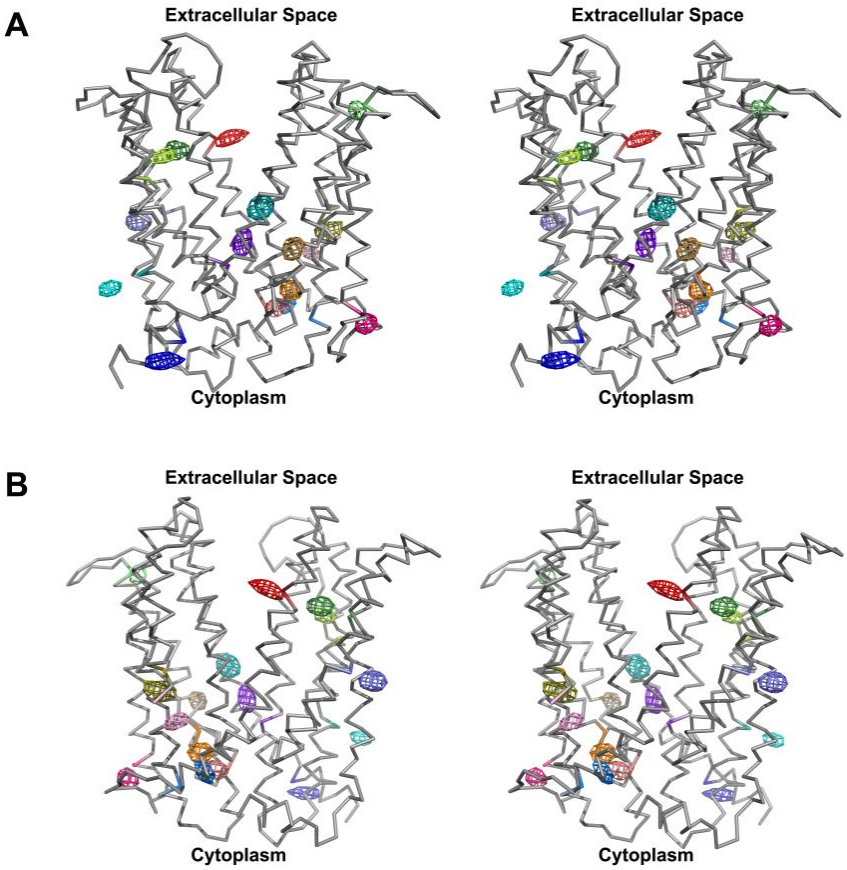



Figure 1.9: Topology verification of NorM_VC structure. Stereo view of NorM1 complementary to NorM2 (shown in Fig. 1.8A-1.8B). (A) Front view showing 16 mercury peaks corresponding to single cysteine mutations as viewed from membrane plane. (B) Back view rotated 180° about the pseudo molecular 2-fold axis relating the two halves of the molecule. Difference Fourier maps were obtained as described for Fig. 1.8A-1.8B from single cysteine mutant crystals soaked with mercury compounds (Table 1.4 and 1.5). The NorM_VC model is rendered as grey ribbon. Each isomorphous difference Fourier peak represents Hg covalently linked to a cysteine residue introduced by single-site mutation. The mutation positions and sigma cutoffs of the difference maps are: A149C: 5.0σ (orange); A260C: 4.0σ (red); A296C: 4.5σ (purple blue); E91C: 4.0σ (marine); F429C: 4.0σ (forest); K10C: 3.5σ (blue); L101C: 4.0σ (pink); M164C: 5.0σ (sand); M323C: 5.5σ (cyan); S26C: 4.0σ (deep teal); S103C: 4.5σ (olive); S397C: 5.0σ (slate); V76C: 7.0σ (deep salmon); V182C: 4.5σ (lime); V216C: 6.0σ (hotpink); Y367C: 5.0σ (yellow green).

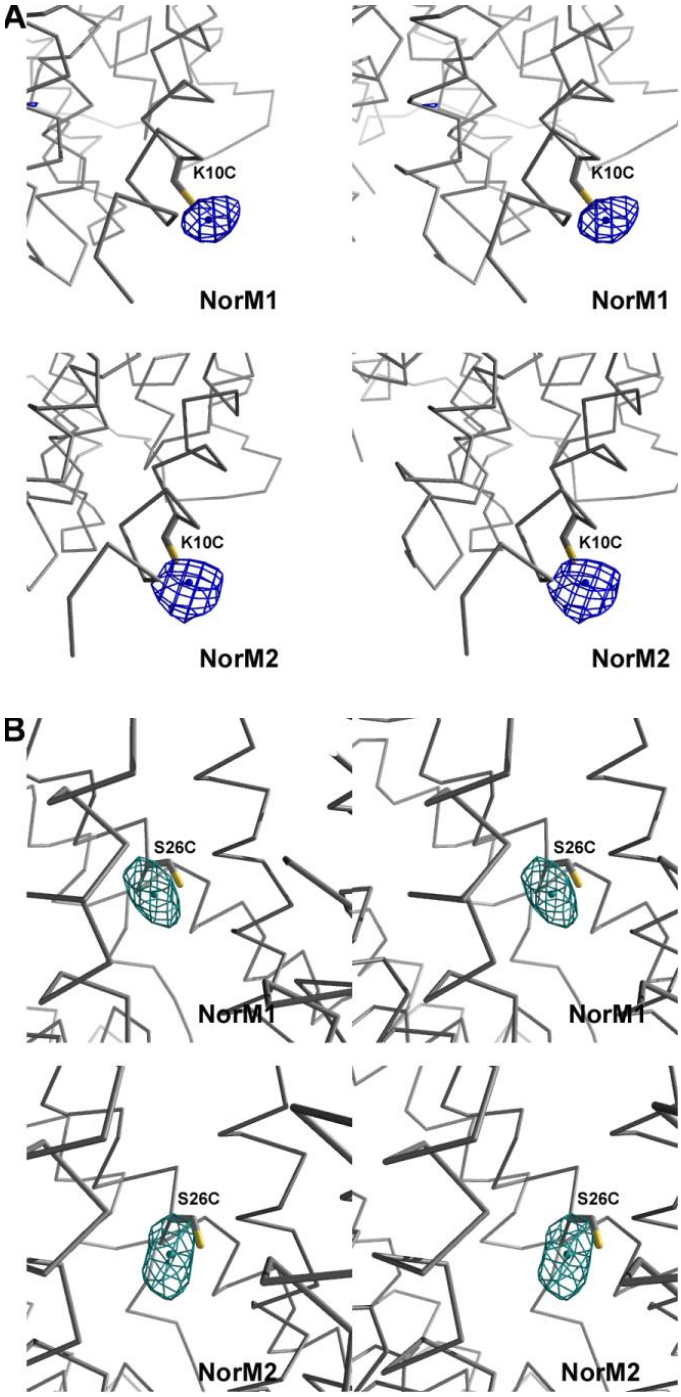



Figure 1.10: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1.9). The relative positions of the Hg sites for both NorM1 and NorM2 are shown, validating both the position and identity of each residue for each monomer in the asymmetric unit. Fourier syntheses were generated as described, and statistics are presented in Table 1.4. The mutation positions and sigma cutoffs of the difference maps for both monomers are the following. (A) K10C, 3.5σ for NorM1 and 4.0σ for NorM2. (B) S26C, 4.0σ NorM1 and 3.5σ for NorM2 Mesh coloring is the same as in Fig. 1.8A-1.8B and Fig. 1.9.

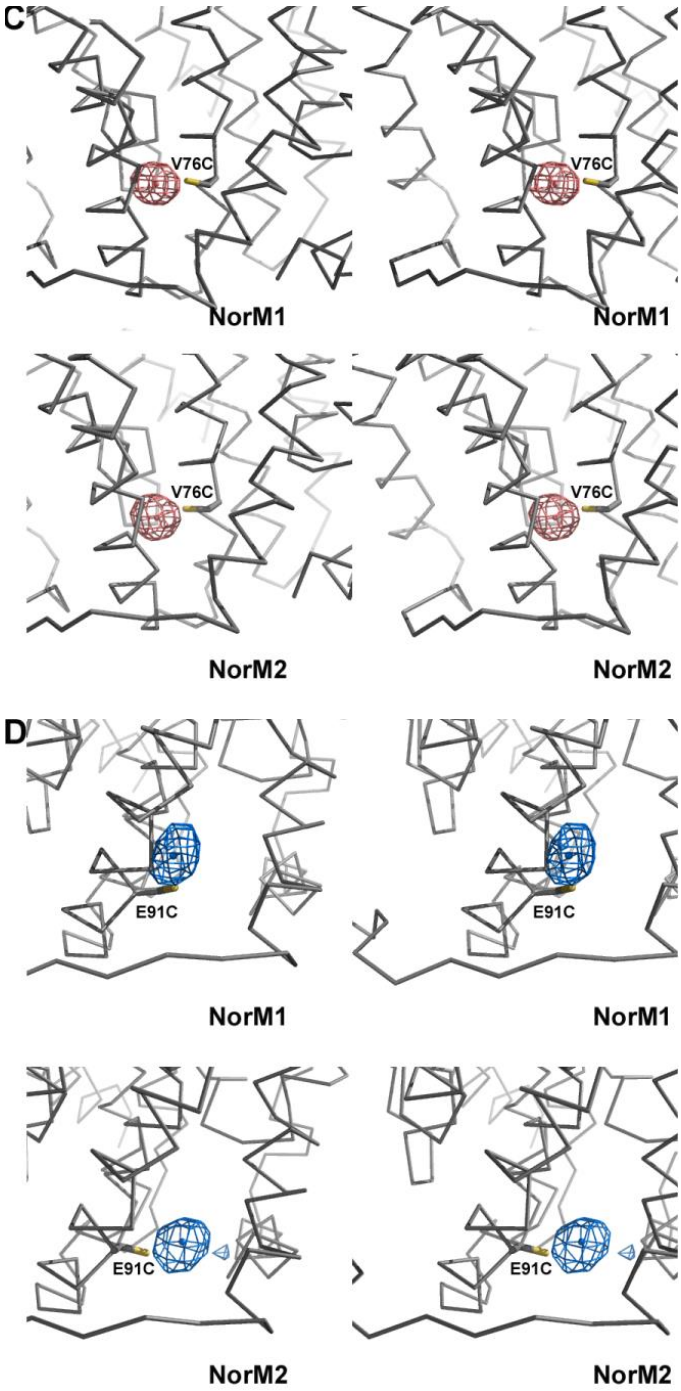



Figure 1.10: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1.9), continued. (C) V76C, 7.0σ. (D) E91C, 4.0σ.

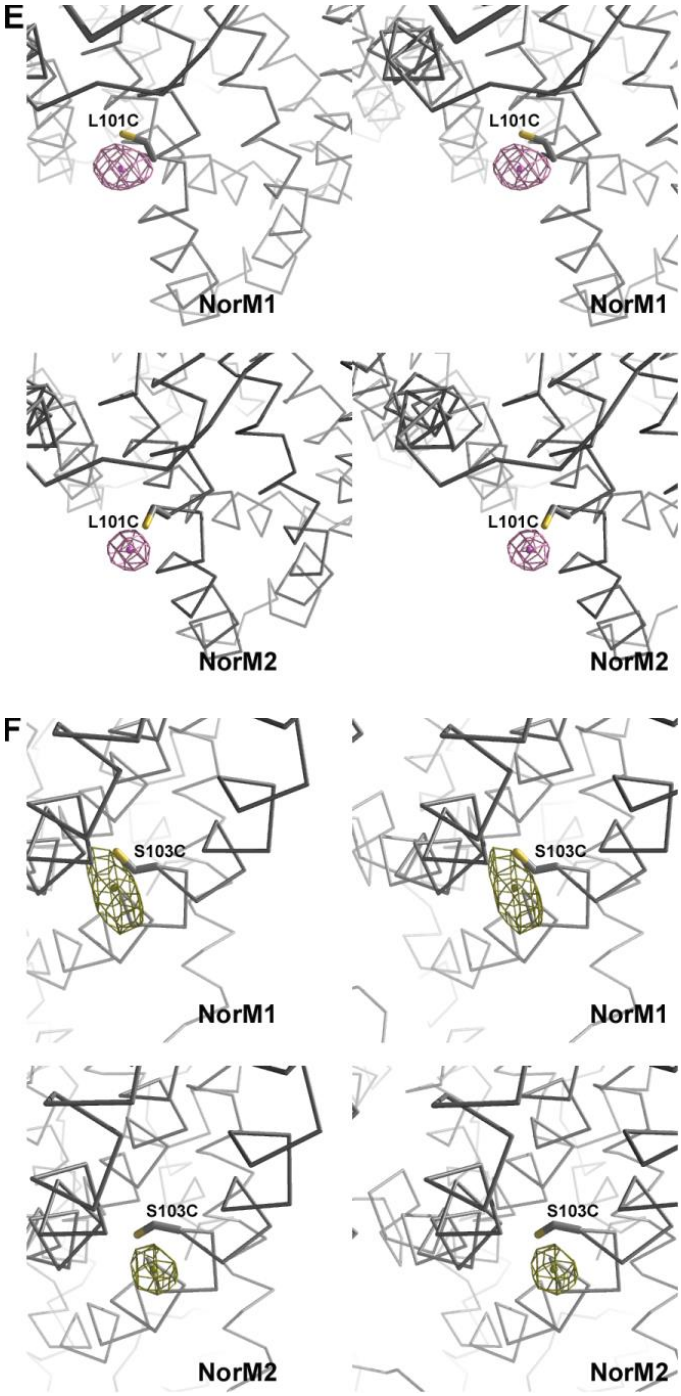



Figure 1.10: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1.9), continued. (E) L101C, 4.0σ for NorM1 and 4.5σ for NorM2. (F) S103C, 4.5σ for NorM1 and 4.0σ for NorM2.

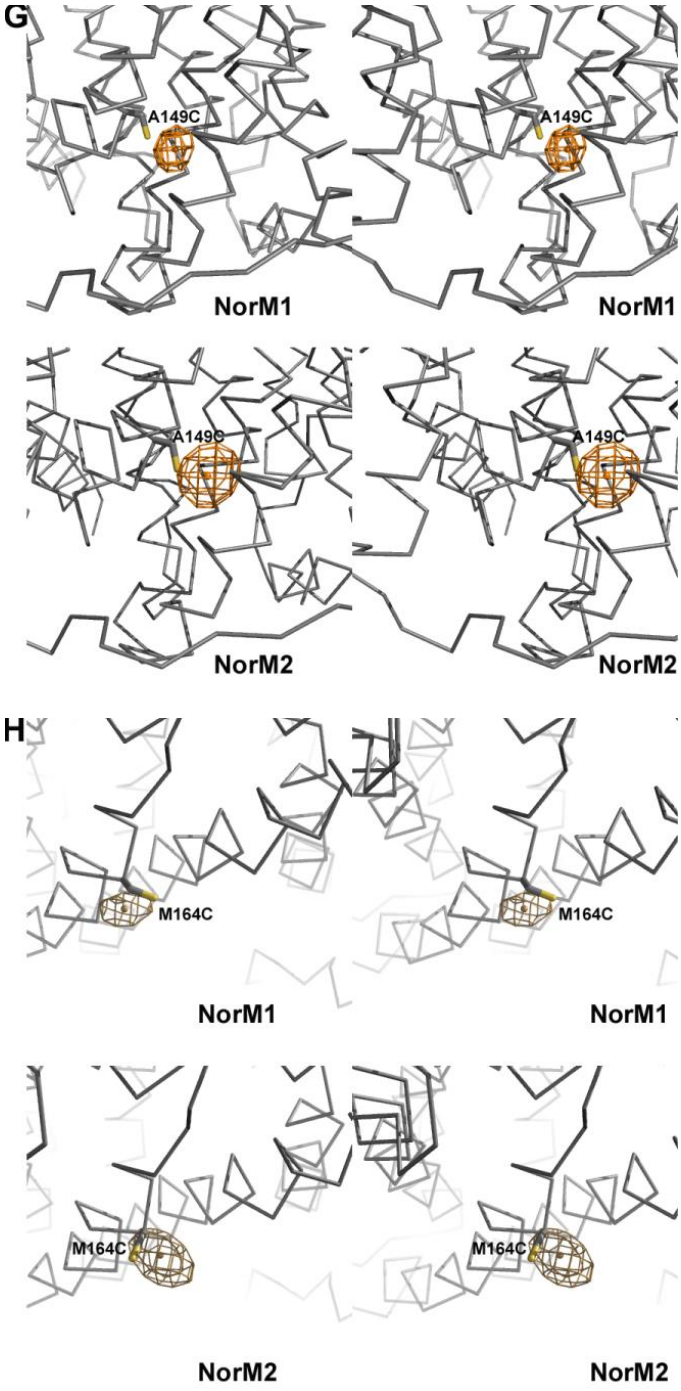



Figure 1.10: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1.9), continued. (G) A149C, 5.0σ. (H) M164C, 5.0σ for NorM1 and 4.0σ for NorM2.

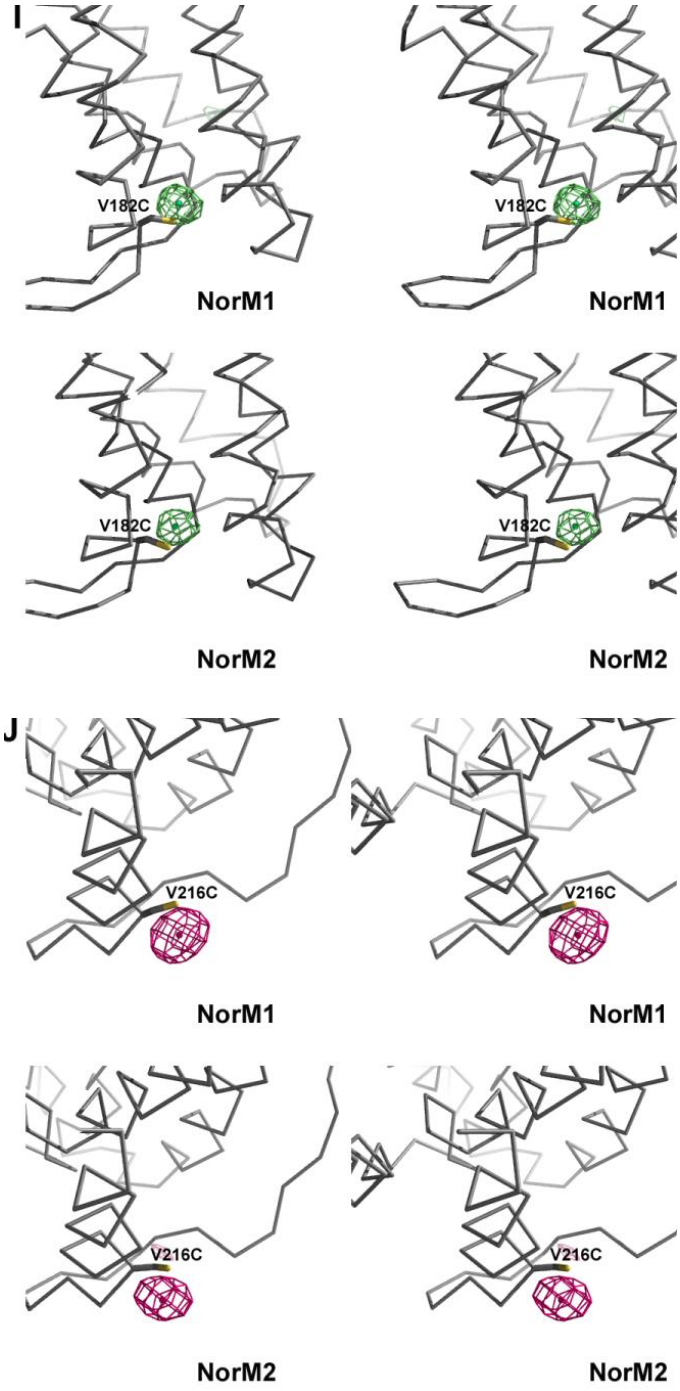



Figure 1.10: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1.9), continued. (I) V182C, 4.5σ. (J), V216C, 6.0σ for NorM1 and 5.0σ for NorM2.

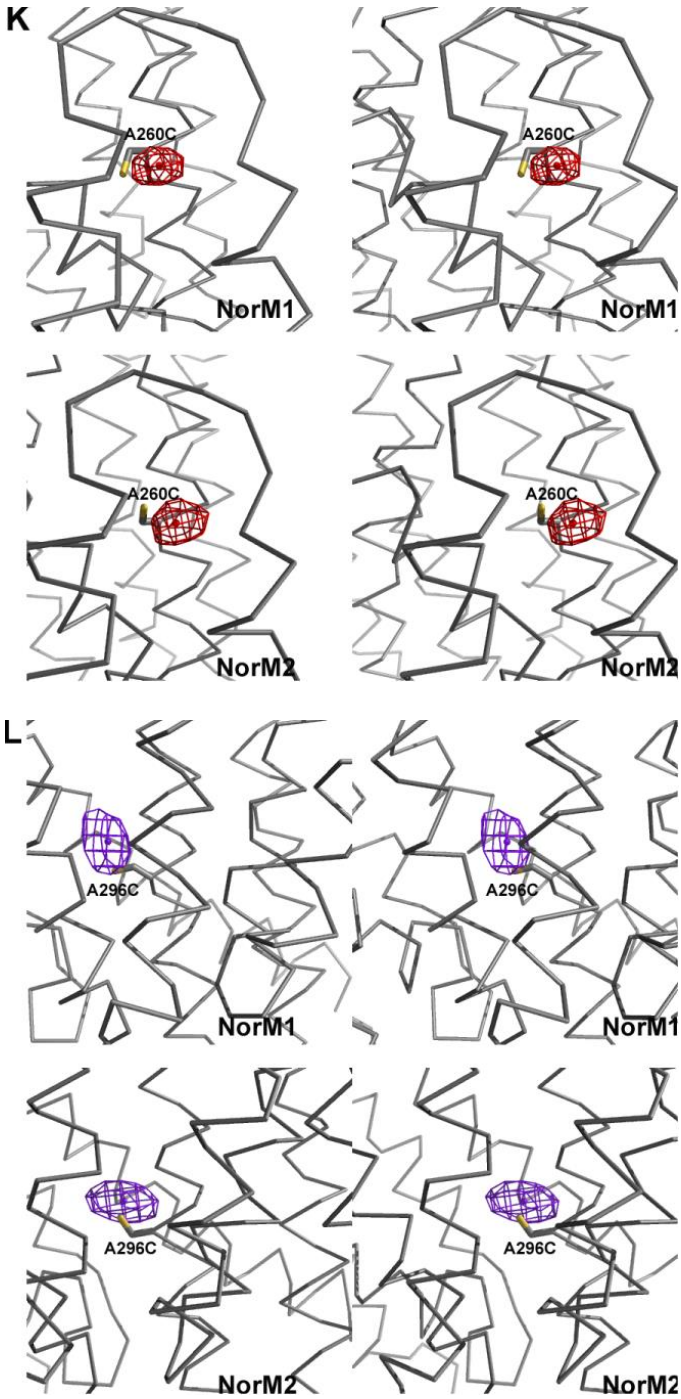



Figure 1.10: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1.9), continued. (K) A260C, 4.0σ for NorM1 and 3.5σ for NorM2. (L) A296C, 4.5σ.

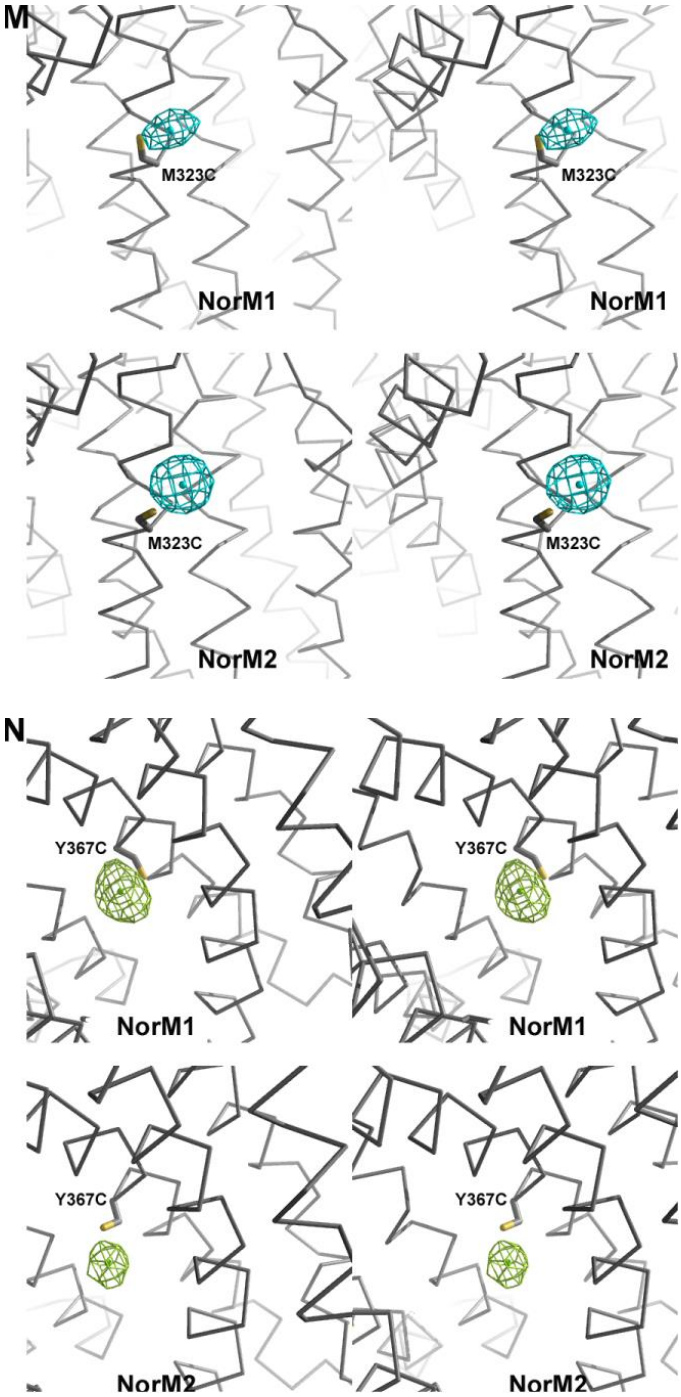



Figure 1.10: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1.9), continued. (M) M323C, 5.5σ. (N) Y367C, 5.0σ for NorM1 and 4.0σ for NorM2.

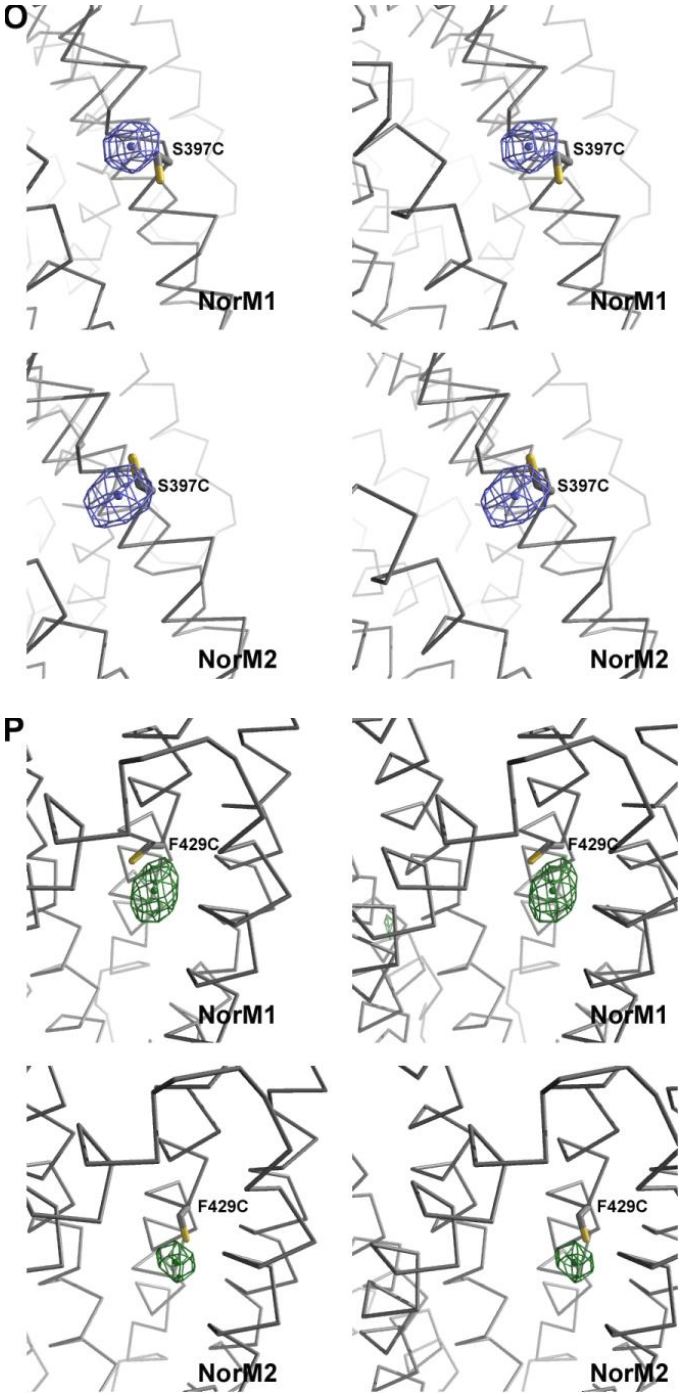



Figure 1.10: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1.9), continued. (O) S397C, 5.0σ for NorM1 and 4.5σ for NorM2. (P) F429C, 4.0σ.

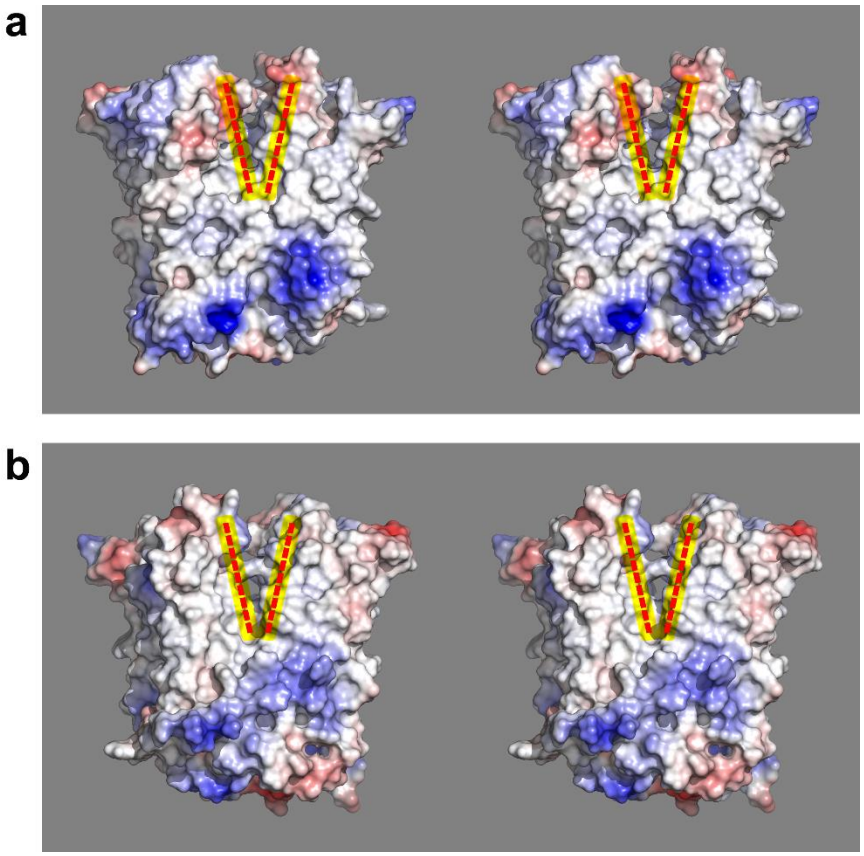



Figure 1.11: Electrostatic potential surface representation of NorM_VC. a,. Stereo views of the portals formed by TMs 1 and 8 and b, TMs 2 and 7, opening to the outer membrane leaflet highlighted and marked by dashed lines (yellow). The same views are shown in Figs. 1.4A and 1.4B, respectively. The surface of NorM is colored according to amino acid residue charge ranging from blue (positive) to red (negative). Hydrophobic and aromatic residues are colored in white.

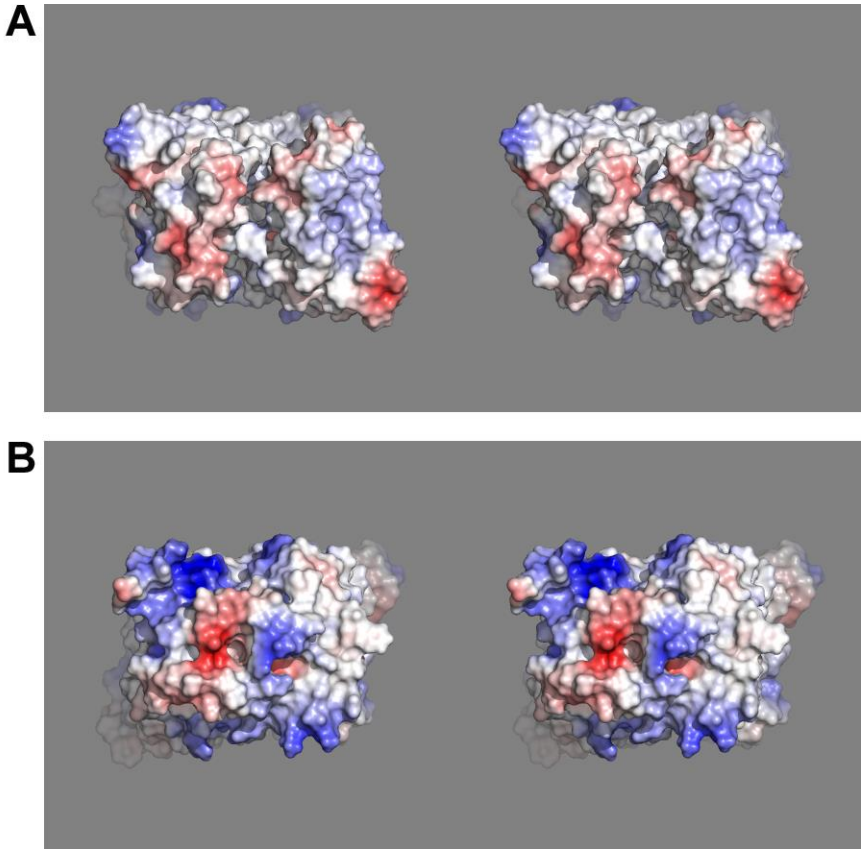



Figure 1.12: Stereo view of electrostatic potential surface of NorM_VC from the (A) extracellular and (B) cytoplasmic sides of the membrane. The internal cavity from the cytoplasmic side of the molecule is occluded in this outward-facing conformation. The surface of NorM is colored according to amino acid residue charge ranging from blue (positive charges) to red (negative charges). Hydrophobic and aromatic residues are colored in white.

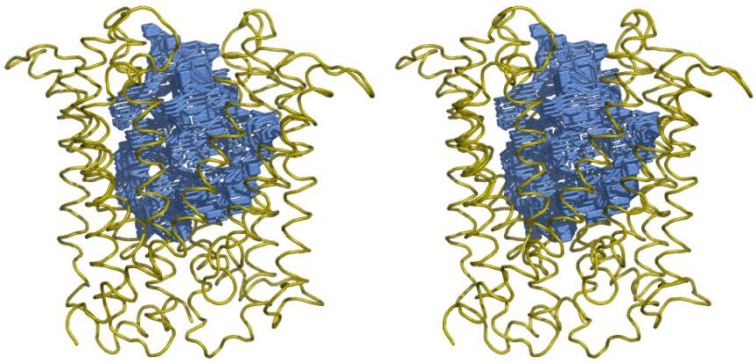
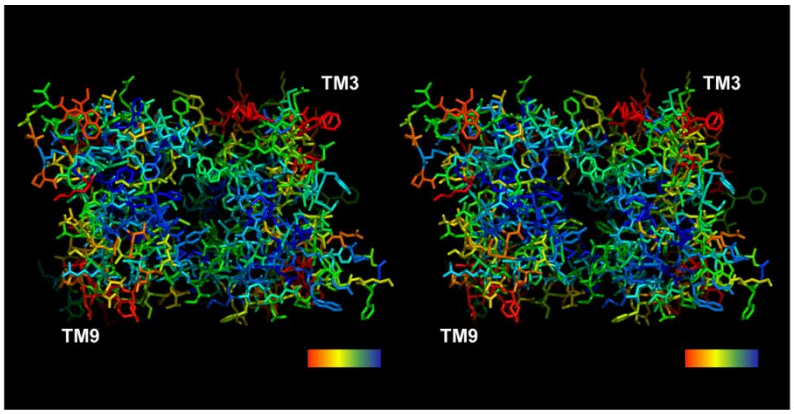


Figure 1.13: NorM_VC internal cavity. The stereo view of the volume of the internal cavity (blue) was generated using the program 3V Channel Extractor with a probe size of 1.6 Å(Voss, Gerstein, Steitz, & Moore, 2006). The Cα trace of the protein is shown as golden ribbon.

A.




B.

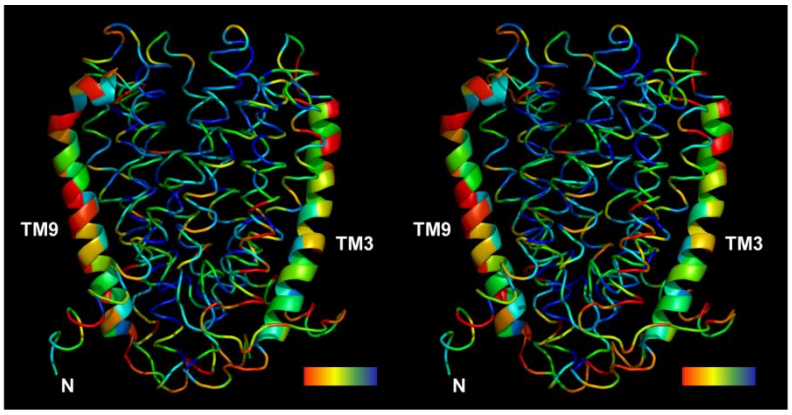



Figure 1.14: Conserved amino acid residues of NorM_VC. (A) Stereo view from the extracellular side showing residue conservation, colored as a rainbow going from most (blue) to least (red) conserved. The most conserved residues are located towards the internal cavity where cation and substrate binding are expected. (B) Side stereo view with TMs 3 and 9 rendered as ribbons for clarity, emphasizing the relatively lower conservation of these helices. The color range is the same as in (A). The analysis was generated using ConSurf (http://consurf.tau.ac.il/) with the MUSCLE option for the sequence alignment and the 50 most recent unique homologous sequences from SWISS-PROT. Three PSI-BLAST iterations were used(Glaser, Rosenberg, Kessel, Pupko, & Ben-Tal, 2005; Goldenberg, Erez, Nimrod, & Ben-Tal, 2009; Landau et al., 2005).

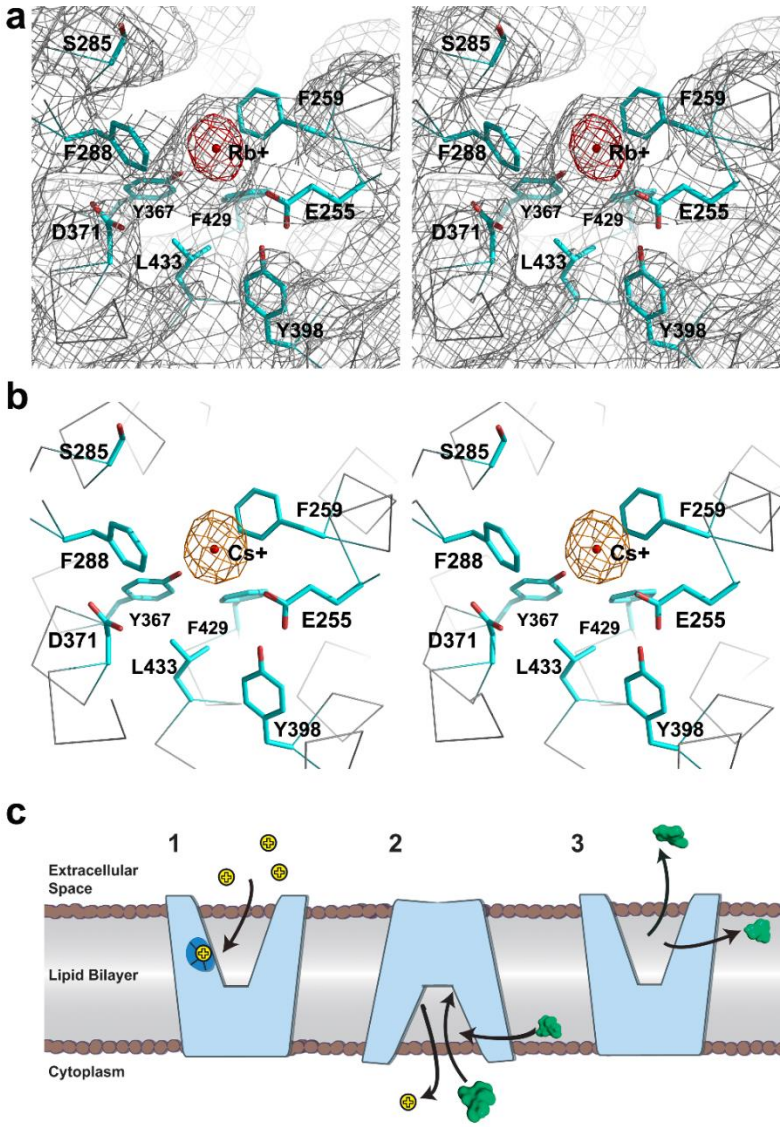



Figure 1.15: The cation-binding site of NorM_VC and mechanism of transport. a, Composite Fo-Fc simulated annealing omit map (gray mesh, contoured at 2.0 σ) of the cation-binding site superimposed with an isomorphous difference Fourier map calculated between Crystal 2 (containing $Rb^+$) and Crystal 3 (native), revealing a peak for $Rb^+$ at 5.5 σ (red mesh) using model phases from Crystal 2 without $Rb^+$. The $Rb^+$ binding site of NorM_VC is composed of all residue centroids within 8.5 Å from the Rb+ site, and include E255, F259, S285, F288, Y367, D371, Y398, F429, and L433 from TMs 7, 8, 10, 11, and 12 (Table 1.5). The composite Fo-Fc omit map was generated as described. b, A 5.0 σ peak (orange mesh) revealing binding of $Cs^+$ in the same view as in a. The isomorphous difference Fourier map was calculated between Crystal 4 (containing $Cs^+$) and Crystal 3 (native). c, Proposed transport mechanism of NorM_VC: In the substrate-free outward-facing conformation, cation (yellow) binds at a structurally and functionally conserved site (blue; step 1). Cation binding induces structural changes to the inward-facing conformation (step 2), which is competent to bind substrate (green) from the inner membrane leaflet or cytoplasm. Substrate binding causes structural changes back to the outward-facing conformation (step 3), allowing substrate export and cation binding.

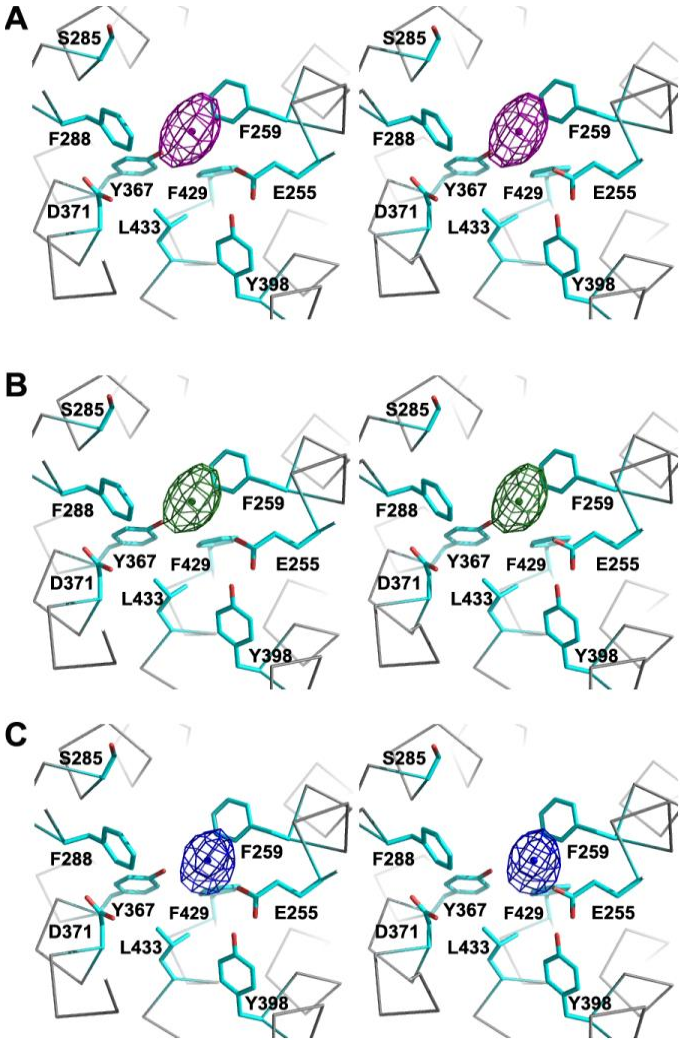



Figure 1.16: Loss of cation-binding as a result of ΔD371N and ΔD371A mutations. (A) Shows a loss of $Rb^+$-binding in the ΔD371N mutant as indicated by a -5.0 σ peak (purple mesh) in the isomorphous difference Fourier map calculated between a RbCl-soaked ΔD371N crystal and Crystal 2, which contains $Rb^+$ (Fig. 1.15A, and Table 1.4). The center of this peak coincides with the Rb+ site in Crystal 2. (B) Shows a loss of $Cs^+$-binding in ΔD371N mutant by a -5.5 σ peak (dark green mesh) in an isomorphous difference Fourier map calculated between a CsCl-soaked ΔD371N crystal and Crystal 4, which contains $Cs^+$ (Fig. 1.15B, and Table 1.4). (C) Shows the loss of $Rb^+$ binding by ΔD371A mutant by a -5.0 σ peak (dark blue mesh) by in an isomorphous difference Fourier map calculated between a RbCl-soaked ΔD371A crystal and Crystal 2. Isomorphous difference Fourier maps are calculated as described. Residues that make up the cation-binding site are marked. Isomorphous difference Fourier maps with native data reveals no peaks for A-C.

Table 1.1: Affinity of NorM_VC to Its Substrates Determined by Fluorescent Polarization

| Compound | $K_d$ or $K_i$, μM | $\lambda_{Ex}$, nm | $\lambda_{Em}$, nm |
|---|---|---|---|
| Doxorubicin | 1.35 ± 0.13 | 485 | 595 |
| Rhodamine 6G | 1.86 ± 0.04 | 485 | 535 |
| TPP † | 2589 * | 485 ** | 535 ** |
| Verapamil | 24 * | 485 *** | 595 *** |

† Tetraphenylphosphonium
* These values were obtained by competitive binding experiments and represent inhibition constant $K_i$.
** The excitation and emission wavelengths are those of Rhodamine 6G; TPP was used as non-fluorescent compound in these conditions.
*** The excitation and emission wavelengths are those of doxorubicin, which was used in competitive binding experiments as a fluorescent substrate; verapamil does not fluoresce in these conditions and was used as non-fluorescent compound.

Table 1.2: Experimental Map Phasing Statistics

| *Data Processing* | | | | | | |
|---|---|---|---|---|---|---|
| Source | CLS BL 08ID-1 | | ALS BL 8.2.2. | | SSRL BL11-1 | |
| Space Group | P $2_12_12$ | | P $2_12_12$ | | P $2_12_12_1$ | |
| Unit Cell | a=163.3 Å, b=242.9 Å, c=45.2 Å, α= β= γ=90.0° | | A=161.0 Å, b=242.5 Å, c=45.9 Å, α= β= γ=90.0° | | a=164.1 Å, b=242.7 Å, c=45.6 Å, α= β= γ=90.0° | |
| | E1 | E2 | E1 | E2 | E1 | E2 |
| Wavelength (Å) | 1.0065 | 1.0099 | 1.0076 | 1.00923 | 0.99826 | 1.00901 |
| Resolution (Å) | 4.6 | 5.2 | 5.6 | 6.4 | 4.3 | 4.5 |
| Redundancy | 5.6 | 5.0 | 5.4 | 4.8 | 2.8 | 3.0 |
| $R_{sym}$[a], % | 0.143 (0.305) | 0.125 (0.304) | 0.075 (0.228) | 0.073 (0.278) | 0.13 (0.334) | 0.125 (0.324) |
| Completeness, % () | 95.6 (85.7) | 94.7 (86.7) | 95.5 (84.8) | 94.9 (85.5) | 89.5 (87.8) | 89.2 (87.8) |

| Phasing and map generation | Combined |
|---|---|
| Overall Phasing Power | 1.7 |
| Figure of Merit | 0.779 |
| Initial NCS Correlation Coefficient | 0.6 |

[a] $R_{sym} = \Sigma|I-\langle I\rangle|\Sigma\langle I\rangle$, where I is the measured intensity of each reflection, and <I> is the intensity averaged from symmetry equivalents.

Table 1.3: Model Refinement Statistics

| Model Building and refinement | Crystal 1 (Native) | Crystal 2 ($Rb^+$-bound) |
|---|---|---|
| Resolution, Å | 3.65 | 4.2 |
| Sigma cutoff | 0.0 | 0.0 |
| $R_{cryst}$[b], %/ $R_{free}$[c], % | 31.2/34.3 | 30.9/34.2 |
| Completeness, % () | 86.5 | 99.0 |
| Rmsd[d] bond lengths, Å | 0.014 | 0.014 |
| Rmsd bond angles, deg | 2.2 | 2.2 |
| <B>, $Å^2$ | 133.6 | 118.0 |
| $R_{shrink}$, $R_{probe}$ (Å) | 0.5, 0.5 | 0.5, 0.5 |
| Bulk Solvent Param: K, $B_{sol}$ ($Å^2$) | 0.25, 50.0 | 0.1, 10.0 |

[b] $R_{cryst} = \Sigma|F_o-F_c|/\Sigma|F_c|$, where $F_o$ and $F_c$ are observed and calculated structure factors, respectively.
[c] $R_{free}$ was calculated from a subset of data (10%) omitted from the refinement.
[d] Rmsd, root mean square deviation.

Table 1.4: Data Processing of Mutants and Additional Data Sets: Summary of mutant, and native datasets used for topological and cation site verification. Superscript denotes compound used for derivation where a=$Hg(OAc)_2$, b= Baker's dimercurial, c= $Hg(CN)_2$, d= RbCl, and e= CsCl. Subscript denotes corresponding dataset used for isomorphous Fourier synthesis.

| Data Set | Source | Unit Cell | | | Resolution (Å) | Redundancy | $R_{sym}$ [a] | Completeness % |
|---|---|---|---|---|---|---|---|---|
| | | a(Å) | b(Å) | c(Å) | | | | |
| 1: ΔK10C[a](11) | APS 23ID-B | 153.0 | 239.7 | 45.6 | 5.5 | 3.6 | 0.093 | 90.7 |
| 2: ΔS26C[a](7) | CLS08ID-1 | 165.1 | 243.1 | 45.7 | 5.2 | 6.7 | 0.133 | 93.1 |
| 3: ΔV76C[a](8) | CLS08ID-1 | 156.9 | 244.2 | 46.2 | 5.8 | 4.0 | 0.119 | 96.6 |
| 4: ΔE91C[a](18) | APS 23ID-B | 144.4 | 241.9 | 46.0 | 4.3 | 3.5 | 0.159 | 91.7 |
| 5: ΔL101C[a](3) | CLS08ID-1 | 159.5 | 244.2 | 46.2 | 4.6 | 6.1 | 0.151 | 93.8 |
| 6: ΔS103C[a](7) | CLS08ID-1 | 160.1 | 243.0 | 45.4 | 4.4 | 3.0 | 0.115 | 88.4 |
| 7: ΔA149C[a](6) | APS 23ID-B | 164.0 | 244.0 | 45.5 | 5.2 | 2.1 | 0.122 | 93.5 |
| 8: ΔM164C[a](3) | APS 23ID-B | 151.4 | 241.8 | 45.6 | 4.3 | 3.3 | 0.127 | 93.5 |
| 9: ΔV182C[a](13) | APS 23ID-B | 159.5 | 243.0 | 46.1 | 4.7 | 2.9 | 0.130 | 90.9 |
| 10: ΔV216C[a](18) | APS 23ID-B | 156.4 | 242.2 | 45.9 | 4.7 | 4.7 | 0.120 | 91.7 |
| 11: ΔA260C[a](1) | APS 23ID-B | 149.0 | 239.1 | 46.0 | 4.6 | 2.8 | 0.154 | 84.2 |
| 12: ΔA296C[b](13) | SSRL 11-1 | 155.1 | 241.8 | 46.0 | 3.8 | 4.8 | 0.107 | 95.7 |
| 13: ΔM323C[c](19) | SSRL11-1 | 159.4 | 243.9 | 46.7 | 4.2 | 3.2 | 0.088 | 94.5 |
| 14: ΔY367C[b](3) | SSRL 11-1 | 154.2 | 244.0 | 46.4 | 4.6 | 3.5 | 0.118 | 91.3 |
| 15: ΔD371A[d](crystal2) | APS 23ID-D | 159.8 | 243.5 | 46.4 | 4.3 | 3.1 | 0.098 | 97.9 |
| 16: ΔD371N[d](crystal2) | APS 23ID-D | 159.1 | 243.1 | 46.4 | 4.6 | 3.4 | 0.126 | 95.2 |
| 17: ΔD371N[e](21) | APS 23ID-D | 153.8 | 237.5 | 45.3 | 4.3 | 3.3 | 0.147 | 97.4 |
| 18: ΔS397C[a](4) | SSRL 11-1 | 153.7 | 241.1 | 46.1 | 5.3 | 2.2 | 0.114 | 83.8 |
| 19: ΔF429C[b](13) | SSRL 11-1 | 159.6 | 242.7 | 46.1 | 4.6 | 3.4 | 0.161 | 93.2 |
| 20: Crystal3 | SSRL 11-1 | 157.8 | 243.4 | 46.7 | 4.0 | 3.3 | 0.106 | 95.1 |
| 21: Crystal4[e] | APS 23ID-D | 154.0 | 239.4 | 45.1 | 4.9 | 3.1 | 0.081 | 98.6 |

$R_{sm}$[a] = Σ(I-<I>)|Σ<I>, where I is the measured intensity of each reflection, and <I> is the intensity averaged from symmetry equivalents.

Table 1.5: Residues in the Internal Cavity: Residues in the internal cavity, where Ala, Ile, Leu, Phe, Trp, Tyr, and Val are classified hydrophobic; Asn, Cys, Gln, Ser, Thr, Tyr Polar; Asp, His, Glu charged; and Phe, Trp, and Tyr aromatic. The $Rb^+$ column denotes residues bounding the cation site.

| | Position | Residue | Hydrophobic | Polar | Charged | Aromatic | $Rb^+$ |
|---|---|---|---|---|---|---|---|
| TM1 | 22 | VAL | X | | | | |
| | 25 | ALA | X | | | | |
| | 26 | SER | | X | | | |
| | 28 | ALA | X | | | | |
| | 29 | GLN | | X | | | |
| | 30 | THR | | X | | | |
| | 32 | MET | | | | | |
| | 33 | GLY | | | | | |
| | 36 | ASP | | | X | | |
| | 37 | THR | | X | | | |
| | 38 | ILE | X | | | | |
| TM2 | 50 | ALA | X | | | | |
| | 51 | VAL | X | | | | |
| | 53 | ILE | X | | | | |
| | 54 | ALA | X | | | | |
| | 55 | ALA | X | | | | |
| | 57 | ILE | X | | | | |
| | 58 | TRP | X | | | X | |
| | 60 | PRO | | | | | |
| | 61 | SER | | X | | | |
| | 63 | LEU | X | | | | |
| | 64 | PHE | X | | | X | |
| | 67 | GLY | | | | | |
| | 68 | LEU | X | | | | |
| | 71 | ALA | X | | | | |
| | 72 | LEU | X | | | | |
| | 75 | VAL | X | | | | |
| TM4 | 135 | HIS | | | X | | |
| | 138 | ILE | X | | | | |
| | 141 | VAL | X | | | | |
| | 142 | PRO | | | | | |
| | 144 | TYR | X | X | | X | |
| | 145 | LEU | X | | | | |
| | 148 | GLN | | X | | | |
| | 152 | SER | | X | | | |
| | 153 | PHE | X | | | X | |
| TM5 | 163 | ALA | X | | | | |
| | 166 | ILE | X | | | | |
| | 167 | GLY | | | | | |
| | 170 | GLY | | | | | |
| | 174 | ASN | | X | | | |
| | 178 | ASN | | X | | | |
| | 182 | VAL | X | | | | |
| TM6 | 196 | CYS | | X | | | |
| | 199 | ALA | X | | | | |
| | 200 | THR | | X | | | |
| | 203 | VAL | X | | | | |
| | 206 | ILE | X | | | | |
| | 207 | MET | | | | | |

| | Position | Residue | Hydrophobic | Polar | Charged | Aromatic | $Rb^+$ |
|---|---|---|---|---|---|---|---|
| TM7 | 248 | VAL | X | | | | |
| | 251 | ALA | X | | | | |
| | 252 | LEU | X | | | | |
| | 253 | PHE | X | | | X | |
| | 255 | GLU | | | X | | X |
| | 256 | VAL | X | | | | |
| | 259 | PHE | X | | | X | X |
| | 260 | ALA | X | | | | |
| | 263 | ALA | X | | | | |
| TM8 | 273 | VAL | X | | | | |
| | 274 | VAL | X | | | | |
| | 276 | ALA | X | | | | |
| | 277 | HIS | | | X | | |
| | 278 | GLN | | X | | | |
| | 280 | ALA | X | | | | |
| | 281 | LEU | X | | | | |
| | 282 | ASN | | X | | | |
| | 285 | SER | | X | | | X |
| | 286 | LEU | X | | | | |
| | 288 | PHE | X | | | X | X |
| | 289 | MET | | | | | |
| | 292 | MET | | | | | |
| | 293 | SER | | X | | | |
| | 295 | GLY | | | | | |
| | 296 | ALA | X | | | | |
| | 299 | SER | | X | | | |
| TM10 | 363 | PHE | X | | | X | |
| | 367 | TYR | X | X | | X | X |
| | 371 | ASP | | | X | | X |
| | 374 | GLN | | X | | | |
| | 375 | VAL | X | | | | |
| | 378 | ALA | X | | | | |
| | 379 | GLY | | | | | |
| TM11 | 391 | PHE | X | | | X | |
| | 395 | PHE | X | | | X | |
| | 398 | TYR | X | X | | X | X |
| | 403 | LEU | X | | | | |
| TM12 | 426 | TRP | X | | | X | |
| | 429 | PHE | X | | | X | X |
| | 433 | LEU | X | | | | X |
| | 436 | ALA | X | | | | |
| | 440 | LEU | X | | | | |
| Total | | | | | | | |
| | Residues | | | | | | |
| | 91 | | 57 | 20 | 5 | 14 | 9 |

Chapter 2:

# Conformational Flexibility and a Novel Epitope of Mouse P-Glycoprotein

## 2.1: Introduction

The mammalian adenosine triphosphate (ATP)-binding cassette (ABC) transporter P-glycoprotein (P-gp) is present in many tissues (I. Cascorbi, 2011), and can detoxify cells by pumping xenobiotics across the membrane (Eckford & Sharom, 2009). A hallmark feature of this transporter is its ability to bind and transport a wide array of structurally diverse molecules ranging in size from 100-4,000 daltons D(a) (MW) (Schinkel & Jonker, 2003). In fact, P-gp was discovered in a cell line where it prevented the permeation of several different drugs (Gottesman & Ling, 2006). While its physiological function is not fully understood, the well-recognized role of P-gp in mediating multidrug resistance (MDR) in many types of cancers has made it a valid therapeutic target (Marquez & Van Bambeke, 2011). P-gp also plays a significant role in the pharmacokinetics and bio-availability of drugs by mediating their transport in the liver, intestines, and across the blood brain barrier (I. Cascorbi, 2011; Eckford & Sharom, 2009; Lee, Cook, Reyner, & Smith, 2010).

P-gp is a ~170 kDa molecule comprised of two pseudo-symmetric halves, each containing a nucleotide-binding domain (NBD) and a transmembrane domain (TMD). Currently, the "alternating access" model is the most widely accepted paradigm explaining the mechanics of transport by ABC transporters (Gutmann, Ward, Urbatsch, Chang, & van Veen, 2010; A. Ward et al., 2007). According to the model, binding of ATP at the NBDs drives conformational changes in the TMDs, and switches the

transporter's overall conformation from inward-facing to outward-facing (inward/outward refer to the opening of the drug-binding pocket relative to the cell). This ATP-driven switch results in the vectorial transport of substrates out of the cell The hydrolysis of ATP and release of Pi/ADP are essential for resetting the transporter back to the inward-facing conformation.

The overall protein topology of the inward-facing conformation of P-gp has been previously established (Aller et al., 2009), and is reiterated in the x-ray structures of related bacterial and mammalian ABC transporters found in the ABCB subfamily (Hohl, Briand, Grutter, & Seeger, 2012b; Jin, Oldham, Zhang, & Chen, 2012; A. Ward et al., 2007), including the human ABCB10 (PDB ID: 3ZDQ, 4AYX, 4AYT, 4AYW) (C. A. Shintre et al., 2013). Furthermore, the inward-facing conformation has been biochemically validated in bacterial ABC exporters, using electron paramagnetic resonance on MsbA (Zou, Bortolus, & McHaourab, 2009) and LmrA (Hellmich et al., 2012), cysteine crosslinking of MsbA (Doshi, Woebking, & van Veen, 2010), and hydrogen/deuterium exchange coupled to mass spectrometry of BmrA (Mehmood, Domene, Forest, & Jault, 2012). The inward-facing conformation is a key intermediate in the alternating access mechanism (Gutmann et al., 2010), as it allows the transporter to scan the inner leaflet of the membrane for substrates. This notion is supported by the inward-facing x-ray structures of P-gp, which have cyclic peptide inhibitors bound to the substrate-binding pocket (Aller et al., 2009).

Owing to its role in MDR, several small molecule inhibitors of P-gp have been developed during the last four decades (Crowley, McDevitt, & Callaghan, 2010;

Eckford & Sharom, 2009), but none have yet been approved for clinical use (Tamaki, Ierano, Szakacs, Robey, & Bates, 2011). In recent years, antibody-derived therapeutics have become increasingly popular because of their specific affinities and relatively high tolerance in humans. Nanobodies (Nbs) are small (~15 kDa) single domain proteins that are derived from, and contain, the unique structural and functional properties of natural heavy chain-only antibodies found in Camelids (Muyldermans et al., 2009). Like conventional antibodies, Nbs achieve antigen-specificity by using three loops called Complementarity Determining Regions (CDRs 1-3).

Nbs are mostly monomeric, encoded by single genes, highly soluble, and can be efficiently produced in prokaryotic and eukaryotic hosts, including bacteria and yeast. They are also capable of penetrating many tissues including the human gut (Muyldermans et al., 2009). These characteristics provide Nbs with great promise as therapeutics against a number of diseases (Muyldermans et al., 2009). Additionally, Nbs can bind to cryptic epitopes and potentially stabilize proteins in certain conformations. These properties make Nbs useful as chaperones for inducing lattice formation during protein crystallization (Lam, Pardon, Korotkov, Hol, & Steyaert, 2009).

Here, we present three new x-ray structures of mouse P-gp (*ABCB1a*) with different inward-facing conformations, all wider than those originally published (Aller et al., 2009). One crystal form comprises P-gp complexed with a single Nb, called Nb592, bound to the first NBD (NBD1). Consistent with its binding location, we show that Nb592 is a potent inhibitor of P-gp's ATP hydrolysis activity. Taken together, these different structures of P-gp reveal a novel epitope and demonstrate the flexibility of the transporter to sample a wide range of inward-facing conformations.

**2.2: Structure Determination**

To structurally map the conformational flexibility of P-gp in the absence of nucleotides and substrates, we determined three x-ray structures (Table 2.1) that crystallized in two different crystal forms (Crystal Forms A and B; Table 2.1). Depending on the crystal growth condition, Crystal Form A yielded crystals that had variations in unit cell dimensions (Table 2.1) with changes mostly in the c- and a- cell edges (Crystal 1 and Crystal 2). To reduce bias, we determined an experimentally derived electron density map of Crystal Form A by generating a heavy atom derivative using ethyl-mercury-chloride (Fig. 2.1, labeled "Experimental Map Crystal") and obtained protein phases using the multiple anomalous dispersion technique (MAD). An initial model was built using this electron density map to 4.5 Å (Fig. 2.1). A higher resolution native crystal (Crystal 1; Fig. 2.2; Table 2.1) later diffracted to 3.8 Å, and was solved by the molecular replacement method. The structure revealed that the distance between the NBDs, as measured by the C□ atom of residues 626 (C-term of NBD1) and 1271 (C-term of NBD2), is ~31 Å, which is much larger than the ~13 Å observed in the previously published mouse P-gp structure (Aller et al., 2009). This represents a large (~18 Å) range of displacement between the NBDs that P-gp can conformationally sample while in the inward-facing state.

A second native crystal of P-gp (Crystal 2) diffracted to 4.0 Å and was determined by molecular replacement using our model derived from Crystal 1. This structure of P-gp is the widest, with a distance of ~36 Å between residues 626 and 1271. Crystals 1 and 2 represent the more extreme ranges of the variations observed for Crystal Form A. The two x-ray structures revealed distinct inward-facing conformations of P-

gp, with the overall structure similar to the one published (Aller et al., 2009) and some smaller local changes. For example, a break in the helical structure of transmembrane helix (TM) 12 in the original more closed inward-facing conformation is fully helical in these structures. The linker (residues 627-683) between NBD1 and TM7 is not observed and presumed to be disordered. Residues 1-33 and 1272-1276 at the N- and C- termini- are also not observed.

### 2.3: Validation of the Structure

To further validate the models of P-gp (Crystals 1-2) derived from Crystal Form A, which is critical at these moderate diffraction resolutions (3-4.5 Å), we made 17 single-site cysteine substitution mutants in P-gp and labeled them with ethyl-mercury-chloride. These mutants included residues S80C, S176C, M188C, A216C, S218C, A250C, R272C, L283C, G284C, S305C, A309C, G342C, A344C, K730C, S876C, S889C, and S975C distributed throughout the TM helices. Figure 2.2 shows a composite view of all experimentally validated positions mapped on to the Crystal 1 model while Fig. 2.3 provides close-up shots of each mercury-labeled single-site mutant. Combined with 7 wild-type cysteine residues of mouse P-gp previously identified (Aller et al., 2009) and observed in these data, we were able to experimentally validate the accuracy of our model using a total of 24 positions.

### 2.4: Nanobody-Bound Structure of P-gp

A second, entirely different crystal form (Crystal Form B) of P-gp was obtained by co-crystallization of mouse P-gp with the nanobody Nb592 (Fig. 2.4; Table 2.1). Size

exclusion chromatography demonstrated that the P-gp and Nb592 form a complex that co-migrates and elutes as a single peak (Fig. 2.5). Fractions from this sample peak were pooled together, concentrated, and crystallized. The x-ray structure of the P-gp-Nb592 complex (Crystal 3) was determined to a resolution of 4.1 Å (Table 2.1) by molecular replacement using the model derived from Crystal 1 and a model of Nb592 generated as described in the Material and Methods section. The CDR loops of Nb592 were found to bind NBD1, thereby potentially precluding the NBDs from dimerizing. This inward-facing conformation is the least wide among the three structures presented here, with a distance of ~30 Å between the C□ atoms of residues 626 and 1271.

The interface between Nb592 and NBD1 of P-gp is significant and constitutes ~745 $A^2$ of the buried surface area. CDR3 (residues 98-107 of Nb592) penetrates into a pocket formed by 3 helices in NBD1 (Fig. 2.4). The P-gp residues within 5 Å of Nb592 include S555, T559, E562, A563, H583, R584, L585, S586, T587, R589, H608, M612, F619, L621, V622, M623, T624, Q625, and T626 on NBD1 and K1260, F1264, S1265, and S1268 on NBD2 (Fig. 2.6A). T626 is the last resolved residue at the C-terminus of NBD1 and is proceeding a flexible linker region presumed disordered but could also potentially interact with Nb592.

The Pgp-Nb592 structure (Crystal 3) elucidates a previously unknown epitope on the NBDs of P-gp. A protein sequence alignment between mouse and human P-gp reveals that the two proteins are very similar in the epitope/Nb592-binding site (Fig. 2.6B). The only differences are at residues 623 and 624 (mouse numbering); mouse: Met-Thr, human: Thr-Met. While these residues are within 5 Å of Nb592, T624 is pointing away from Nb592 and M623 is oriented toward the conserved portion of the

Nb just before and after CDR2 (residues 551-557 of Nb592). Interestingly, Nb592 does not bind to mouse NBD2, which also has a high protein sequence similarity in this region (Fig. 2.6B).

**2.5: Inhibition of P-gp ATPase by Nb592**

The NBDs of ABC transporters are responsible for the binding and hydrolysis of ATP that drives drug transport through the TMDs. The structure of P-gp with Nb592 bound at NBD1 (Fig. 2.4) indicates that this Nb could significantly affect the catalytic function of P-gp. We tested this hypothesis using two biochemical approaches. First, we measured the drug (verapamil)-stimulated ATPase activity of purified mouse P-gp, in the presence of increasing concentrations of Nb592. Our results revealed that Nb592 is indeed a strong inhibitor of ATP hydrolysis with half-maximal inhibition achieved in the nanomolar range ($IC_{50}$ of 520 ± 57 nM) (Fig. 2.7A). This $IC_{50}$ value is comparable to or lower than those reported for most small molecule P-gp inhibitors (Aller et al., 2009; Crowley et al., 2010; Eckford & Sharom, 2009; Shukla, Ohnuma, & Ambudkar, 2011). Nb592 also potently inhibited basal ATPase activity (in the absence of drug) with 90 ± 5% inhibition seen at saturating amounts of 7.5 μM Nb592. More so, the inhibition was fully retained in the presence of 5 mM DTT (Fig. 2.7B), indicating that Nb592 is a robust protein that can withstand reducing agents at concentrations that are typical of the cytoplasm.

We made further measurements suggesting that Nb592 prevents dimerization of the NBDs. ATP hydrolysis requires intimate interactions of residues from both *cis*- and *trans*-NBDs with bound ATP to facilitate the hydrolytic attack on the γ-phosphate

(Seeger & van Veen, 2009). Sodium orthovanadate (Vi) is an inorganic phosphate (Pi)-analogue that can trap ATP in the post-hydrolysis state (ADP•Pi) (J. Chen, Sharma, Quiocho, & Davidson, 2001), resulting in an outward-facing conformation with dimerized NBDs (A. Ward et al., 2007). Using 8-azido-[$\alpha^{32}$P]-ATP, we show that 8-azido-[$\alpha^{32}$P]-ADP is trapped in mouse P-gp by Vi, in the presence of the ATP hydrolysis activator, verapamil (Lanes 1-3, Fig. 2.7c) (Urbatsch, Sankaran, Bhagat, & Senior, 1995). At increasing concentrations of Nb592, the hydrolysis of ATP is effectively inhibited as indicated by the reduced amount of trapped 8-azido-[$\alpha^{32}$P]-ADP in the catalytic sites of P-gp (Lanes 4-10, Fig. 2.7c).

Taken together, both biochemical experiments show that the binding of Nb592 to mouse P-gp inhibits its ATP hydrolysis activity by hindering the formation of an ATP hydrolysis-competent NBD dimer sandwich. These results complement the structure of the Pgp-Nb592 complex (Fig. 2.4), suggesting a molecular basis for the inhibition of ATPase activity. This new mechanism of inhibiting P-gp by specifically targeting the NBDs contrasts the mode of inhibition by several small molecule inhibitors/drugs, which target the substrate-binding cavity in the TMDs (Aller et al., 2009; Crowley et al., 2010; Eckford & Sharom, 2009).

### 2.6: Context of the Novel Structures of P-gp

The overall topology and protein fold of the three new P-gp x-ray structures described in this work are similar to those previously published (Aller et al., 2009), with some small localized differences. When the N-terminal halves of the P-gp models in this study were structurally aligned (residues 33-209, 852-961, and 320-626), the C-

terminal halves showed a significant displacement relative to one another; most notable in the relative positions of NBD2 (Fig. 2.8). The overall change in the structures can be described as a hinge movement where a small angular change at the pivot point causes larger changes farther away towards the NBDs. The movement near the pivot or "TM-hinge region" on the extracellular side is smaller, and comprises Loop 3-4 (L3-4; residues 206-208) and Loop 5-6 (L5-6; residues 319-324), as well as the corresponding Loop 9-10 (L9-10; residues 849-851) and Loop 10-11 (L10-11; residues 963-968) on the other half of the molecule.

The most striking feature of these P-gp structures is a much larger separation between NBD1 and NBD2, suggesting that the transporter can adopt a much wider inward-facing conformation than previously described. The overall range of distances between residues 626 (C-term of NBD1) to 1271 (C-term of NBD2) sampled by the three inward-facing conformations presented here is ~29-36 Å, while the corresponding distance in the original published structure (Aller et al., 2009) is only ~13 Å. Such dynamic conformational flexibility in the inward-facing state has also been previously observed in biochemical, biophysical, and molecular dynamics simulations experiments on bacterial ABC transporters, such as MsbA (Zou et al., 2009) and BmrA (Mehmood et al., 2012), and very recently on mouse P-gp (Wen, Verhalen, Wilkens, McHaourab, & Tajkhorshid, 2013). The enlargement of the portals (formed from TM4/TM6 and TM10/TM12) facing the inner-membrane leaflet side may be required for larger substrates like β-amyloid (~4 kD in size) (Kuhnke et al., 2007), to enter and bind inside the substrate binding pocket.

The wideness of our P-gp structures is comparable to recently published structures of other ABC transporters having similar ABC-B like protein folds (Jin et al., 2012; Park, Pardon, Wu, Steyaert, & Hol, 2012; C. A. Shintre et al., 2013; A. Ward et al., 2007). Interestingly, the structures of human ABCB10 and *T. maritima* TM287/288 (Hohl et al., 2012b) also have nucleotide-analogs bound to separated NBDs, suggesting a state of these transporters just prior to the formation of the ABC-sandwich that is essential for the structural transition to the outward-facing conformation. The degree of NBD separation might also be influenced in part by crystal lattice contacts, and may ultimately be constrained in the physiological context by the thickness of the hydrophobic section of the lipid bilayer and the TMDs of the transporter.

The discovery and structural elucidation of novel epitopes on therapeutic targets has tremendous value in the pharmaceutical industry. Antibodies like UlC2 that have been previously developed against P-gp target epitopes on the TMDs/extracellular surface (Mechetner et al., 1997). The structure of P-gp complexed with Nb592 (Crystal 3) reveals an entirely new epitope located on NBD1 (Fig. 2.4B) that is away from the Walker-A or conserved histidine (H583). The interaction between Nb592 and NBD1 of P-gp appears to be quite specific, as the CDR loops of Nb592 do not bind NBD2, despite a relatively high level of sequence conservation between NBD1 and NBD2 (Fig 2.4D). The binding location of Nb592 on NBD1 (Fig. 2.4) suggests that it sterically prevents the NBDs from dimerizing, and provides a structural basis to support and explain our biochemical observations that Nb592 (i) strongly inhibits ATPase activity, and (ii) completely abolishes Vi-induced ADP-trapping, in mouse P-gp (Fig. 2.7) signifying that NBD dimerization, followed by ATP “occlusion”, is essential to the formation of the

hydrolysis-competent state during the transport cycle. Future biochemical studies may establish whether Nb592 can bind and inhibit other ABC transporters via a similar mechanism.

Nb592 is a very robust protein and will likely bind to its intracellular epitope even when exposed to the reducing glutathiones in the cytoplasm, as our studies suggest (Fig. 2.7). This result is quite remarkable considering the presence of two highly conserved cysteines, C22 and C96 that form a disulfide bond in the Nb, as observed in our structure. Further validation of these types of inhibitors in whole cell settings will be necessary to assess their penetration through cell membranes and cross-reactivity with other cellular ATPases. Cell-penetration can be improved through innovations such as small peptide/chemical tags (Deshayes et al., 2010), while specificity may be enhanced through molecular scaffold engineering. Although the $IC_{50}$ of Nb592 (520 nM) is better than the QZ59 inhibitors that co-crystallized in the original P-gp structures (Aller et al., 2009), it is an order of magnitude less than Tariquidar (43 nM), a P-gp inhibitor that has been used in clinical trials (Martin et al., 1999; Tamaki et al., 2011). Thus, while Nb592 itself might not directly be used as a clinical inhibitor, it provides a template for the future development of this type of molecular scaffold. Several techniques may also be applied to improve the binding affinity of Nb592 (De Genst et al., 2004).

The x-ray structures of P-gp described here, along with those already published (Aller et al., 2009), will be useful for the molecular modeling of conformational trajectories to understand how the substrate-binding pocket changes during the transport cycle (Wen et al., 2013). As human and mouse P-gp share nearly 87% protein sequence

identity, our structures present experimentally-derived check-points useful for docking simulations. The development of these algorithms to more accurately model and predict substrate binding will provide a useful screening tool, complementing *in vivo* and cell-based studies, for understanding P-gp's role in drug pharmacokinetics. Collectively, these models could facilitate the development of therapeutically important compounds evading P-gp in several clinically relevant contexts, like penetration of drugs through the blood-brain barrier.

**2.7: P-gp protein expression and purification**

Gene-optimized mouse P-gp (*ABCB1a*, GenBank JF834158) was expressed as previously described (Aller et al., 2009; Bai et al., 2011) in 10 L cultures of *P. pastoris* grown in a bioreactor (Bioflow 415, New Brunswick Scientific). Protein expression was induced by addition of methanol (3.6 ml per hour per liter of culture volume) overnight. Cells were lysed by a single pass through a cell disrupter (TS-Series, Constant Systems, Inc.) at 40,000 psi. Cell wall and debris were removed by centrifugation (3,500 x g, 35 minutes, 4° C) and crude plasma membranes were obtained by centrifugation at 35,267xg for 2-3 hours at 4° C. The purification procedure is similar to that previously described (Aller et al., 2009) with some modifications. Membranes containing P-gp were resuspended in cold buffer (100 mM NaCl, 15% glycerol, 20 mM Tris pH 8.0, Sigma Protease Inhibitors) and solubilized with a final percentage of ~4.5% Triton® X-100 for 1-2 hours at 4° C. Insoluble material was removed by centrifugation at 38,400 x g, 4°C for 30-60 minutes and the supernatant was poured over a metal resin (Ni-NTA Superflow™, Qiagen) by FPLC (AKTA, GE Life Sciences). Immobilized P-gp was

buffer exchanged into 20 mM Hepes, 20 mM imidazole, 0.04% sodium cholate (Sigma) and 0.0675% β-DDM. Protein was eluted with buffer containing 200 mM imidazole pH 7.5. The eluted protein was then diluted 1:10 in buffer containing 20 mM Hepes pH 8.0, 100 mM NaCl, 0.2 mM TCEP, 0.04% sodium cholate, 0.065% β-DDM, and rebound to a new Ni-NTA column. The column was washed with buffer containing 20 mM imidazole, eluted with 200 mM imidazole, the protein concentrated (Centricon® YM-50 or YM-100; Millipore), spun at 95,000 rpm (TLA120.1 rotor) for 30-60 min at 4° C and subjected to gel filtration chromatography (Superdex™200 16/60; GE Healthcare) at 4° C.

**2.8: Elicitation of Nanobodies Against P-gp**

In order to generate the nanobodies, 2mg of purified, detergent solubilized mouse P-gp was injected into a llama (*Llama glama*) over a period of 6 weeks to elicit an immune response. The immunization, library construction and nanobody selection have been performed following standard procedures according to (Conrath et al., 2001), and a C-terminal His6-tagged nanobody library in pMES4 (genbank GQ907248) of $2x10^8$ independent clones was established using the *PstI/BstEII* site. P-gp-specific phages were recovered by incubating P-gp-coated wells with 100 mM triethylamine, pH 11.0, for 10 minutes. The P-gp-coated wells were then washed once with Tris-HCl, pH 6.8, and several times with phosphate-buffered saline. Finally, freshly grown TG1 cells were added to the wells to recover the non-eluted phage. After two rounds of selection, individual colonies were screened for the expression of P-gp-specific nanobodies: Maxisorb 96-well plates were coated overnight at 4° C with 1 µg/ml P-gp

in sodium bicarbonate buffer, pH 8.2. Residual protein-binding sites in the wells were blocked for 2 hours at room temperature with 2% milk in phosphate-buffered saline. Detection of P-gp bound nanobodies was performed with a mouse anti-His tag monoclonal (Serotec). Subsequent detection of the mouse anti-tag antibodies was done with an alkaline phosphatase anti-mouse IgG conjugate (Sigma). The absorption at 405 nm was measured 30-60 min after adding the enzyme substrate 4-nitrophenyl phosphate. Plasmids were extracted from the positive clones and transformed in *E. coli* WK6 strains.

**2.9: Nanobody Expression and purification**

Nb592 protein was produced in the *E. coli* WK6 strain described above. Bacteria were grown in terrific broth to an OD600 of 0.7 and then expression was induced by 1mM IPTG overnight at 28°C. Bacteria were then pelleted at 7,500xg for 15 minutes at room temperature. Pellets were resuspended in 15 ml TES buffer (0.2 M Tris pH 8.0, 0.5 mM EDTA, 0.5 M sucrose) and kept under slow agitation for 1hr at 4°C. Thirty mls of 4-fold diluted TES buffer was added and the sample osmotically lysed under slow agitation for 45 minutes at 4°C. Samples were then centrifuged for 30 minutes at 4°C and 6,000xg. Supernatant was used for purification on Ni-NTA resin (Qiagen). Binding to the Ni-NTA resin was performed at 4°C for 1h. The column was washed with 50mM phosphate buffer pH6.0, 1 M NaCl, then eluted with 50 mM sodium acetate buffer pH 4.5, 1M NaCl. The protein solution was neutralized using 1 M Tris pH 7.5. The eluted protein was then subjected to gel filtration (Superdex™75 10/300GL; GE Healthcare) at 4°C.

**2.10: Formation of P-gp-Nanobody complex**

The P-gp-Nb592 complex was generated by incubating 3-fold stoichiometric excess nanobody with P-gp purified from the metal resin elution step for 30 minutes at 4°C. The complex was then subjected to gel filtration chromatography as described above to remove excess Nb592.

**2.11: ATPase inhibition**

Purified P-gp was reduced with 1 mM DTT for 30 min on ice, and then excess DTT removed by passage through 1 ml Sephadex G-50 centrifuge columns equilibrated in 20 mM Hepes pH 7.4, 10% glycerol, 250 mM NaCl and 0.1% DDM as described in (Urbatsch, Gimi, Wilke-Mounts, & Senior, 2000). Protein was activated with 1% (w/v) *E. coli* polar lipids (Avanti) for 10 min at room temperature. Lipid-activated P-gp (0.5 to 2 µg) was pre-incubated with increasing amounts of Nb592 for 30 min on ice, then ATP hydrolysis assayed with 10 mM MgATP and 100 µM verapamil in a final volume of 50 µl for 15 min at 37ºC. Release of inorganic phosphate was determined by the Malachite green method (Urbatsch et al., 2000). Negative controls containing P-gp, Nb592, or both were assayed for 1 min at 37ºC, and were subtracted as background values. In some experiments P-gp was substituted for Cys-less P-gp (Tombline et al., 2006), purified in the absence of reducing agents, which gave essentially the same results. The data were fitted using nonlinear regression on Sigmaplot (v11) with the equation: $Y=d-(a*x^b/(c^b+x^b))$, where a is the verapamil-stimulated activity in the absence of Nb592, b is the Hill coefficient, c is the concentration for half-maximal inhibition ($IC_{50}$), d is the maximal inhibition, and x is the concentration of Nb592.

For vanadate trapping experiments, P-gp was activated with 1 mM DTT and *E. coli* lipids (2:1, w/w) for 10 min at room temperature (Urbatsch et al., 2000). Lipid-activated Pgp (3.3 µg, 3-5 µl) was pre-incubated with increasing concentrations of Nb592 for 5 min, and reacted with 200 µM 8-azido-[α32P]-ATP, 2 mM $Mg^{2+}$, and/or 100 µM verapamil and/or 200 µM orthovanadate in a final volume of 100 µl for 15 min at 37ºC. Unbound nucleotide was removed by passage through centrifuge columns, and the eluates UV crosslinked for 7 min on ice as described in (Urbatsch et al., 2000). Samples were resolved on 10% SDS-PAGE gels stained with Coomassie Brilliant Blue and the dried gels exposed to film.

**2.12: Crystallization, data collection, and structure determination of P-gp and P-gp-Nb592 complex**

Purified P-gp or P-gp-Nb592 were isolated after gel filtration chromatography at a protein concentration of 1-2mg/ml and subjected to reductive methylation (Rayment, 1997). Freshly made borane and formaldehyde were added to the protein solution at final concentrations of 50mM and 100mM, respectively and incubated in the dark for 2 hours at 4° C with gentle shaking. The reaction was quenched with the addition of ice cold 2.5mM glycine and incubated for 30 minutes at 4° C. Methylated P-gp was then concentrated to 1ml (Centricon® YM-50 or YM-100; Millipore 4) and subsequently diluted with 9ml of quench buffer (20 mM Tris pH 7.5, 100 mM NaCl, 0.2 mM TCEP, 0.04% sodium cholate, 0.065% β-DDM). The concentration/dilution step was repeated two times. In some cases, 2mM methyl-β cyclodextrin was added to

the dilute protein and the mixture was concentrated for crystallization.

Immediately before crystallization, P-gp and the P-gp-Nb592 complex were concentrated to 8-12mg/ml. P-gp-only crystals (Crystal Form A) described in this work were grown using the sitting drop method at 4º C by combining protein and precipitant at 1:1 (volume:volume). Crystal1 and Crystal2, as well as all point mutants of Pgp (Table 2.2), were grown at a protein concentration of 10-12mg/ml using 0.1 M Hepes (pH 7-8), 50 mM lithium sulfate, 10 mM EDTA, and 25-29.5% (w/v) PEG 600 at 4C. These crystals typically appeared after one to three days and continued to grow to full size in approximately two weeks. Crystals of the P-gp-Nb592 complex were grown using 0.1 M Hepes (pH 7-8) and 22-27% (w/v) PEG 600.

X-ray diffraction data were collected on cryo-cooled crystals at the Stanford Synchrotron Radiation Laboratory (SSRL) (BL 11-1), the Advanced Light Source (ALS) (BL5.0.1) and the Advanced Photon Source (APS) (23-ID-B, and 23-ID-D). Datasets were processed with HKL2000 (*HKL Research, Inc.*) and mosflm (Winn et al., 2011b). Experimentally derived protein phases were obtained via the multiple anomalous dispersion technique using diffraction data collected on the mercury LIII edge and inflection point (Table 2.1) calculated by the program PHASES . The overall combined phasing power was 1.9, with a figure of merit of 0.732. The experimental electron density map revealed that there was only one molecule in the asymmetric unit. An initial P-gp model was positioned using PDB 3G5U using the program Coot (Emsley, Lohkamp, Scott, & Cowtan, 2010b). The structures of Crystal1 and Crystal2 were determined by molecular replacement using the package MolRep as part of the CCP4 suite (Winn et al., 2011b). Similar to the original published structures of P-gp the

N-terminus (residues 1-33) was not visualized and no electron density was present for most of the linker region (residues 627-683), which is likely a flexible region connecting the two halves of P-gp.

To further validate these structures of P-gp (Crystal1 and Crystal2), we introduced single site cysteine mutants throughout the TM regions (Table 2.2). PCR-based mutagenesis was performed with pairs of complementary mutagenic primers that carry the desired cysteine (codon TGT) using the *pLIC-Opti-Pgp* expression plasmid (Bai et al., 2011) as a template, *Pfx50* DNA polymerase (Invitrogen) and the In-Fusion HD cloning kit (Clontech). All plasmids were verified by sequencing. Each mutant was expressed in *P. pastoris* strain KM71H (Aller et al., 2009; Bai et al., 2011), purified and crystallized as described above. Crystals of P-gp mutants were soaked with 5 mM ethyl-mercury chloride for 1-2 hour, flash cooled, and x-ray data collected at synchrotrons and using our in-house x-ray source (Bruker, Inc.). The identity and position of the corresponding mercury labeled cysteine residues yielded peaks in the anomalous difference Fourier (Figs. 2.2b and 2.3).

For the structure determination of the P-gp-Nb592 complex, the molecular replacement method was used followed by rigid body refinement of the TMD and NBD regions using Crystal1 as the model. The position of Nb592 was determined using a 2Fo-Fc difference map using phases derived from only a P-gp model. A homology model for Nb-592 was built from chain A of PDB 1HCV using Swiss Model (Schwede, Kopp, Guex, & Peitsch, 2003) and idealized using the Chiron server (http://troll.med.unc.edu/chiron/login.php). The preliminary nanobody model was manually fit into difference density and refined against the x-ray data using rigid body

refinement.

Crystallographic refinement using native data from all crystals of P-gp was accomplished using the simulated annealing protocol (mlf target; CNS v1.3) and later using Phenix v1.8.2-1309 (Adams et al., 2010a). A final round of group B-factor refinement and bulk solvent correction produced chemical models for Crystal1-3. The chemical geometry of the refined P-gp and P-gp-Nb592 complex structures were corrected and checked with molprobity (V. B. Chen et al., 2010a) yielding models with no violations in the Ramachandran phi-psi plot and good bond angle/bond-length geometry (Table 2.1) comparable with other structures in this moderate resolution range. All structures of P-gp were validated using a sigma-A weighted 2Fo-Fc composite simulated annealing (SA) omit map (iteratively omitting 5% of the model) as well as multiple Fo-Fc SA difference maps. The Fo-Fc maps were calculated using CNS v1.3 systematically omitting eleven consecutive residues throughout the model generously omitting a neighboring sphere size of 4.0 Å and a map cushion surrounding the omitted region of 2.0 Å. All models of P-gp and P-gp-Nb592 were also validated using Fo-Fc difference maps Figs. 2.9-2.11. Figures were generated using PyMOL (Otsuka et al., 2005) and Adobe Photoshop 7.0.

Chapter 2 is derived from the publication, of which I was a primary author: "Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide binding domain" Proc Natl Acad Sci, 110:13386-13391, 2013

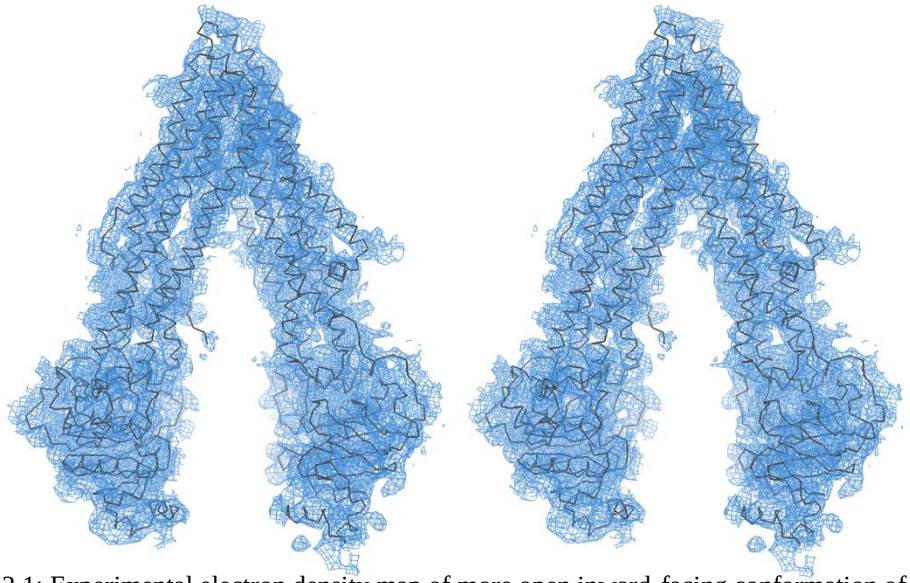


Figure 2.1: Experimental electron density map of more open inward-facing conformation of P-gp. Stereo view of 4.5Å resolution electron density map (blue mesh) contoured at 1σ. The P-gp monomer is presented in gray ribbon. The map was generated using the program PHASES (Furey & Swaminathan, 1997) as described.

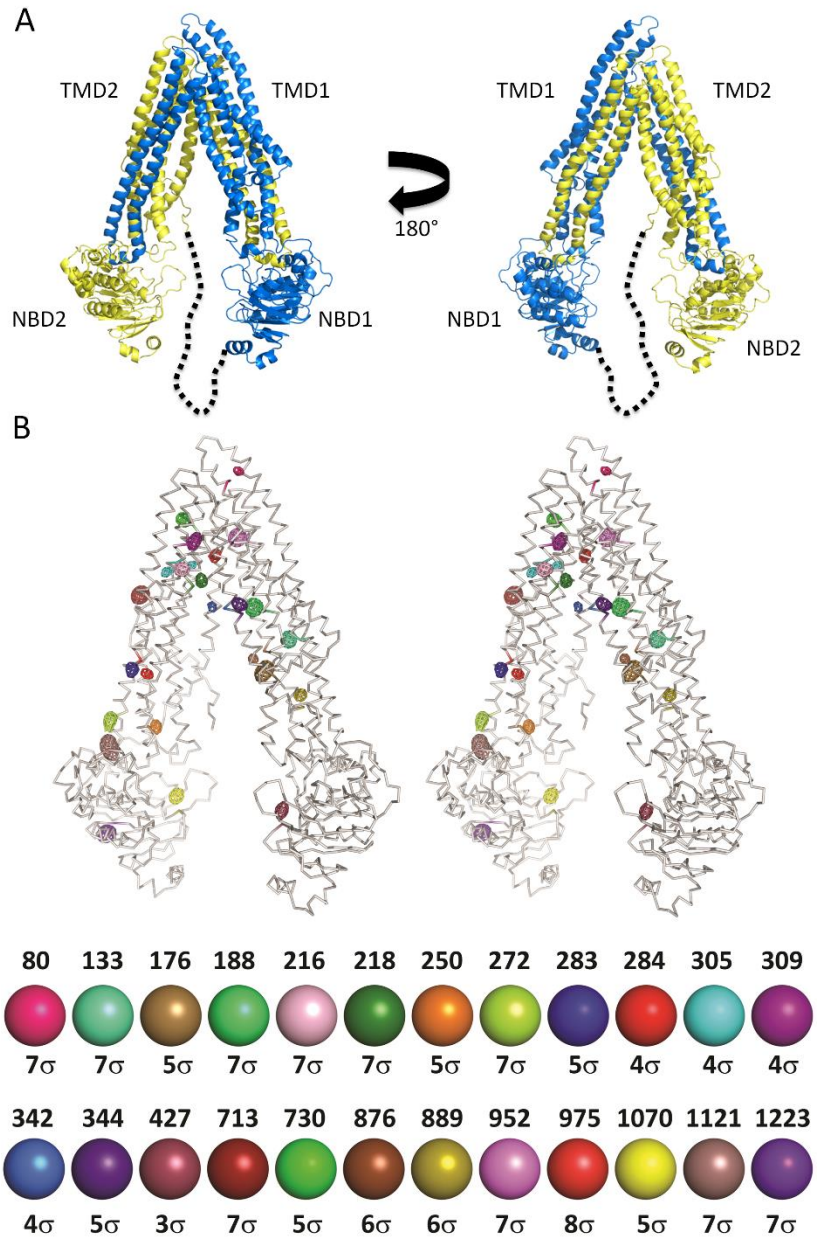



Figure 2.2: Structure of inward-facing P-gp. (A) Open inward-facing conformation (Crystal1) with the NBDs far apart. The N-terminal half of the protein (blue) and the C-terminal half (yellow) are connected by a flexible linker region (black dashed line) that is disordered in the structure. (B) Superposition of anomalous Fourier mercury peaks validating the position of knocked-in and wild-type cysteine residue positions. Relative peak positions were determined by aligning the corresponding TMD or NBD subdomains of the model derived from the mutant diffraction data (Table 2.2) with the model derived from Crystal1. The locations of the 24 anomalous mercury peaks (mesh, sigma values) were used to confirm the registration of amino acids in the structure.

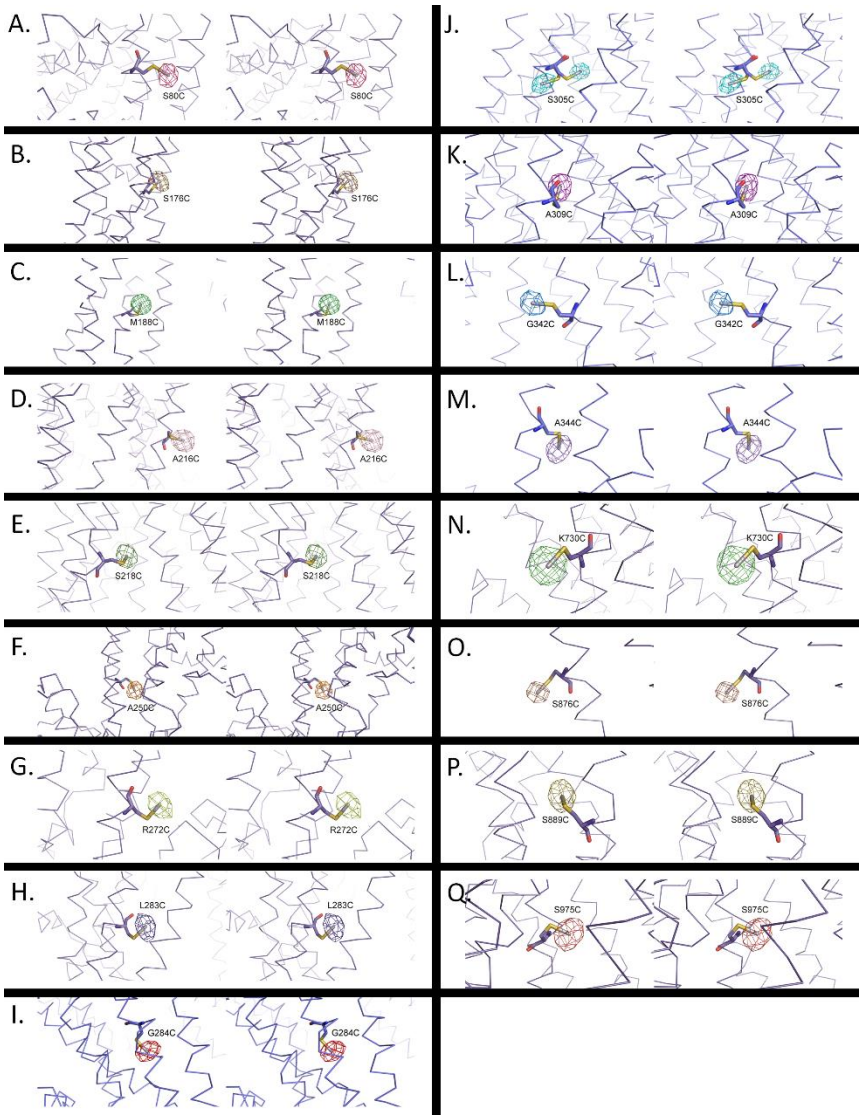



Figure 2.3: Close up views of mercury sites resolved from Hg soaked crystals of cysteine substitution mutants used for topology verification (Fig. 2.2B). The relative position of the Hg site is shown validating both the position and identity of each residue. Fourier syntheses were generated as described. The mutation positions and sigma cutoffs of the difference maps are: S80C, 7σ, S176C, 5σ; M188C, 7σ; A216C, 7σ; S218C, 7σ; A250C, 5σ; R272C, 7σ; L283C, 5σ; G284C, 4σ; S305C, 4σ; A309C, 4σ; G342C, 4σ; A344C, 5σ, K730C, 5σ; S876C, 6σ; S889C, 6σ; S975C, 8σ. Native cysteine sites include: C133, 7σ; C427, 3σ; C713, 7σ; C952, 7σ; C1070, 5σ; C1121, 7σ; C1223; 7σ. For residue 305 located on TM5 of the model, we observed two mercury peaks corresponding to two different rotamers of the mutated cysteine. Both of these positions are superimposed as shown in Fig. 2.3J. Interestingly, one position is on the inside of the molecule while the other is outside, demonstrating that the P-gp may have some degree of local flexibility in structure.

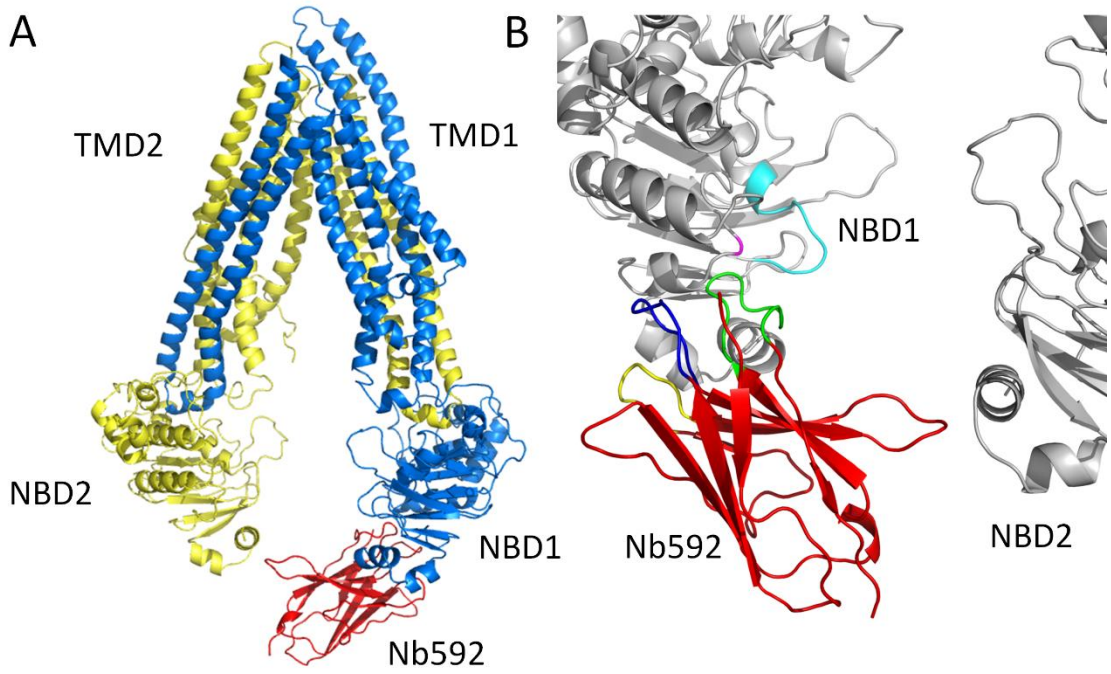
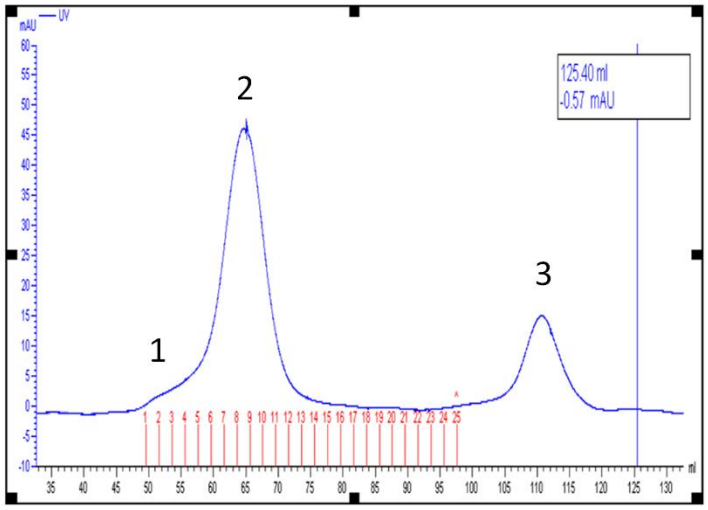



Figure 2.4: X-ray structure of P-gp in complex with Nb592. (A) Overview of the entire structure. Nb592 (red) binds to the C-terminus of NBD1. There are additional interactions with NBD2 (yellow). The binding site of the nanobody precludes the ABC domains from coming together, explaining its potent ATPase inhibition properties. (B) Close up view of Nb592 binding site on P-gp. The view is rotated 180° from A. The complementarity determining regions (CDR1: residues 25-33 for Nb592, blue; CDR2: residues 51-57 for Nb592, yellow; CDR3: residues 98-107, green) of the nanobody all interact with the C-terminal portion of NBD1. CDR3 inserts into a shallow pocket formed by three helices in NBD1. The walker-A motif (residues 423-430) located on NBD1 is colored in cyan. The conserved H583 is also shown in violet.

A.




B.

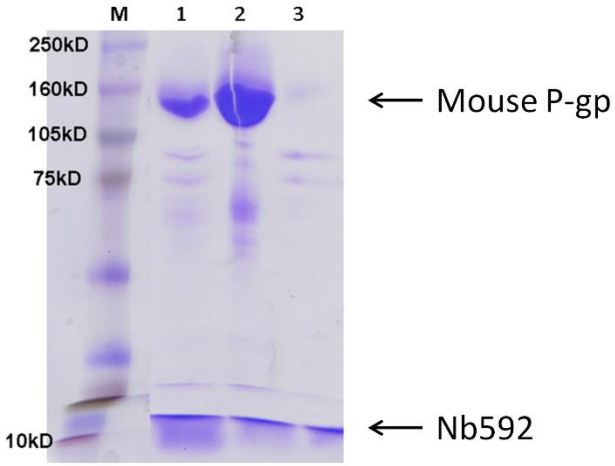



Figure 2.5: Size-exclusion chromatography (SEC) plot of P-gp-Nb592 complex. (A) Protein sample was injected on a Superdex™200 16/60 (GE Healthcare). The P-gp-Nb592 complex ran as a monodispersed peak eluting at 65 mls (2). The excessive Nb592 sample eluted at 111 mls (3). (B) The Coomassie stained SDS-PAGE gel of fractions at selected points shown corresponding to the region marked in the SEC plot.

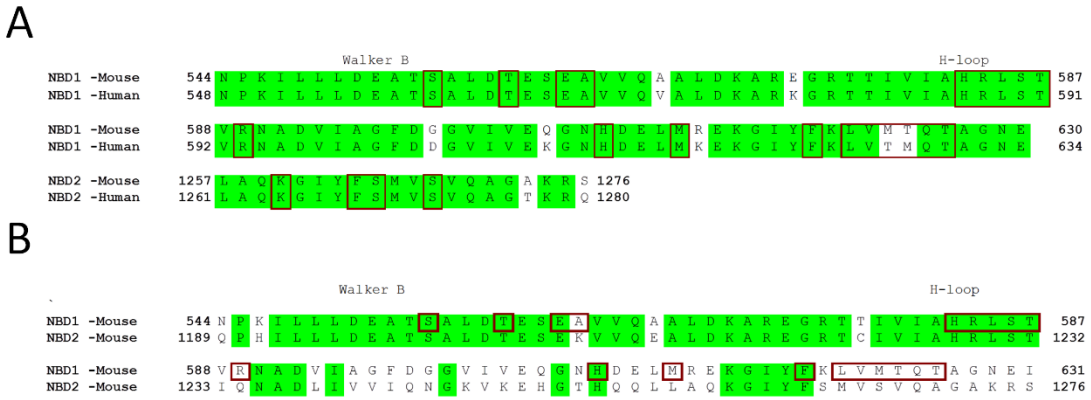



Figure 2.6: Sequence alignment between mouse and human P-gp (A) Protein sequence alignment between mouse and human P-gp showing only the regions bound by Nb592 CDR loops. Residues within 5Å of Nb592 in the mouse P-gp structure, and the corresponding human residues are boxed in red. Identical residues are indicated in green. (B) Protein sequence alignment of mouse NBD1 with mouse NBD2 in the same scheme as in (A). Nb592 does not bind mouse NBD2.

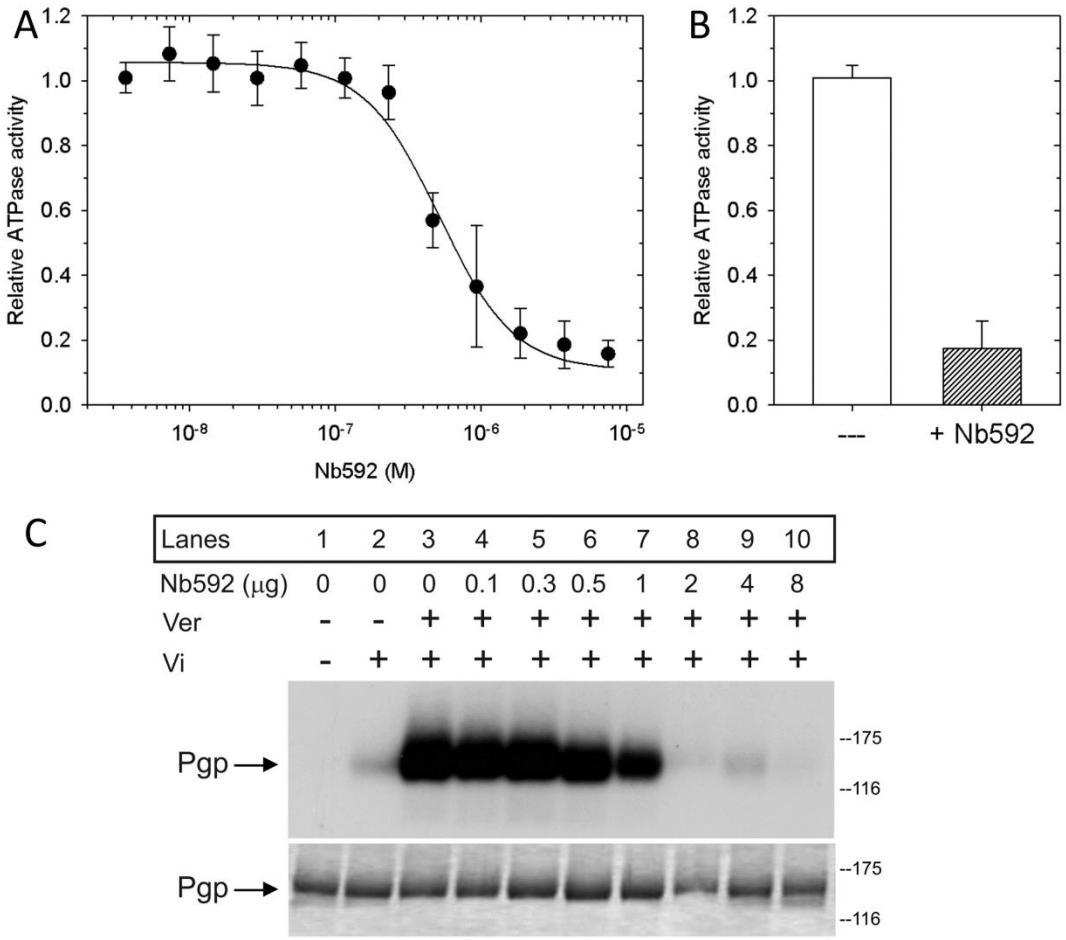



Figure 2.7: Nb592 is a strong inhibitor of P-gp's ATPase activity. (A) Nb592 inhibits verapamil-stimulated ATPase activity with an $IC_{50}$ of 520 ± 57 nM. Data points indicate the average activity ± SEM from four independent experiments, relative to P-gp's activity in the absence of Nb592. Lines represent non-linear regression analysis of the data points; $R^2$ value for the fit was 0.95. (B) Inhibition of verapamil-stimulated ATPase activity in the presence of 5 mM DTT, without or with 7.5 μM Nb592 (n=4). (C) Nb592 prevents vanadate induced 8-azido-[$\alpha^{32}$P]-ADP trapping in P-gp's catalytic sites. P-gp was incubated with 100 μM verapamil (VER) and 200 μM orthovanadate (Vi) in the 8-azido-[$\alpha^{32}$P]-ATP hydrolysis/trapping reaction as indicated above the lanes. Top, autoradiogram; bottom, Coomassie-stained SDS-PAGE gel showing the presence of P-gp (loading control).

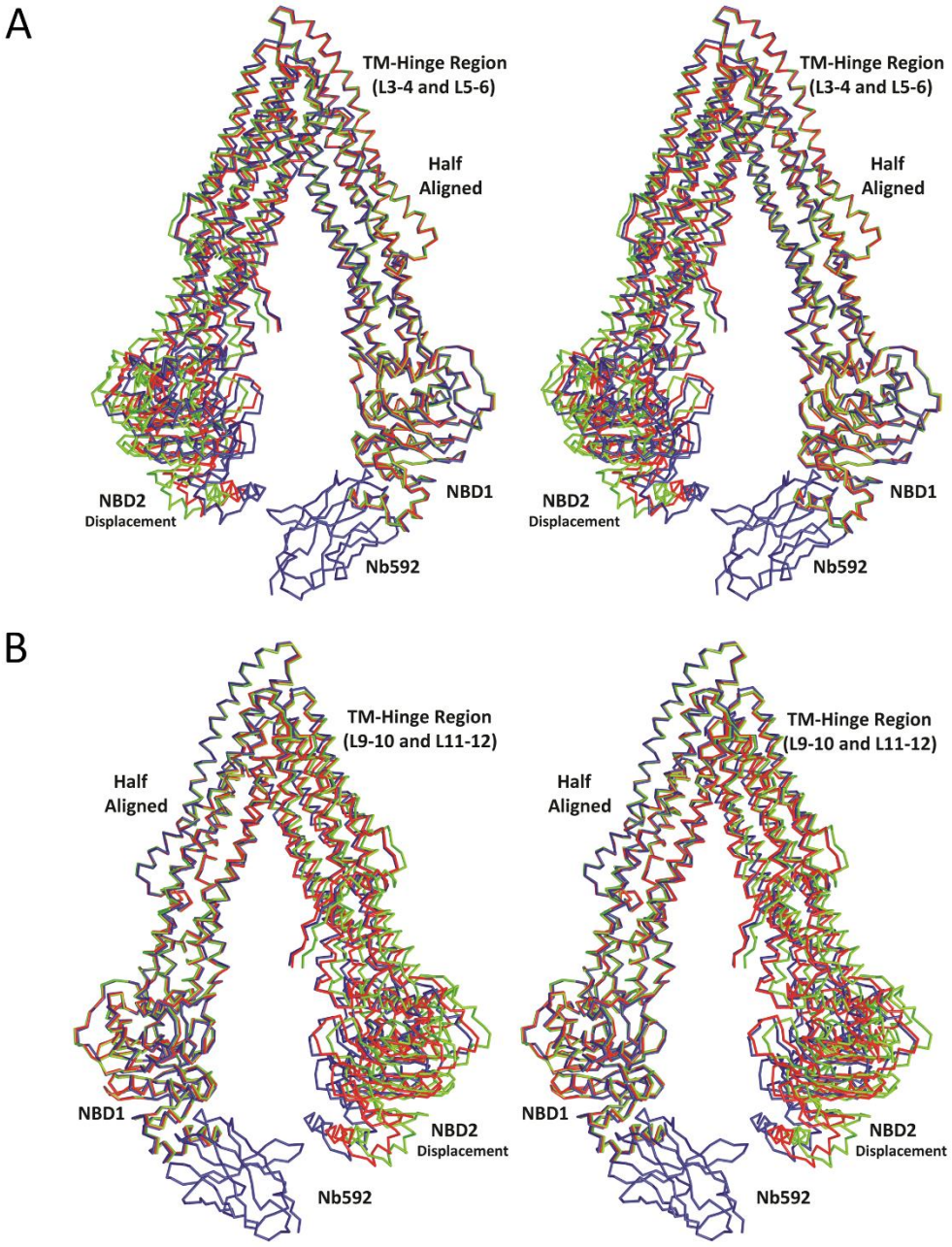



Figure 2.8: Conformational changes by P-gp. (A) Stereo view of the three P-gp structures described (Crystal1: red, Crystal2: green, P-gp-Nb592 complex [Crystal3]: blue) aligned using residues in TMD1 and NBD1 (residues 33-209, 852-961, and 320-626; designated as "Half Aligned"). The RMSD on Cα atoms for the aligned portion was 0.11 Å between Crystal2 and Crystal1 and 0.28 Å between P-gp-Nb592 complex (Crystal3) and Crystal1. The "TM-hinge regions" for this half of the molecule (L3-4 and L5-6 as described in the text) are marked. The relatively large displacement of the other half of the molecule, including NBD2, is clearly shown. The relative position of the Nb592 is marked in interaction with NBD1. It also makes a smaller contact with NBD2. The P-gp-Nb592 complex is the most closed inwardfacing conformation described in this study. (B) Same structural alignment as (A) turned 180°.

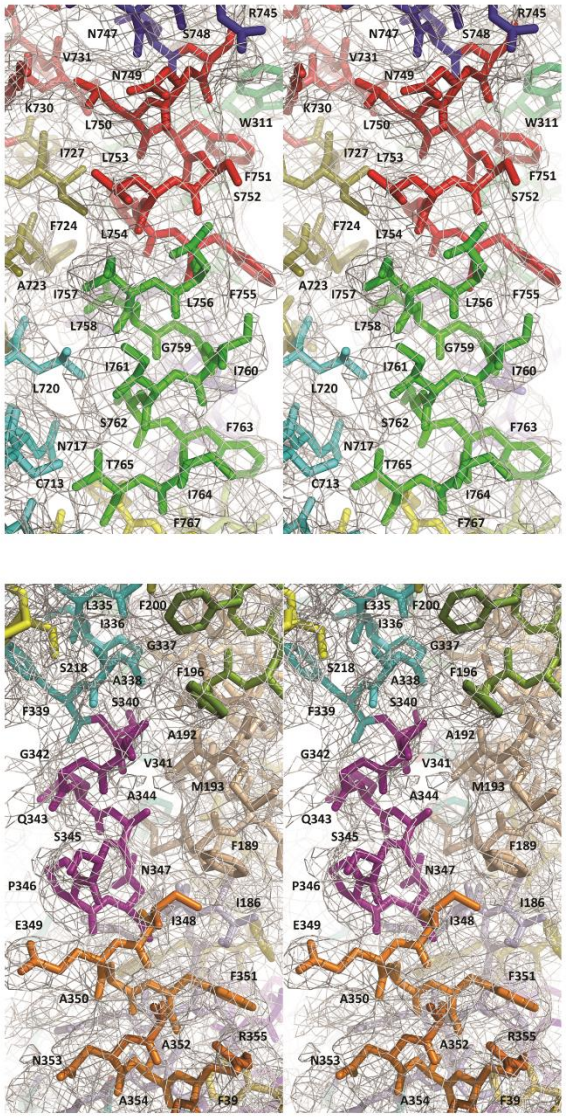



Figure 2.9: Stereo views of Fo-Fc simulated annealing omit map density for Crystal1 with multiple density maps superimposed. The Fo-Fc density maps (gray) were generated as described in the text. Difference density for the entire P-gp molecule was achieved by superimposing individual Fo-Fc maps calculated using sequential segments of 9 residues that were iteratively omitted from the model calculated with a neighboring sphere size of 4Å and a map cushion of 2Å. The windows of residues omitted from the model are shown in different colors. Maps are contoured at 1.0σ.

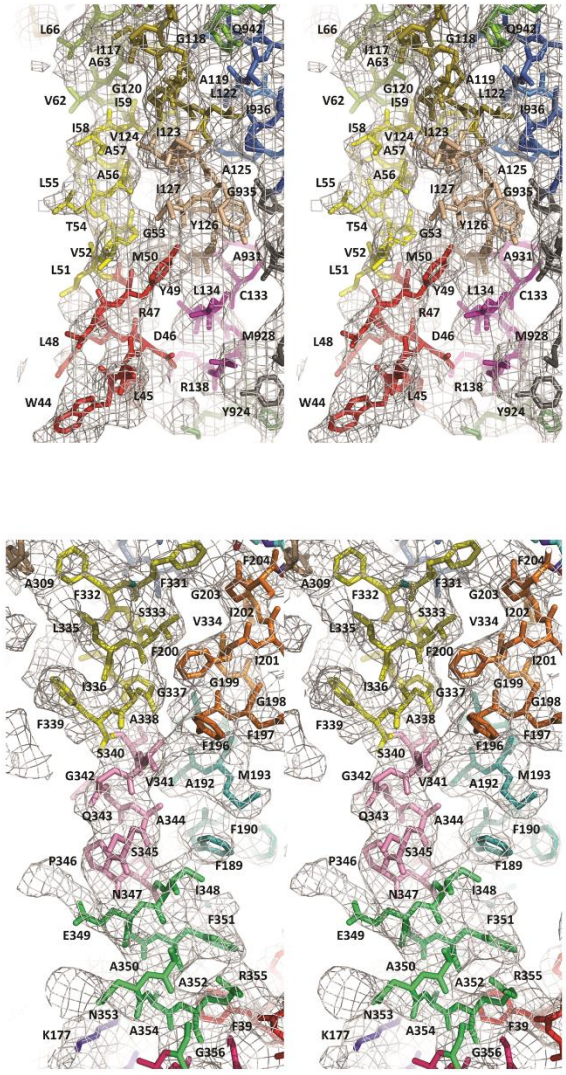



Figure 2.10: Stereo views of Fo-Fc simulated annealing omit map density for Crystal2 with multiple density maps superimposed. The Fo-Fc density maps (gray) were generated in Fig. 2.10. The windows of residues omitted from the model are shown in different colors. Maps were contoured at 1.0 σ.

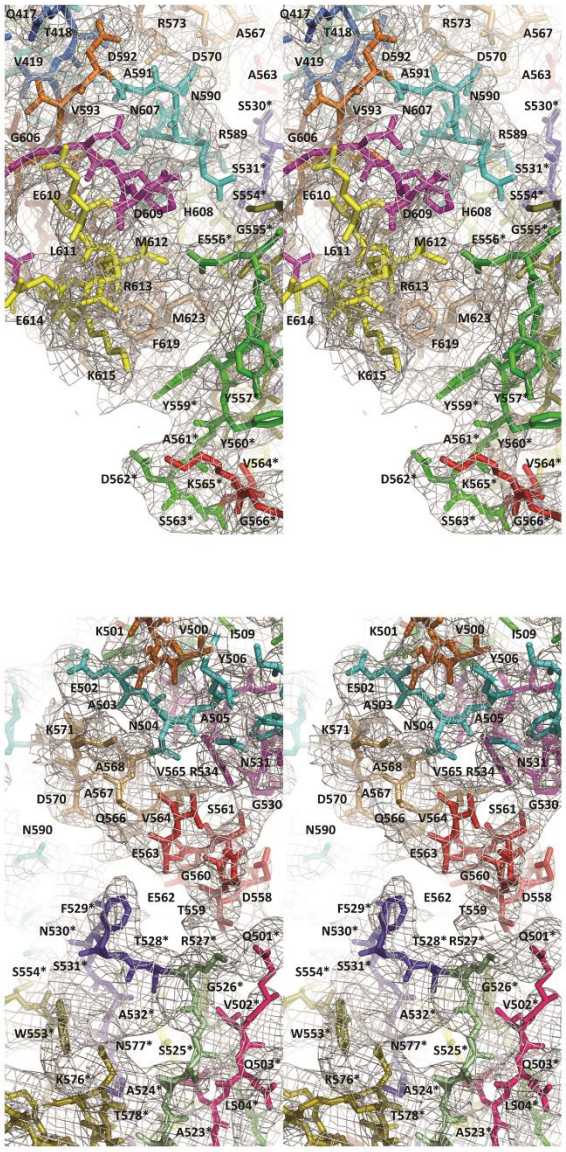



Figure 2.11: Stereo views of Fo-Fc simulated annealing omit map density for the P-gp-Nb592 complex with multiple density maps superimposed. The residues of Nb592 are marked with asterisk (*). The Fo-Fc density maps (gray) were generated as in Fig. 2.9. The windows of residues omitted from the model are shown in different colors. Maps were contoured at 1.0 σ.

Table 2.1: Data Collection, Phasing, and Refinement Statistics

| | **Crystal Form A** | **Crystal Form A** | **Crystal Form B** | **Crystal Form A** | |
|---|---|---|---|---|---|
| | Crystal 1 | Crystal 2 | Crystal 3 (Nb-bound) | Experimental Map Crystal | |
| **Data Collection**: | | | | | |
| Space Group: | P212121 | P212121 | P212121 | P212121 | |
| Unit Cell | | | | E1 | E2 |
| a,b,c (Å) | A=87.40, b=138.65, c=185.13 | a=90.65, b=138.29, c=194.72 | a=87.10, b=102.47, c=312.10 | a=87.92, b=137.37, c=181.44 | a=87.92, b=137.41, c=181.83 |
| α,β,γ (°) | α,β,γ=90 | α,β,γ=90 | α,β,γ=90 | α,β,γ=90 | α,β,γ=90 |
| Wavelength (Å) | 1.00511 | 0.97945 | 1.009 | 1.007 | 1.009 |
| Resolution (Å) | 3.8 | 4 | 4.1 | 4.2 | 4.2 |
| Redundancy | 3.4(3.6) | 2.4(2.3) | 5.3(5.3) | 4.9(4.9) | 4.9(4.9) |
| Rsym, %[a] | 0.066(0.414) | 0.038(0.234) | 0.072(0.350) | 0.076(0.469) | 0.076(0.478) |
| Completeness, % | 96.1(98.4) | 88.7(91.9) | 96.8(97.9) | 98.4(99.2) | 98.4(99.2) |
| **Model Building and Refinement**: | | | | | |
| Resolution (Å) | 3.8 | 4 | 4.1 | | |
| Sigma Cutoff | 2.04 | 2.02 | 2.05 | | |
| R Value %[b] | 32.47 | 31.7 | 32.39 | | |
| (Working + | | | | | |
| R Value %[b] | 32.31 | 31.58 | 32.28 | | |
| (Working Set) | | | | | |
| Free R Value[c] | 35.66 | 33.77 | 34.45 | | |
| Completeness, | 95.29 | 86.9 | 95.7 | | |
| Root Mean Square Deviations: | | | | | |
| Bond Lengths | 0.004 | 0.004 | 0.005 | | |
| Bond Angles (°) | 0.869 | 0.909 | 1.072 | | |
| Wilson B, (Å2) | 98.63 | 113.94 | 96.84 | | |
| Shrinkage | 0.9 | 0.9 | 0.9 | | |
| Solvent Radius (Å) | 1.11 | 1.11 | 1.11 | | |
| **Ramachandran Statistics (Molprobity)**: | | | | | |
| Outliers % | 0 | 0 | 0 | | |
| Allowed % | 8.83 | 8.91 | 9.89 | | |
| Favored % | 91.17 | 91.09 | 90.11 | | |
| Rotomer | 0 | 0 | 0 | | |
| Cbeta | 0 | 0 | 0 | | |

[a] $R_{sym}$ = S|I-<I>|S<I>, where I is the measured intensity of each reflection, and <I> is the intensity averaged from symmetry equivalents.

[b] $R = S|F_o-F_c|/S|F_c|$, where $F_o$ and $F_c$ are observed and calculated structure factors, respectively.

[c] $R_{free}$ was calculated from a subset of data (5%) omitted from the refinement.

Table 2.2: Data Collection Statistics of Single Site Mutations Used for Topological Validation of P-gp Model

| Data Set | Source | Unit Cell | | | Resolution(Å) | Multiplicity | Rsym[a] | Completeness |
|---|---|---|---|---|---|---|---|---|
| | | a(Å) | b(Å) | c(Å) | | | | |
| 1: S80C | APS23ID-D | 88.48 | 137.9 | 186.27 | 4.3 | 4.3 | 0.062 | 99.7 |
| 2: S176C | SSRL11-1 | 87.81 | 138.37 | 185.17 | 4.3 | 7.2 | 0.051 | 94.2 |
| 3: M188C | APS23ID-D | 89.14 | 138.86 | 188.93 | 4.3 | 4.5 | 0.055 | 97.3 |
| 4: A216C | APS23ID-D | 88.81 | 137.66 | 188.89 | 4.4 | 4.2 | 0.08 | 99.7 |
| 5: S218C | APS23ID-D | 89.95 | 138.76 | 191.96 | 4.2 | 4.3 | 0.071 | 99.7 |
| 6: A250C | ALS5.0.1 | 86.04 | 137.79 | 184.13 | 4.7 | 4 | 0.078 | 95.7 |
| 7: R272C | SSRL11-1 | 88.82 | 140.02 | 188.99 | 4.15 | 3.5 | 0.105 | 95.7 |
| 8: L283C | SSRL11-1 | 88.08 | 138.68 | 185.9 | 4.8 | 3.6 | 0.076 | 92.3 |
| 9: G284C | SSRL11-1 | 88.71 | 137.89 | 188.31 | 4.2 | 4.3 | 0.062 | 83.1 |
| 10: S305C | ALS5.0.1 | 90.9 | 138.8 | 195.96 | 4.7 | 4.4 | 0.057 | 93.9 |
| 11: A309C | SSRL11-1 | 86.76 | 138.01 | 187.71 | 6 | 3.4 | 0.091 | 85.4 |
| 12: G342C | SSRL11-1 | 87.31 | 138 | 183.71 | 4.65 | 3.6 | 0.074 | 97.7 |
| 13: A344C | SSRL11-1 | 89.96 | 138.47 | 189.95 | 4.7 | 3.4 | 0.057 | 95.8 |
| 14: K730C | SSRL11-1 | 90.22 | 138.34 | 192.73 | 4.25 | 4.1 | 0.057 | 81.9 |
| 15: S876C | APS23ID-D | 88.94 | 138.25 | 188.61 | 4.4 | 4.6 | 0.069 | 94.5 |
| 16: S889C | SSRL11-1 | 89.4 | 137.86 | 186.61 | 4.8 | 2.8 | 0.09 | 85.6 |
| 17: S975C | SSRL11-1 | 90.27 | 138.72 | 194.05 | 4.6 | 3.5 | 0.053 | 95.9 |

Chapter 3:

# Snapshots of Ligand Entry, Malleable Binding, and Induced Helical Movement in P-glycoprotein

## 3.1: Introduction

Membrane-embedded transporter proteins mediate the passage of metabolites and toxins across cellular membranes. The transporter P-glycoprotein (P-gp, ABCB1) is expressed in the intestines, liver, kidney, blood-brain barrier and affects the bioavailability, pharmacokinetics, and efficacy of drugs (Ingolf Cascorbi, 2006). P-gp also causes cellular multidrug resistance, hindering the treatment of several diseases, including cancers and HIV (Eckford & Sharom, 2009; Falasca & Linton, 2012; Gottesman & Ling, 2006). As such, the US Food and Drug Administration (FDA) now mandates documentation of P-gp-drug interactions for approval of any new drug (Giacomini et al., 2010; U.S. Food and Drug Administration, 2012). Thus, inhibition or evasion of P-gp without compromising therapeutic efficacy is a major goal of the pharmaceutical industry.

P-gp, and related ATP-Binding Cassette (ABC) transporters power protein conformational changes driving substrate translocation using ATP hydrolysis at the cytosolic nucleotide-binding domains (NBDs). Because P-gp can transport a wide array of structurally diverse compounds, a molecular framework for understanding how different ligands enter and bind is crucial to overcoming P-gp-mediated drug efflux.

previous structures of P-gp bound to two enantiomeric cyclopeptides provided the first structural view detailing how P-gp recognizes and binds ligands (Aller et al., 2009). The relatively lower resolution of these structures (4.4 and 4.35 Å) confirmed the presence of the ligand, but limited the interpretation of ligand/P-gp interactions. In this study, we report higher resolution (3.4-3.8 Å), structures of P-gp bound to a series of rationally engineered compounds to probe P-gp's ligand-binding pocket. In doing so, we reveal how (1) drug substrates may enter the transporter, (2) bind at multiple, shared-yet-discrete sites in the TMDs, and (3) might transmit this information to the NBDs to stimulate ATP hydrolysis. Together, these new structures provide snapshots of how structurally similar ligands, differing only in R-group size, can bind prior to drug-efflux. These higher resolution P-gp structures also resolve, to an extent, the issue of TM registry (Jin et al., 2012; J. Li, Jaimes, & Aller, 2014).

### 3.2: Synthesis of QZ-Homotrimers

*N*-Boc-(*S*)-amino acid selenazole esters were prepared from corresponding *N*-Boc-(*S*)-amino acids according to a previously published procedure (Tao et al., 2011). The selenazole ester derived from N-Boc-(*S*)-Ala (500 mg, 1.44 mmol) was subjected to hydrolysis by a solution of NaOH in mixed solvents (THF/MeOH/$H_2O$ = 5/1/2) and then Boc-deprotection by 50% trifluoroacetic acid in $CH_2Cl_2$. Resulting amino acid (~ 1.4 mmol) was dissolved in anhydrous acetonitrile (14 mL), treated with diisopropylethyl amine (DIEA, 1.0 mL, 6.1 mmol) and pentafluorophenol diphenylphosphinate (FDPP, 1.1 g, 2.9 mmol). The reaction was stirred at room temperature for 24 h before being concentrated under vacuum. The residue was

dissolved in $CH_2Cl_2$, washed with $NaHCO_3$, 5% HCl aqueous solution and NaCl solution successively. Organic phases were concentrated and the residue was purified by silica gel chromatography. The major product was collected and identified as desired product QZ-Ala (92 mg, 32%). Other cyclopeptides were obtained by following the same procedure.

**QZ-Ala:** while solid, 32%. $^{13}C$ NMR (75 MHz, $CDCl_3$) δ (ppm): 178.1 (3C), 160.0 (3C), 149.7 (3C), 131.4 (3C), 50.3 (3C), 25.2 (3C). $^1H$ NMR (300 MHz, $CDCl_3$) δ (ppm): 8.86 (s, 3 × 1H), 8.50 (d, *J* = 7.9 Hz, 3 × 1H), 5.67 (dq, *J* = 13.7, 6.8 Hz, 3 × 1H), 1.70 (d, *J* = 6.8 Hz, 3 × 3H). MS (ESI) calcd. for $C_{18}H_{19}N_6O_3Se_3$ $(M+H)^+$ 606.9, found: 606.9.

**QZ-Leu:** while solid, 45%. $^{13}C$ NMR (75 MHz, $CDCl_3$) δ (ppm): 177.3 (3C), 160.1 (3C), 150.1 (3C), 130.9 (3C), 51.7 (3C), 47.9 (3C), 25.2 (3C), 23.0 (3C), 22.5 (3C). $^1H$ NMR (300 MHz, $CDCl_3$) δ (ppm): 8.76 (s, 3 × 1H), 8.25 (d, *J* = 9.6 Hz, 3 × 1H), 5.85 – 5.56 (m, 3 × 1H), 1.98 – 1.87 (m, 3 × 1H), 1.84 – 1.68 (m, 3 × 2H), 1.04 (d, *J* = 5.9 Hz, 3 × 3H), 1.00 (d, *J* = 6.1 Hz, 3 × 3H). MS (ESI) calcd. for $C_{27}H_{37}N_6O_3Se_3$ $(M+H)^+$ 733.0, found: 733.0.

**QZ-Phe:** while solid, 40%. $^{13}C$ NMR (75 MHz, $CDCl_3$) δ (ppm): 175.0 (3C), 160.1 (3C), 149.5 (3C), 136.3 (3 × 2C), 131.5 (3C), 130.0 (3 × 2C), 128.9 (3C), 127.5 (3C), 55.5 (3C), 44.0 (3C). $^1H$ NMR (300 MHz, $CDCl_3$) δ (ppm): 8.72 (s, 3 × 1H), 8.48 (d, *J* = 8.5 Hz, 3 × 1H), 7.32 – 7.21 (m, 3 × 3H), 7.17 – 7.11 (m, 3 × 2H), 5.76 (td, *J* = 9.0, 4.8 Hz, 3 × 1H), 3.50 (dd, *J* = 13.1, 4.7 Hz, 3 × 1H), 3.04 (dd, *J* = 13.0, 9.4 Hz, 3 × 1H). MS (ESI) calcd. for $C_{36}H_{31}N_6O_3Se_3$ $(M+H)^+$ 835.0, found: 835.0.

### 3.3: Measurement of ATPase Activity

The ATPase activity of P-gp was measured at 37 °C using the ATP-regenerating system described by Vogel and Steinhart (Vogel & Steinhart, 1976) as modified by Urbatsch et al (Urbatsch, Sankaran, Weber, & Senior, 1995). Briefly, 1 μg P-gp was added to 100 μL 50 mM Tris-HCl (pH 7.5) buffer containing 10 mM ATP, 12 mM $MgCl_2$, 6 mM phosphoenolpyruvate, 1 mM NADH, 10 units lactate dehydrogenase, 10 units pyruvate kinase and test compounds over a range of concentrations. ATP hydrolysis was determined by the decrease in NADH absorbance at 340 nm using a Filtermax F5 Multiplate Spectrophotometer. ATPase activity was calculated using the following equation: $\Delta OD/(\varepsilon \times [protein] \times time)$, where $\Delta OD$ is the change in absorbance and $\varepsilon$ is the extinction coefficient of NADH. The concentration of purified P-gp was estimated by comparing the SDS/PAGE intensity of Coomassie-stained protein bands with known amounts of BSA. $EC_{50}$ values were calculated using GraphPad Prism – nonlinear regression (curve fit) from entire concentration range.

### 3.4: Calcein-AM Transport Assay

Calcein-AM (Calcein acetoxymethyl ester) is a membrane permeable P-gp substrate while its free acid, Calcein hydrolyzed by endogeneous esterases, is trapped in the cytosol and exhibits strong fluorescence. It has been commonly used in fluorescence-based transport assays for P-gp in whole cells systems (M. K. al-Shawi & Senior, 1993). Chinese hamster ovary cells CR1R12 were cultured in the presence of 5 μg/mL of colchicine to maintain P-gp over-expression. In general, ~ $5 \times 10^5$/well CR1R12 cells were pretreated with test compounds in serially increasing concentrations

at room temperature for 15 min, then Calcein-AM (0.25 μM) was added and incubated for an additional 15 min at room temperature while monitoring the fluorescence intensity (excitation at 485 nm, emission at 535 nm). 100% was the maximum fluorescence achieved by full inhibition of P-gp in CR1R12 cells. The means and standard deviations were obtained from quadruplet experiments.

### 3.5: Sensitization assay

CR1R12 cells were grown in the presence of increasing concentrations of colchicine for three days at given concentrations of test compounds. Cell densities were determined by using the sulforhodamine B colorimetric assay (Vichai & Kirtikara, 2006). 100% is the growth in the absence of colchicine. Parental AUXB1 cells were included as controls.

### 3.6: Expression, Purification Reductive Methylation, and Crystallization of P-gp

The expression, purification, reductive methylation, and crystallization of P-gp was performed as previously described (Ward et al., 2013), with the addition of an incubation step of P-gp with individual cyclopeptides following reductive methylation, but preceding crystallization. Briefly, methylated P-gp at a concentration of ~1-2 mg/ml was incubated with 2 mM cyclopeptide overnight at 4º C, then concentrated (Centricon® YM-50 or YM-100; Millipore 4) to ~10 mg/ml, diluted ten-fold in wash buffer (20 mM Tris pH 7.5, 100 mM NaCl, 0.2 mM TCEP, 0.01% sodium cholate, 0.035% □-DDM), then concentrated and diluted again before finally being concentrated to ~10-12 mg/ml for crystallization trials.

### 3.7: X-ray Data Collection, Structure Determination and Refinement of P-gp Cyclopeptide Cocrystal Structures

X-ray diffraction data were collected on cryo-cooled crystals at either the Stanford Synchrotron Radiation Laboratory (BL 11–1) or the Canadian Light Source (08ID-1). Diffraction data were processed with MOSFLM (Battye, Kontogiannis, Johnson, Powell, & Leslie, 2011) and reduced with SCALA (Evans, 2006) within the CCP4 suite of programs (Winn et al., 2011a). The 3.4 Å structure of P-gp was initially solved by molecular replacement (MR) with the program PHASER (McCoy et al., 2007) using a previously determined P-gp structure (pdb code: 4KSC (Ward et al., 2013)) as a search model. Commensurate with the improved resolution, the new electron density features guided adjustments to our model when compared to the same more 'open' crystal form we reported in 2013 (Ward et al., 2013). Residues 30-32 were located in the electron density, and resulted in a subsequent shift in the registration of residues in the first helix (residues 30-43) preceding TM1. Amendments were made to the topology of intracellular helix (IH) 1 (residues 154-168), extracellular loop (ECL) 3 (residues 318-338) and a portion of TM6 leading into the first NBD (residues 358-387). Within NBD1, residues 398-404, 424-427, 520-526, and 597-602 were rebuilt. EH2 was rebuilt from residue 689 to 708. A registry issue was amended from ECL4 (residue 738) to TM8 (residue 760) and another that constitutes segments of TM9, ECL5 and the beginning of TM10 (residues 826-855). The topology of IH3 was adjusted (residues 795-806), as was ECL6 (residues 961-967) and a portion of TM12 (residues 972-984). Further modifications were made in the region leading into and contributing to NBD2 (residues 1010-1028, 1042-1047, 1129-1137 and 1165-1172). Residues 1272 and 1273

were also located in the electron density maps at the C-terminus. As with all structures of P-gp determined to date, the ‘linker’ region (residues 627-688) was not located in the electron density. Many of the structural adjustments are in general agreement with recent corrections (J. Li et al., 2014) made to the model of the more ‘closed’ conformation of P-gp, first reported in 2009 (Aller et al., 2009). During the refinement process, the model underwent rigid body and restrained positional refinement, with hydrogens applied in their riding positions, using PHENIX.REFINE (Adams et al., 2010b) with TLS. Rounds of refinement were interspersed with manual inspection and correction against Sigma-A weighted electron density maps in COOT (Emsley, Lohkamp, Scott, & Cowtan, 2010a) and improvements to model geometry and stereochemistry were monitored using MOLPROBITY (V. B. Chen et al., 2010b). Subsequent cyclopeptide cocrystal structures were solved either by MR or rigid body refinement using the refined 3.4 Å model with residues from TM4 (218-243) and EH2 (689-694) removed to avoid biasing their placement within the electron density maps. These structures were then refined in using a similar fashion to the 3.4 Å structure. Ligand description dictionaries were calculated using PHENIX.ELBOW (Adams et al., 2010b) and the crystallographic positions of the incorporated seleniums validated using anomalous scattering methods.

### 3.8: Rational Engineering and Functional Characterization of Cyclopeptide P-gp Ligands

Previous functional studies have identified at least four, and potentially up to seven, sometimes overlapping, binding sites for substrates and inhibitors in the greasy,

polyspecific binding cavity of P-gp (Martin et al., 2000; Shapiro & Ling, 1997). To probe this phenomenon, the cyclic peptide QZ59-SSS (here named QZ-Val), previously reported in a co-crystal structure of mouse P-gp/Mdr1a (Aller et al., 2009), was taken as a base to engineer a series of selenium-labelled homotrimeric cyclopeptides (Figs. 3.1, and 3.2). The R-groups of this series were systematically varied to generate alanine-, valine-, leucine-, and phenylalanine-derived compounds (QZ-Ala, -Val, -Leu, -Phe), of increasing R-group size and hydrophobicity.

We measured the effect of each ligand on the basal level of ATP hydrolysis of purified P-gp. Our results revealed that the smaller compounds were more stimulatory and, most notably, the smallest ligand, QZ-Ala, potently stimulated ATP hydrolysis in a dose-dependent manner at the tested concentrations, similar to verapamil (Fig. 3.1b). Consistent with these substrate-like interactions, P-gp mediated mild resistance to QZ-Ala in cells (Fig. 3.3). These data suggest that the binding of QZ-Ala and, to a lesser degree QZ-Val at the TMDs, generates a signal to the NBDs accelerating nucleotide hydrolysis. We also characterized the ligands using P-gp-overexpressing CR1R12 cells. A pattern emerged in P-gp-mediated calcein-AM transport out of cells, whereby the ligand's potency to inhibit export was inversely related to R-group size (Fig. 3.1c). The smaller QZ-Ala inhibited calcein-AM export ($IC_{50}$=140 nM) the best compared to QZ-Val ($IC_{50}$=1.7 μM), QZ-Leu ($IC_{50}$=5.4 μM), and QZ-Phe ($IC_{50}$=24 μM) (Fig. 3.1c). Fitting the data in each case to the Hill equation all gave Hill coefficients >1, suggesting positive cooperativity for their binding to at least two sites. All four compounds also prevented P-gp-mediated export of the anticancer drug colchicine and sensitized CR1R12 cells in a dose-dependent manner (Fig. 3.4).

### 3.9: Structures of P-gp in Complex with Cylcopeptide Ligands

Prior to determining P-gp:ligand co-crystal structures, an improved structure for apo mouse P-gp/Mdr1a was determined to 3.4 Å. To date, this model is the highest resolution structure reported for a mammalian ABCB1/MDR1-type transporter (87% identity to human P-gp; Fig. 3.5; Table 3.1). This model is similar to those recently described (Ward et al., 2013), and differs significantly from the original reported in 2009 (Aller et al., 2009). The increase in resolution has also facilitated electron density map-driven improvements to the model (see methods; Fig. 3.5, Table 3.1) resulting in better refinement statistics.

To interrogate R-group-dependent variations in ligand binding, co-crystal structures of P-gp with each ligand were determined (Figs. 3.6a-c, and Fig. 3.7). Ligands were placed using difference electron densities and strong anomalous scattering from the three triangulated selenium atoms (Fig. 3.6d). All ligands bound to P-gp, and the majority of the coordinating side-chains, are well resolved in the electron density maps (Fig. 3.6d). The ligands make a number of interactions with binding pocket residues (Fig. 3.7), burying a high proportion of the ligands solvent accessible surface area (Table 3.2).

### 3.10: Comparison of Cylcopeptide Ligands in Complex with P-gp

Our P-gp:ligand co-crystal structures demonstrate the wide range of possible binding modes for similar substrates (Fig. 3.6a). Within the binding pocket, our structures group the four ligands into two subsets that correlate with size and hydrophobicity (Figs. 3.6b-c, and Fig. 3.7). The smaller ligands QZ-Ala and QZ-Val

(Subset A) share an upper and lower binding site (Fig. 3.6b), while the larger and more hydrophobic ligands QZ-Leu and QZ-Phe (Subset B) share a different upper binding site, with QZ-Phe also binding to a second, unique lower site (Fig. 3.6c, and Fig. 3.7). P-gp consists of two pseudo-symmetric halves, each containing six transmembrane (TM) helix bundles. Viewed from the plane of the membrane, ligands bound to the upper sites from Subsets A and B are wedged in the apex of the cavity, engaging TM helices from both pseudo-halves (Fig. 3.6b-c). In contrast, ligands in the lower sites interact with distinct pseudo-halves of P-gp that are mutually exclusive to each ligand subset (Fig. 3.6b-c, and Fig. 3.7).

### 3.11: Movement of TM4 Upon Ligand Binding

In previous structures of apo mouse P-gp, and also for ligands pertaining to Subset B, TM4 adopts a mostly straight-helical conformation characterised by weaker electron density; indicative of regional flexibility (Ward et al., 2013). Subset A ligands not only reveal an ordered TM4, but a large conformational change (up to 11 Å when comparing corresponding Cα positions) (Fig. 3.8a). This structural kink is not likely a consequence of lattice contacts as the crystal form is the same (Table 3.1). Thus, we must conclude that binding of Subset A ligands induces these changes in TM4. This ligand binding-induced kinking of TM4 begins at Pro219 before returning to the apo-WT topology at Tyr243 in the ball-and-socket region close to NBD2 (Loo, Bartlett, & Clarke, 2013).

The observed divergence of TM4 upon substrate/ligand binding may have significant biological implications. TM4 and TM6 comprise an intramembranous portal

for substrate entry to the binding cavity and a conformational change in these helices may influence the entry and binding of ligands (Loo & Clarke, 1994, 2005; Woebking et al., 2008). In the recent structure of *Cyanidioschyzon merolae* P-gp, a portion of TM4 in the wild-type protein was inherently disordered (Kodan et al., 2014). Mutations in this region not only resulted in a well-ordered, straight-helical conformation, but also functionally disrupted substrate transport, suggesting a role for TM4 in facilitating substrate entry and/or binding (Kodan et al., 2014). For co-crystal structures pertaining to Subset A ligands, the movement of TM4 brings residues 221-228 closer to the bound ligands in the lower binding sites, fostering an intermolecular interaction with Trp228, a residue that has been implicated in steroid binding to P-gp (Gruol, King, & Kuehne, 2002). We propose that the movement of TM4 upon binding of Subset A compounds provides a structural glimpse of the induced-fit model of drug binding proposed for P-gp nearly a decade ago (Loo, Bartlett, & Clarke, 2003b). Taken in the context of our biochemical data and published crosslinking and FRET studies on P-gp and its homologs (Doshi & van Veen, 2013; Loo, Bartlett, & Clarke, 2003a), the induced-fit model provides a starting framework to rationalize how substrate-stimulated ATP hydrolysis in P-gp may be initiated.

### 3.12: Ligand Binding Site at the Membrane Interface

Previous biochemical experiments have led to the proposal that P-gp extrudes drugs from the inner leaflet of the plasma membrane, functioning as a so-called “hydrophobic vacuum cleaner” (Bolhuis et al., 1996; de Graaf, Sharma, Mechetner, Schimke, & Roninson, 1996; Raviv, Pollard, Bruggemann, Pastan, & Gottesman, 1990).

In our QZ-Val co-crystal structure, we observed an additional binding site on the exterior of P-gp bounded by residues from TM8, TM12, and elbow helix (EH) 2 (Fig. 3.8b). The site faces away from the transporter, but lies close to the predicted membrane-water interface and intramembranous substrate entry portal. Drug binding near the EH has been reported using electron paramagnetic resonance on a bacterial P-gp homolog that transports lipids (Smriti, Zou, & McHaourab, 2009). These data lend strong credence to the proposal of an initial lower-affinity "ON-site" for a ligand near the inner leaflet of the lipid bilayer preceding the higher-affinity "ON-site(s)" within the central binding cavity (M. Al-Shawi & Omote, 2005; Dey, Ramachandra, Pastan, Gottesman, & Ambudkar, 1997).

### 3.13 Context of the Presented Structures

Understanding drug-transporter interactions is indispensable for engineering drugs to inhibit or evade P-gp. The findings presented here complement and extend on related X-ray structures (Aller et al., 2009; Hohl, Briand, Grutter, & Seeger, 2012a; Jin et al., 2012; Kodan et al., 2014; Chitra A. Shintre et al., 2013; Andrew Ward, Christopher L. Reyes, Jodie Yu, Christopher B. Roth, & Geoffrey Chang, 2007; Ward et al., 2013), as well as many biochemical/biophysical studies (M. Al-Shawi & Omote, 2005; Juliano & Ling, 1976). We have conducted the first detailed structure-activity-relation (SAR) study on how altering the size and hydrophobicity of R-groups to a known P-gp ligand can impact its interactions with this clinically important protein. The mechanistic concepts delivered through this structural study include (1) ligand entry *via* the elbow helix binding site, (2) distinct, and shared binding modes, and (3) ligand

binding-induced fit that could cause transmission coupled to ATP catalysis. These structural observations provide an excellent basis to biochemically test hypotheses for further studies on this transporter.

Chapter 3 is derived from the follow publication, which is currently under submission to ActaD: “Snapshots of Ligand Entry, Malleable Binding, and Induced Helical Movement in P-glycoprotein” in submission.

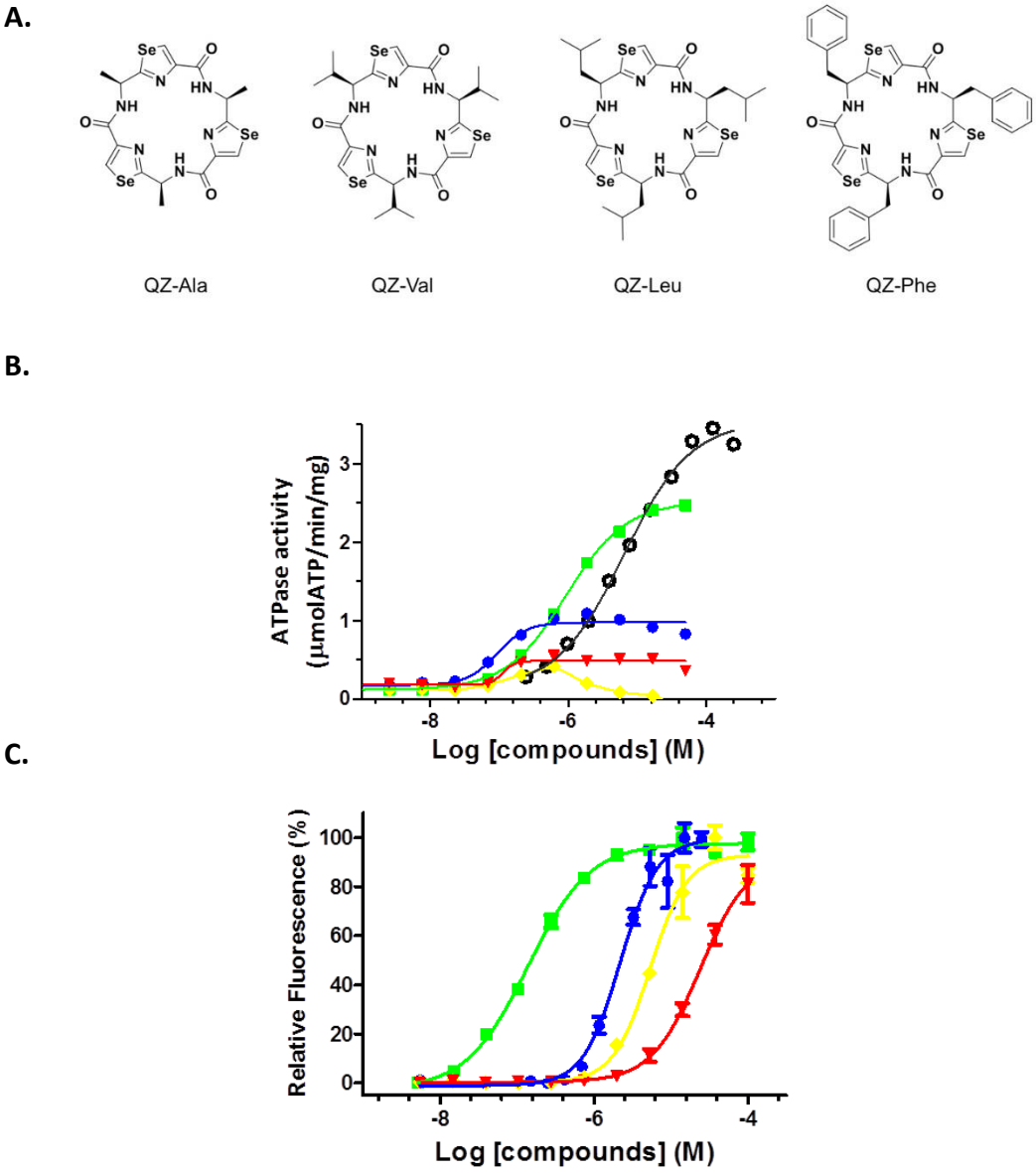



Figure 3.1: Structure and function of selenium-labelled homotrimeric cylcopeptides. (A) Structure of the cyclopeptides bear an identical backbone with side chains systematically increased in size and hydrophobicity. (B) Stimulation of the basal ATPase activity of purified P-gp. QZ-Ala (green) conferred the highest degree of stimulation (~16-fold) relative to basal activity with an $EC_{50}$ value of 0.92µM followed by QZ-Val (blue, ~7-fold). Data for verapamil (black), QZ-Leu (yellow), and QZ-Phe (red) are shown. (C) Inhibition of Calcein-AM transport in P-gp-overexpressing CR1R12 cells. Same colour scheme is applied for each compound as in (B), and the data was fitted using Hill equation. The mean and SD of triple and quadruplet experiments are shown in (B) and (C), respectively.

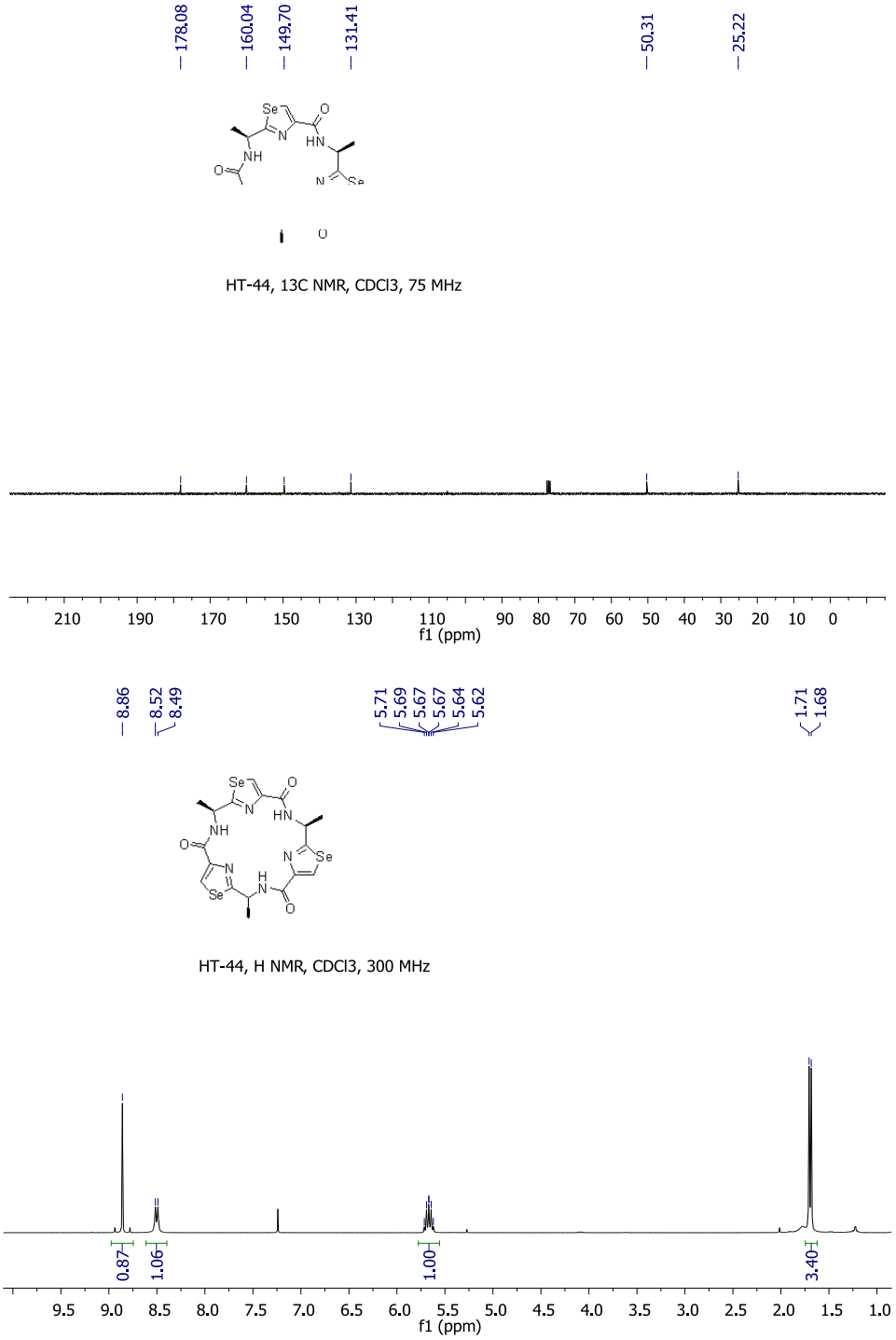



Figure 3.2: NMR spectra of QZ-Ala, QZ-Leu and QZ-Phe, $^{13}C$ (top) and $^{1}H$ (bottom) spectra are shown for each compound.

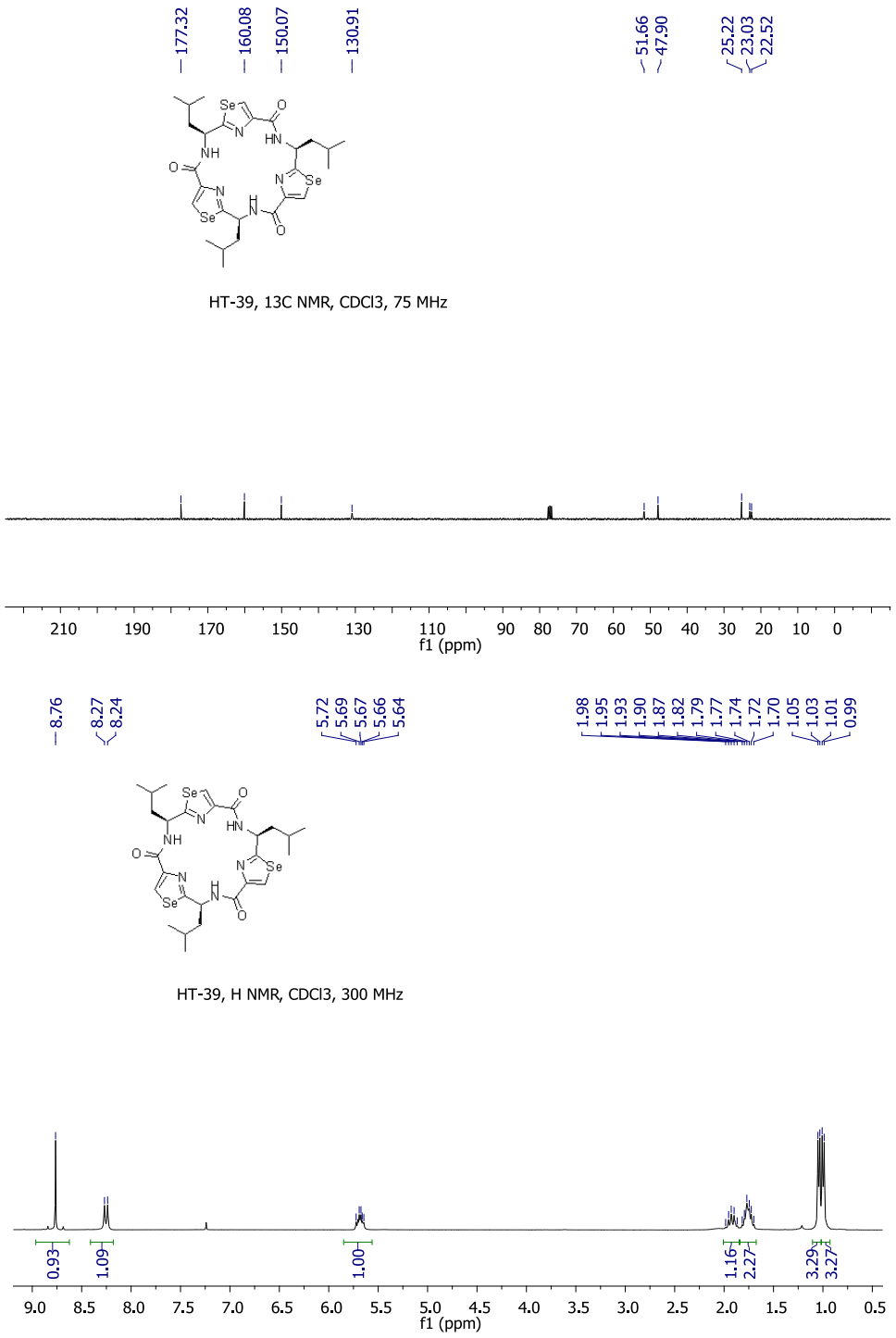



Figure 3.2: NMR spectra of QZ-Ala, QZ-Leu and QZ-Phe, $^{13}C$ (top) and $^{1}H$ (bottom) spectra are shown for each compound, continued.

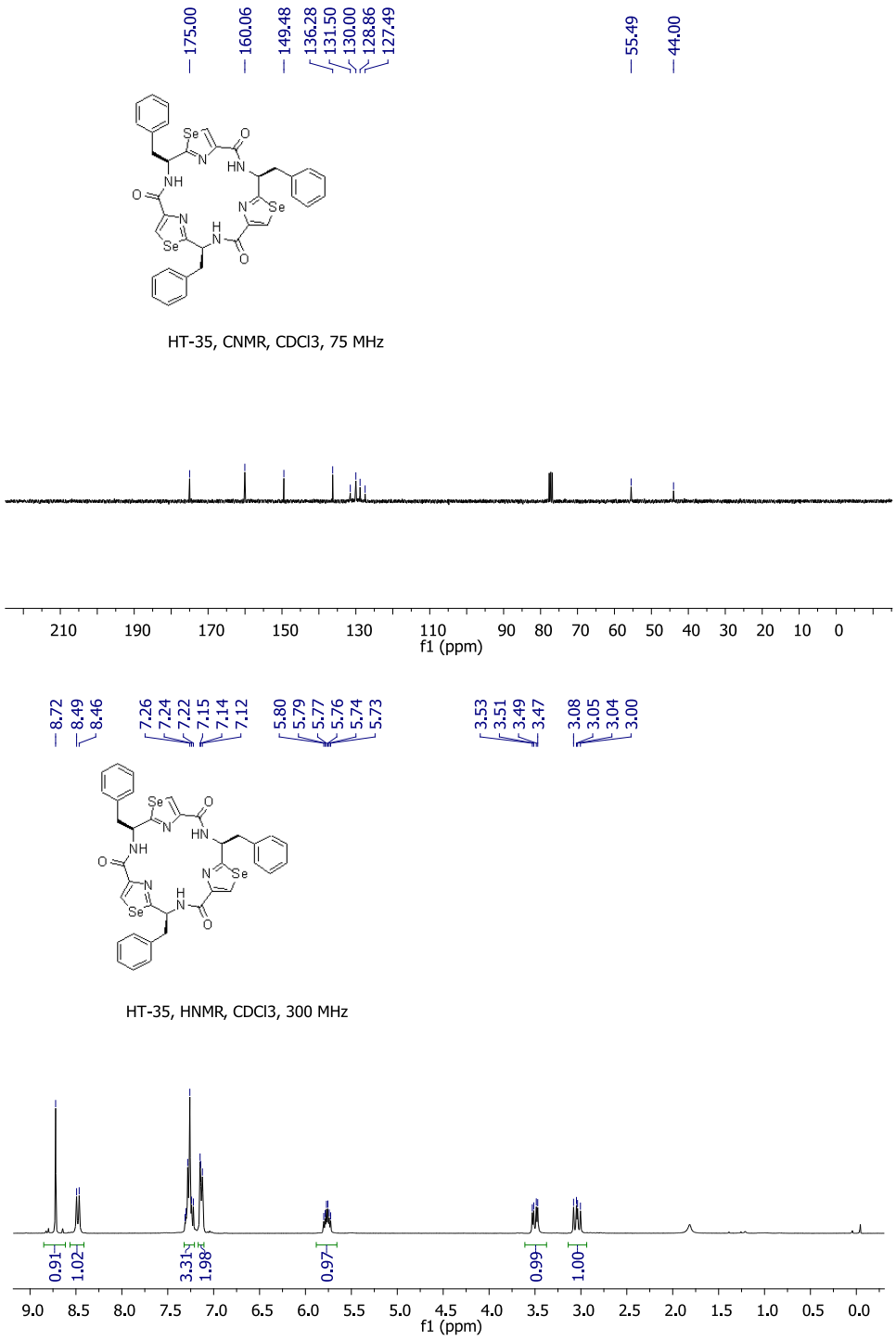
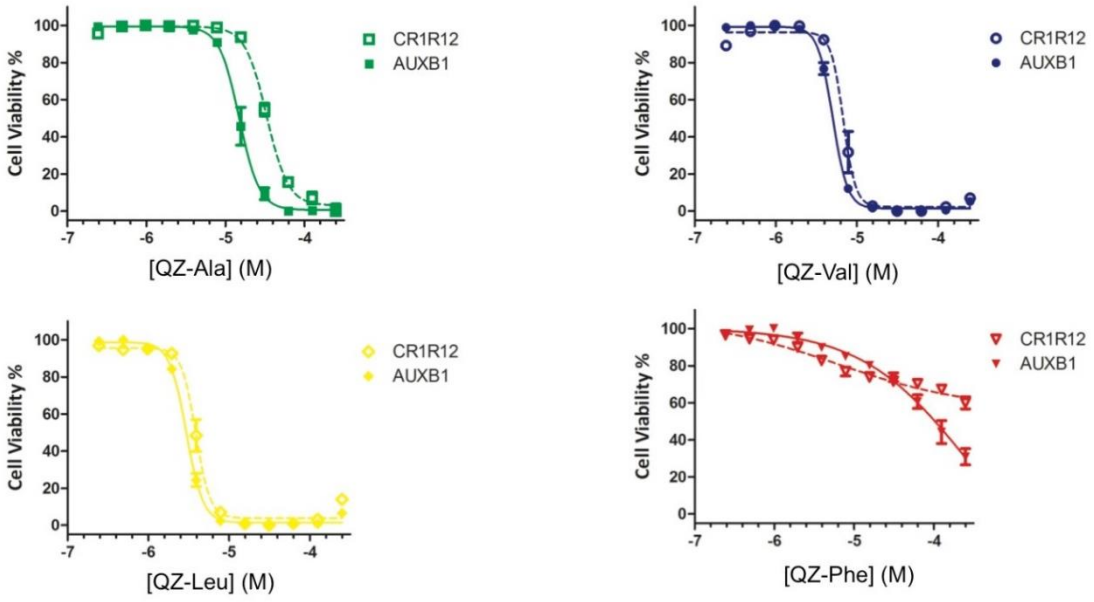



Figure 3.2: NMR spectra of QZ-Ala, QZ-Leu and QZ-Phe. $^{13}C$ (top) and $^{1}H$ (bottom) spectra are shown for each compound, continued.




| | toxicity ($IC_{50}$ μM) CR1R12 / AUXB1 |
|---|---|
| QZ-Ala | 32.1 ±0.9 / 14.8±0.5 |
| QZ-Val | 7.1±0.3 / 5.3±0.1 |
| QZ-Leu | 3.9±0.1/ 3.1±0.1 |
| QZ-Phe | >100 μM/ ~60 μM |

Figure 3.3: Cytotoxicity of the homotrimeric cyclopeptide compounds to P-gp-over-expressing CR1R12, and parental control AuxB1, cells. P-gp mediates mild resistance to QZ-Ala. For QZ-Phe, full growth of inhibition was not observed due to limited solubility of this compound at high concentrations. The mean and SD of quadruplet experiments are shown; data were fit using the Hill equation. QZ-Ala, QZ-Val, QZ-Leu and QZ-Phe are drawn in green, blue, yellow and red curves, respectively. The concentrations are plotted in log scale.

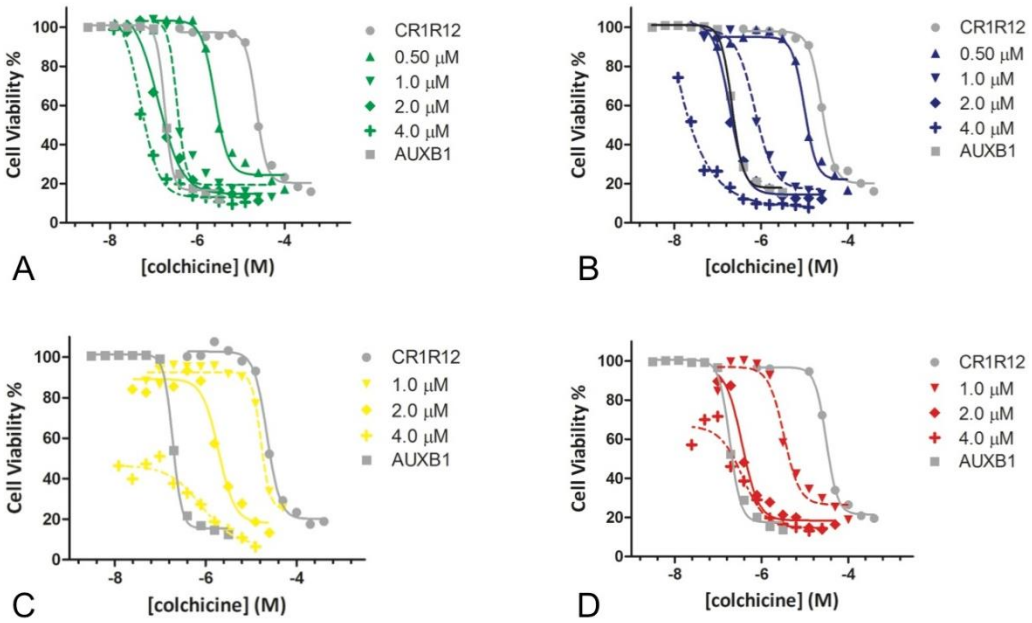



Figure 3.4: Sensitization of CR1R12 cells to the anti-cancer drug colchicine by homotrimeric cyclopeptides. Relative to parental control AUXB1 cells, P-gp-overexpressing CR1R12 cells are ~200 times more resistant to the anticancer drug colchicine (gray lines in all panes). Their 50% growth inhibition ($GI_{50}$) values are 0.14 μM and 27 μM respectively. Presence of (a) QZ-Ala, green, (b) QZ-Val, blue, and (d) QZ-Phe, red, in the growth media sensitized the CR1R12 cells to colchicine in a concentration-dependent manner with full reversal of P-gp mediated multidrug resistance seen at 2 μM; this concentration gave essentially the same $GI_{50}$ value as seen in AuxB1 cells. Presence of increasing concentrations of (c) QZ-Leu, yellow, also sensitized CR1R12 cells to colchicine, where full reversal of multidrug resistance was not observed because this compound became cytotoxic at higher concentrations. Shown are representative data from three independent experiments. The concentrations are plotted in log scale.

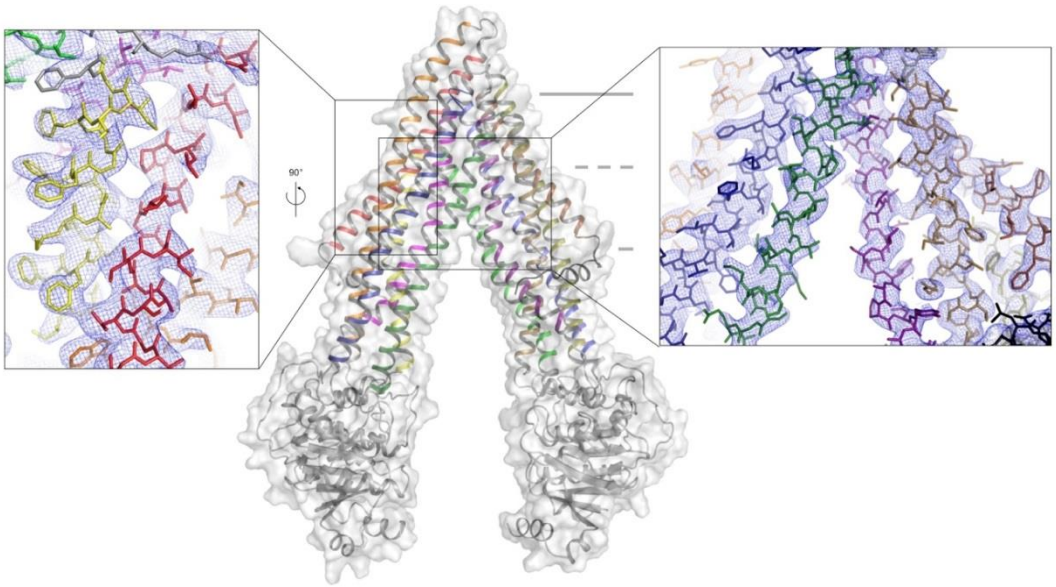



Figure 3.5: Overview of mouse P-gp at 3.4 Å resolution. Magnified insets at different orientations are shown with the resulting 2m$F_o$-D$F_c$ electron density contoured at 1σ, individual transmembrane helices are shown as different colors.

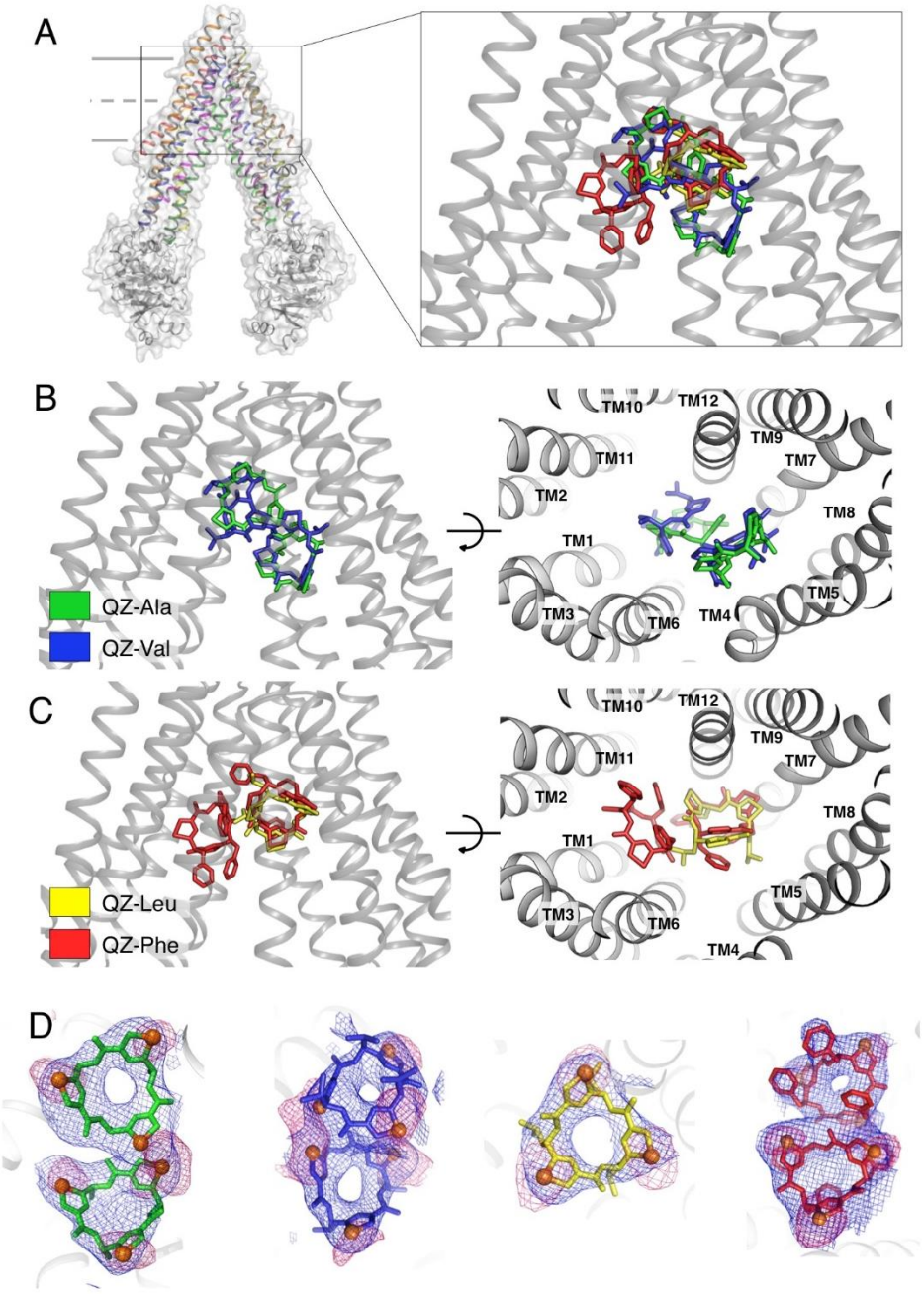



Figure 3.6: Overview of P-gp cyclopeptide co-crystal structures. (A) Superposition of homotrimeric cyclopeptide compounds bound to P-gp, showing their relative location and orientation in the substrate-binding pocket. (B) Two orientations of Subset A ligands (QZ-Ala and QZ-Val; displayed as sticks) bound in the substrate-binding pocket of P-gp. (C) Two orientations of Subset B ligands (QZ-Leu and QZ-Phe) bound in the substrate-binding pocket of P-gp. (D) Close-up view of ligands (coloring as in panes C and D), with the resulting $2mF_o$-$DF_c$ electron density (where m is the figure of merit and D is the Sigma-A weighting factor) in blue (contour level of 1.0σ) and anomalous difference density peaks in pink (σ contour level of 3.0 for QZ-Ala, -Val, -Phe and 4.0 for QZ-Leu) for the selenium atoms (orange spheres).

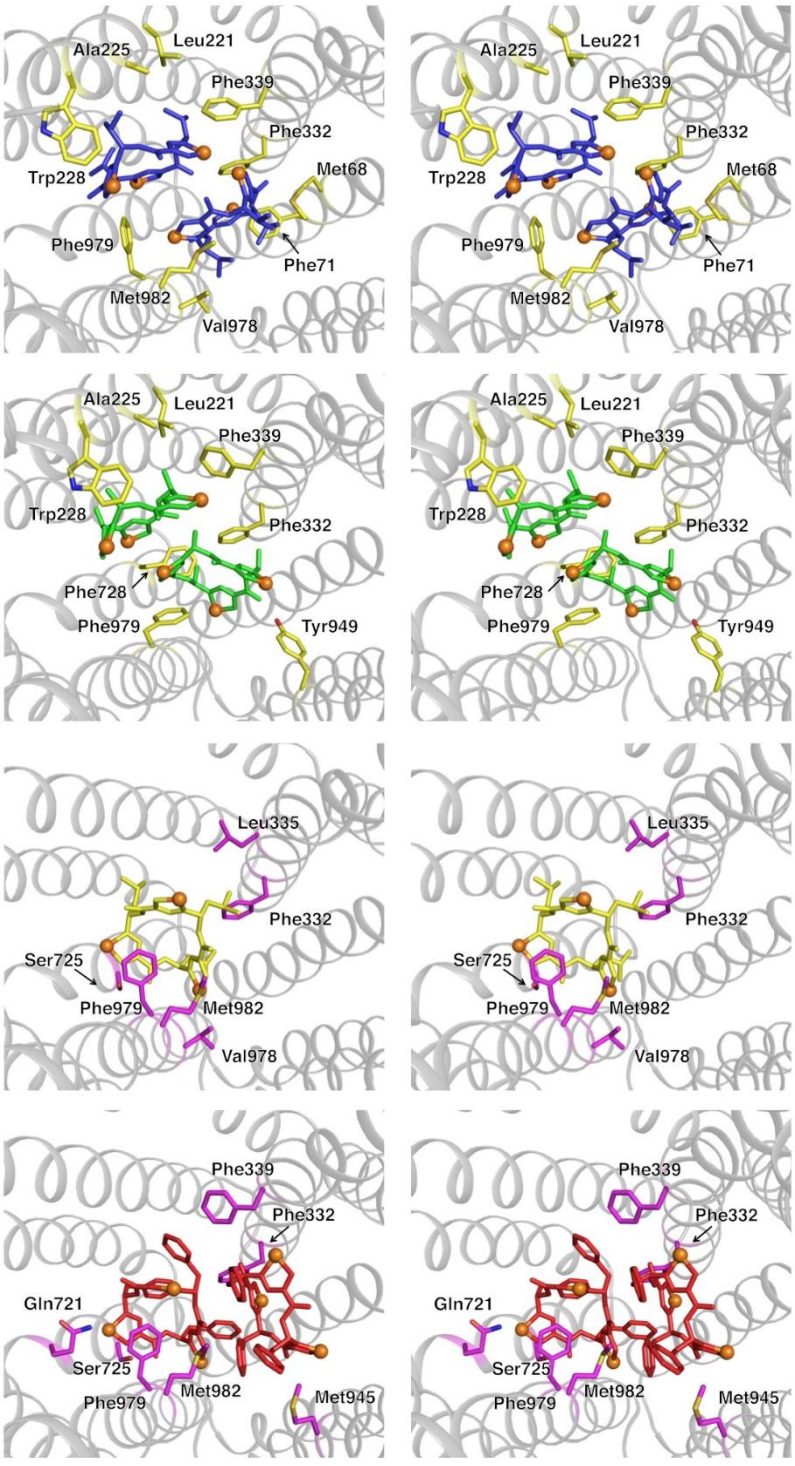



Figure 3.7: Key P-gp residues involved in binding homotrimeric cyclopeptides. Stereoviews of the binding pocket are shown perpendicular to the membrane and viewed from the cytosol. P-gp residues are shown as sticks, with those involved with binding Subset A ligands colored in yellow, and those involved with binding Subset B ligands in magenta.

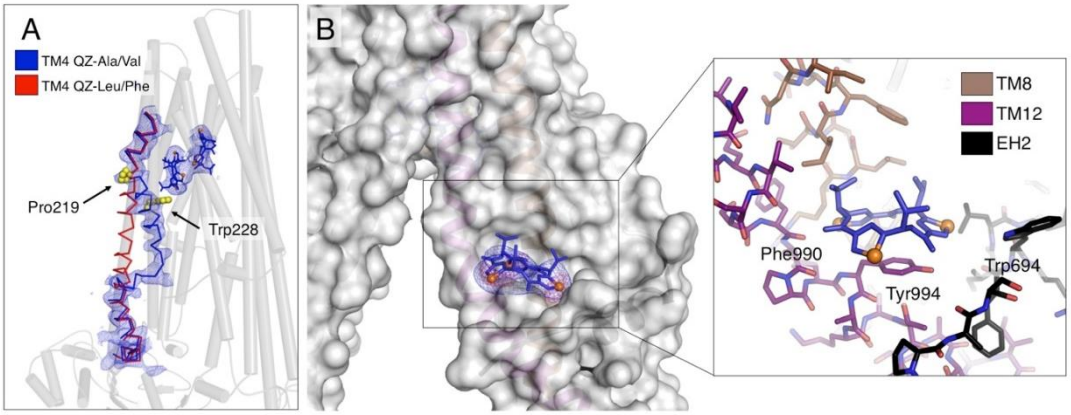



Figure 3.8: Overview of EH2 ligand binding site and ligand induced movement of TM4. (A) The kinking of TM4 in response to Subset A ligands is shown (blue ribbon) in comparison to its 'straight' topology in the Subset B co-crystal structures (red ribbon). The position of QZ-Val ligands and the resulting 2m$F_o$-D$F_c$ electron density (contoured at 1.0σ) are displayed. Residues Pro219 and Trp228 are as yellow spheres. (B) QZ-Val bound at the elbow-helix 2 site with P-gp rendered as a molecular surface representation. The resulting 2m$F_o$-D$F_c$ electron density for the ligand is shown at a sigma contour level of 1.0σ and the resulting anomalous difference peaks for the seleniums (orange spheres) are shown at 3.5σ. The close-up inset view depicts the surrounding TMs 8 and 12 as brown and purple sticks, respectively, and elbow helix 2 as black sticks. Residues in close vicinity to the ligand (Trp694, Phe990 and Tyr994) are labelled.

Table 3.1: Data Collection and Refinement Statistics

| | 3.4 Å | QZ-Ala | QZ-Val | QZ-Leu | QZ-Phe |
|---|---|---|---|---|---|
| **Data collection** | | | | | |
| Wavelength | 0.9795 | 0.9796 | 0.9794 | 0.9785 | 0.9796 |
| Beamline | SSRL11-1 | CLS08ID | CLS08ID | SSRL11-1 | SSRL11-1 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ |
| Cell dimensions | | | | | |
| *a*, *b*, *c* (Å) | 88.0, 139.2, 186.0 | 86.0, 138.7, 184.1 | 88.7, 138.7, 190.0 | 91.3, 138.5, 195.6 | 86.3, 138.4, 185.0 |
| $\alpha, \beta, \gamma$ (°) | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 |
| Resolution (Å) | 3.4 | 3.8 | 3.6 | 3.8 | 3.8 |
| Number of crystals | 1 | 2 | 1 | 1 | 1 |
| $R_{merge}$ (%) | 8.4 (73.0)* | 9.3 (85.3) | 7.9 (64.9) | 6.5 (88.4) | 5.8 (69.6) |
| $R_{pim}$ | 4.4 (38.3) | 3.5 (37.0) | 4.1 (35.0) | 3.3 (43.9) | 3.1 (36.6) |
| Mean ((I/sd(I)) | 8.0 (1.8) | 11.2 (2.6) | 8.3 (2.0) | 10.8 (1.9) | 10.8 (2.2) |
| Completeness (%) | 99.4(100) | 98.3 (99.4) | 97.4 (98.5) | 94.5 (97.7) | 99.4 (100) |
| Redundancy | 4.7 (4.6) | 8.0 (6.7) | 4.5 (4.3) | 3.7 (3.8) | 4.4 (4.4) |
| | | | | | |
| **Refinement** | | | | | |
| Resolution (Å) | 3.4 | 3.8 | 3.6 | 3.8 | 3.8 |
| $R_{work/}$ $R_{free}$ | 26.8/29.4 | 25.2/28.7 | 26.1/28.3 | 28.8/31.8 | 25.4/29.2 |
| R.m.s deviations | | | | | |
| Bond lengths (Å) | 0.005 | 0.004 | 0.004 | 0.005 | 0.005 |
| Bond angles (°) | 0.904 | 0.764 | 0.801 | 0.849 | 0.822 |
| Ramachandran statistics | | | | | |
| Outliers % | 0 | 0 | 0 | 0 | 0 |
| Favoured % | 95.93 | 95.76 | 95.27 | 95.52 | 95.69 |
| Allowed % | 4.07 | 4.24 | 4.73 | 4.48 | 4.31 |
| Rotomer | 0.92 | 0.36 | 0.51 | 0.41 | 0.20 |
| Outliers % | | | | | |
| Cβ Deviations | 0 | 0 | 0 | 0 | 0 |

*Highest resolution shell is shown in parenthesis.

Table 3.2: Solvent Accessible Surface Area of Cyclopeptide Ligands Buried in the Pocket of P-gp as Calculated by PISA(Krissinel & Henrick, 2007)

| | Accessible Surface Area (Å²) | Buried Surface Area (Å²) |
|---|---|---|
| QZ-Ala Upper | 628.7 | 451.6 |
| QZ-Ala Lower | 623.3 | 409.1 |
| QZ-Val Upper | 723.1 | 513.6 |
| QZ-Val Lower | 690.5 | 475.4 |
| QZ-Val Outer | 723.9 | 426.7 |
| QZ-Leu Upper | 824.1 | 577.5 |
| QZ-Phe Upper | 760.4 | 577.2 |
| QZ-Phe Lower | 783.9 | 474.8 |

References:

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., . . . Zwart, P. H. (2010a). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 2), 213-221.

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., . . . Zwart, P. H. (2010b). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography, 66*(Pt 2), 213-221.

Al-Shawi, M., & Omote, H. (2005). The remarkable transport mechanism of P-glycoprotein: A multidrug transporter. *Journal of Bioenergetics and Biomembranes, 37*(6), 489-496.

al-Shawi, M. K., & Senior, A. E. (1993). Characterization of the adenosine triphosphatase activity of Chinese hamster P-glycoprotein. *J Biol Chem, 268*(6), 4197-4206.

Aller, S. G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., . . . Chang, G. (2009). Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science, 323*(5922), 1718-1722.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res, 25*(17), 3389-3402.

Altschul, S. F., Wootton, J. C., Gertz, E. M., Agarwala, R., Morgulis, A., Schaffer, A. A., & Yu, Y. K. (2005). Protein database searches using compositionally adjusted substitution matrices. *FEBS J, 272*(20), 5101-5109.

Bai, J., Swartz, D. J., Protasevich, II, Brouillette, C. G., Harrell, P. M., Hildebrandt, E., . . . Urbatsch, I. L. (2011). A gene optimization strategy that enhances production of fully functional P-glycoprotein in Pichia pastoris. *PLoS One, 6*(8), e22577.

Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., & Leslie, A. G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography, 67*(Pt 4), 271-281.

Becker, M. L., Visser, L. E., van Schaik, R. H., Hofman, A., Uitterlinden, A. G., & Stricker, B. H. (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.

Bolhuis, H., van Veen, H. W., Molenaar, D., Poolman, B., Driessen, A. J., & Konings, W. N. (1996). Multidrug resistance in *Lactococcus lactis*: evidence for ATP-dependent drug extrusion from the inner leaflet of the cytoplasmic membrane. *The EMBO journal, 15*(16), 4239-4245.

Brunger, A. T. (2007). Version 1.2 of the Crystallography and NMR system. *Nat Protoc, 2*(11), 2728-2733.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., . . . Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr, 54*(Pt 5), 905-921.

Cascorbi, I. (2006). Role of pharmacogenetics of ATP-binding cassette transporters in the pharmacokinetics of drugs. *Pharmacology & Therapeutics, 112*(2), 457-473.

Cascorbi, I. (2011). P-glycoprotein: tissue distribution, substrates, and functional consequences of genetic variations. *Handb Exp Pharmacol*(201), 261-283.

Chen, J., Sharma, S., Quiocho, F. A., & Davidson, A. L. (2001). Trapping the transition state of an ATP-binding cassette transporter: evidence for a concerted mechanism of maltose transport. *Proc Natl Acad Sci U S A, 98*(4), 1525-1530.

Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., . . . Richardson, D. C. (2010a). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 1), 12-21.

Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., . . . Richardson, D. C. (2010b). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography, 66*(Pt 1), 12-21. doi: 10.1107/S0907444909042073

Chen, Y. J., Pornillos, O., Lieu, S., Ma, C., Chen, A. P., & Chang, G. (2007). X-ray structure of EmrE supports dual topology model. *Proc Natl Acad Sci U S A, 104*(48), 18999-19004.

Conrath, K. E., Lauwereys, M., Galleni, M., Matagne, A., Frere, J. M., Kinne, J., . . . Muyldermans, S. (2001). Beta-lactamase inhibitors derived from single-

domain antibody fragments elicited in the camelidae. *Antimicrob Agents Chemother, 45*(10), 2807-2812.

Crowley, E., McDevitt, C. A., & Callaghan, R. (2010). Generating inhibitors of P-glycoprotein: where to, now? *Methods Mol Biol, 596*, 405-432.

De Genst, E., Handelberg, F., Van Meirhaeghe, A., Vynck, S., Loris, R., Wyns, L., & Muyldermans, S. (2004). Chemical basis for the affinity maturation of a camel single domain antibody. *J Biol Chem, 279*(51), 53593-53601.

de Graaf, D., Sharma, R. C., Mechetner, E. B., Schimke, R. T., & Roninson, I. B. (1996). P-glycoprotein confers methotrexate resistance in 3T6 cells with deficient carrier-mediated methotrexate uptake. *Proceedings of the National Academy of Sciences, 93*(3), 1238-1242.

DeLano, W. L. (2009).

Deshayes, S., Konate, K., Aldrian, G., Crombez, L., Heitz, F., & Divita, G. (2010). Structural polymorphism of non-covalent peptide-based delivery systems: highway to cellular uptake. *Biochim Biophys Acta, 1798*(12), 2304-2314.

Dey, S., Ramachandra, M., Pastan, I., Gottesman, M. M., & Ambudkar, S. V. (1997). Evidence for two nonidentical drug-interaction sites in the human P-glycoprotein. *Proceedings of the National Academy of Sciences, 94*(20), 10594-10599.

Doshi, R., & van Veen, H. W. (2013). Substrate binding stabilizes a pre-translocation intermediate in the ATP-binding cassette transport protein MsbA. *Journal of Biological Chemistry, 288*(30), 21638-21647.

Doshi, R., Woebking, B., & van Veen, H. W. (2010). Dissection of the conformational cycle of the multidrug/lipidA ABC exporter MsbA. *Proteins: Structure, Function, and Bioinformatics, 78*(14), 2867-2872.

Eckford, P. D., & Sharom, F. J. (2009). ABC efflux pump-based resistance to chemotherapy drugs. *Chem Rev, 109*(7), 2989-3011.

Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography, 60*, 2126-2132.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010a). Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography, 66*(Pt 4), 486-501.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010b). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 4), 486-501.

Evans, P. (2006). Scaling and assessment of data quality. *Acta crystallographica. Section D, Biological crystallography, 62*(Pt 1), 72-82.

Falasca, M., & Linton, K. J. (2012). Investigational ABC transporter inhibitors. *Expert Opinion on Investigational Drugs, 21*(5), 657-666.

Furey, W., & Swaminathan, S. (1997). PHASES-95: a program package for processing and analyzing diffraction data from macromolecules. *Methods Enzymol, 277*, 590-620.

Giacomini, K. M., Huang, S. M., Tweedie, D. J., Benet, L. Z., Brouwer, K. L., Chu, X., . . . Zhang, L. (2010). Membrane transporters in drug development. *Nat Rev Drug Discov, 9*(3), 215-236.

Glaser, F., Rosenberg, Y., Kessel, A., Pupko, T., & Ben-Tal, N. (2005). The ConSurf-HSSP database: The mapping of evolutionary conservation among homologs onto PDB structures. *Proteins-Structure Function and Bioinformatics, 58*(3), 610-617. doi: Doi 10.1002/Prot.20305

Goldenberg, O., Erez, E., Nimrod, G., & Ben-Tal, N. (2009). The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. *Nucleic Acids Research, 37*, D323-D327.

Gottesman, M. M., & Ling, V. (2006). The molecular basis of multidrug resistance in cancer: the early years of P-glycoprotein research. *FEBS Lett, 580*(4), 998-1009.

Gruol, D. J., King, M. N., & Kuehne, M. E. (2002). Evidence for the locations of distinct steroid and Vinca alkaloid interaction domains within the murine mdr1b P-glycoprotein. *Mol Pharmacol, 62*(5), 1238-1248.

Gutmann, D. A. P., Ward, A., Urbatsch, I. L., Chang, G., & van Veen, H. W. (2010). Understanding polyspecificity of multidrug ABC transporters: closing in on the gaps in ABCB1. *Trends in biochemical sciences, 35*(1), 36-42.

Hellmich, U. A., Lyubenova, S., Kaltenborn, E., Doshi, R., van Veen, H. W., Prisner, T. F., & Glaubitz, C. (2012). Probing the ATP hydrolysis cycle of the ABC multidrug transporter LmrA by pulsed EPR spectroscopy. *Journal of the American Chemical Society, 134*(13), 5857-5862.

Hohl, M., Briand, C., Grutter, M. G., & Seeger, M. A. (2012a). Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nature structural & molecular biology, 19*(4), 395-402.

Hohl, M., Briand, C., Grutter, M. G., & Seeger, M. A. (2012b). Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nat Struct Mol Biol, 19*(4), 395-402.

Hvorup, R. N., Winnen, B., Chang, A. B., Jiang, Y., Zhou, X. F., & Saier, M. H. (2003). The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. *European Journal of Biochemistry, 270*(5), 799-813.

Jardetzk.O. (1966). Simple Allosteric Model for Membrane Pumps. *Nature, 211*(5052), 969-&.

Jin, M. S., Oldham, M. L., Zhang, Q., & Chen, J. (2012). Crystal structure of the multidrug transporter P-glycoprotein from Caenorhabditis elegans. *Nature, 490*(7421), 566-569.

Juliano, R. L., & Ling, V. (1976). A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim Biophys Acta, 455*(1), 152-162.

Kaatz, G. W., McAleese, F., & Seo, S. M. (2005). Multidrug Resistance in Staphylococcus aureus Due to Overexpression of a Novel Multidrug and Toxin Extrusion (MATE) Transport Protein. *Antimicrob. Agents Chemother., 49*(5), 1857-1864.

Kodan, A., Yamaguchi, T., Nakatsu, T., Sakiyama, K., Hipolito, C. J., Fujioka, A., . . . Kato, H. (2014). Structural basis for gating mechanisms of a eukaryotic P-glycoprotein homolog. *Proceedings of the National Academy of Sciences of the United States of America, 111*(11), 4049-4054.

Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology, 372*(3), 774-797.

Kuhnke, D., Jedlitschky, G., Grube, M., Krohn, M., Jucker, M., Mosyagin, I., . . . Vogelgesang, S. (2007). MDR1-P-Glycoprotein (ABCB1) Mediates Transport of Alzheimer's amyloid-beta peptides--implications for the mechanisms of Abeta clearance at the blood-brain barrier. *Brain Pathol, 17*(4), 347-353.

Lam, A. Y., Pardon, E., Korotkov, K. V., Hol, W. G., & Steyaert, J. (2009). Nanobody-aided structure determination of the EpsI:EpsJ pseudopilin heterodimer from Vibrio vulnificus. *J Struct Biol, 166*(1), 8-15.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., & Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Research, 33*, W299-W302.

Lee, C. A., Cook, J. A., Reyner, E. L., & Smith, D. A. (2010). P-glycoprotein related drug interactions: clinical importance and a consideration of disease states. *Expert Opin Drug Metab Toxicol, 6*(5), 603-619.

Leslie, A. G. W. (1992). Joint CCP4 + ESF-EAMCB. *Newsletter on Protein Crystallography, 26*.

Li, J., Jaimes, K. F., & Aller, S. G. (2014). Refined structures of mouse P-glycoprotein. *Protein science : a publication of the Protein Society, 23*(1), 34-46.

Li, X. Z., Poole, K., & Nikaido, H. (2003). Contributions of MexAB-OprM and an EmrE homolog to intrinsic resistance of Pseudomonas aeruginosa to aminoglycosides and dyes. *Antimicrob Agents Chemother, 47*(1), 27-33.

Loo, T. W., Bartlett, M. C., & Clarke, D. M. (2003a). Drug binding in human P-glycoprotein causes conformational changes in both nucleotide-binding domains. *Journal of Biological Chemistry, 278*(3), 1575-1578.

Loo, T. W., Bartlett, M. C., & Clarke, D. M. (2003b). Substrate-induced conformational changes in the transmembrane segments of human P-glycoprotein: Direct evidence for the substrate-induced fit mechanism for drug binding. *Journal of Biological Chemistry, 278*(16), 13603-13606.

Loo, T. W., Bartlett, M. C., & Clarke, D. M. (2013). Human P-glycoprotein contains a greasy ball-and-socket joint at the second transmission interface. *Journal of Biological Chemistry, 288*(28), 20326-20333.

Loo, T. W., & Clarke, D. M. (1994). Mutations to amino acids located in predicted transmembrane segment 6 (TM6) modulate the activity and substrate specificity of human P-glycoprotein. *Biochemistry, 33*(47), 14049-14057.

Loo, T. W., & Clarke, D. M. (2005). Do drug substrates enter the common drug-binding pocket of P-glycoprotein through "gates"? *Biochemical and biophysical research communications, 329*(2), 419-422.

Marquez, B., & Van Bambeke, F. (2011). ABC multidrug transporters: target for modulation of drug pharmacokinetics and drug-drug interactions. *Curr Drug Targets, 12*(5), 600-620.

Martin, C., Berridge, G., Higgins, C. F., Mistry, P., Charlton, P., & Callaghan, R. (2000). Communication between multiple drug binding sites on P-glycoprotein. *Molecular Pharmacology, 58*(3), 624-632.

Martin, C., Berridge, G., Mistry, P., Higgins, C., Charlton, P., & Callaghan, R. (1999). The molecular interaction of the high affinity reversal agent XR9576 with P-glycoprotein. *Br J Pharmacol, 128*(2), 403-411.

McAleese, F., Petersen, P., Ruzin, A., Dunman, P. M., Murphy, E., Projan, S. J., & Bradford, P. A. (2005). A Novel MATE Family Efflux Pump Contributes to the Reduced Susceptibility of Laboratory-Derived Staphylococcus aureus Mutants to Tigecycline. *Antimicrob. Agents Chemother., 49*(5), 1865-1871.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography, 40*(Pt 4), 658-674. doi: 10.1107/S0021889807021206

Mechetner, E. B., Schott, B., Morse, B. S., Stein, W. D., Druley, T., Davis, K. A., . . . Roninson, I. B. (1997). P-glycoprotein function involves conformational transitions detectable by differential immunoreactivity. *Proceedings of the National Academy of Sciences, 94*(24), 12908-12913.

Mehmood, S., Domene, C., Forest, E., & Jault, J.-M. (2012). Dynamics of a bacterial multidrug ABC transporter in the inward- and outward-facing conformations. *Proceedings of the National Academy of Sciences, 109*(27), 10832-10836.

Morita, M., Shitan, N., Sawada, K., Van Montagu, M. C., Inze, D., Rischer, H., . . . Yazaki, K. (2009). Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in Nicotiana tabacum. *Proc Natl Acad Sci U S A, 106*(7), 2447-2452.

Murakami, S., Nakashima, R., Yamashita, E., & Yamaguchi, A. (2002). Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature, 419*(6907), 587-593.

Muyldermans, S., Baral, T. N., Retamozzo, V. C., De Baetselier, P., De Genst, E., Kinne, J., . . . Saerens, D. (2009). Camelid immunoglobulins and nanobody technology. *Vet Immunol Immunopathol, 128*(1-3), 178-183.

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

Otsuka, M., Yasuda, M., Morita, Y., Otsuka, C., Tsuchiya, T., Omote, H., & Moriyama, Y. (2005). Identification of essential amino acid residues of the

NorM Na+/multidrug antiporter in Vibrio parahaemolyticus. *J Bacteriol, 187*(5), 1552-1558.

Otwinowski, Z., & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A, 276*, 307-326.

Park, Y. J., Pardon, E., Wu, M., Steyaert, J., & Hol, W. G. (2012). Crystal structure of a heterodimer of editosome interaction proteins in complex with two copies of a cross-reacting nanobody. *Nucleic Acids Res, 40*(4), 1828-1840.

Raviv, Y., Pollard, H. B., Bruggemann, E. P., Pastan, I., & Gottesman, M. M. (1990). Photosensitized labeling of a functional multidrug transporter in living drug-resistant tumor cells. *Journal of Biological Chemistry, 265*(7), 3975-3980.

Rayment, I. (1997). Reductive alkylation of lysine residues to alter crystallization properties of proteins. *Macromolecular Crystallography, Pt A, 276*, 171-179.

S. Wood, K. L. S., S. J. Scherr (2000). *Pilot analysis of global ecosystems : agroecosystems.* (Vol. xii): World Resources Institute, Washington, D.C., 2000.

Sack, J. S. (1988). Chain - a Crystallographic Modeling Program. *Journal of Molecular Graphics, 6*(4), 224-225.

Schinkel, A. H., & Jonker, J. W. (2003). Mammalian drug efflux transporters of the ATP binding cassette (ABC) family: an overview. *Adv Drug Deliv Rev, 55*(1), 3-29.

Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res, 31*(13), 3381-3385.

Seeger, M. A., & van Veen, H. W. (2009). Molecular basis of multidrug transport by ABC transporters. *Biochim Biophys Acta, 1794*(5), 725-737.

Shapiro, A. B., & Ling, V. (1997). Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities. *European journal of biochemistry / FEBS, 250*(1), 130-137.

Shintre, C. A., Pike, A. C., Li, Q., Kim, J. I., Barr, A. J., Goubin, S., . . . Carpenter, E. P. (2013). Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proc Natl Acad Sci U S A, 110*(24), 9710-9715.

Shintre, C. A., Pike, A. C. W., Li, Q., Kim, J.-I., Barr, A. J., Goubin, S., . . . Carpenter, E. P. (2013). Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proceedings of the National Academy of Sciences, 110*(24), 9710-9715.

Shukla, S., Ohnuma, S., & Ambudkar, S. V. (2011). Improving cancer chemotherapy with modulators of ABC drug transporters. *Curr Drug Targets, 12*(5), 621-630.

Smriti, Zou, P., & McHaourab, H. S. (2009). Mapping daunorubicin-binding Sites in the ATP-binding cassette transporter MsbA using site-specific quenching by spin labels. *J Biol Chem, 284*(20), 13904-13913. doi: 10.1074/jbc.M900837200

Tamaki, A., Ierano, C., Szakacs, G., Robey, R. W., & Bates, S. E. (2011). The controversial role of ABC transporters in clinical oncology. *Essays Biochem, 50*(1), 209-232.

Tao, H., Weng, Y., Zhuo, R., Chang, G., Urbatsch, I. L., & Zhang, Q. (2011). Design and synthesis of Selenazole-containing peptides for cocrystallization with P-glycoprotein. *Chembiochem : a European journal of chemical biology, 12*(6), 868-873.

Terwilliger, T. C. (2000). Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr, 56*(Pt 8), 965-972.

Terwilliger, T. C., & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr, 55*(Pt 4), 849-861.

Tombline, G., Urbatsch, I. L., Virk, N., Muharemagic, A., White, L. B., & Senior, A. E. (2006). Expression, purification, and characterization of cysteine-free mouse P-glycoprotein. *Arch Biochem Biophys, 445*(1), 124-128.

Tsuda, M., Terada, T., Mizuno, T., Katsura, T., Shimakura, J., & Inui, K. (2009). Targeted Disruption of the Multidrug and Toxin Extrusion 1 (Mate1) Gene in Mice Reduces Renal Secretion of Metformin. *Molecular Pharmacology, 75*(6), 1280-1286.

U.S. Food and Drug Administration, C. D. E. R. (2012). *Guidance for industry: drug interaction studies - study design, data analysis, implications for dosing, and labeling recommendations*. Retrieved from http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm292362.pdf.

Urbatsch, I. L., Gimi, K., Wilke-Mounts, S., & Senior, A. E. (2000). Conserved walker A Ser residues in the catalytic sites of P-glycoprotein are critical for catalysis and involved primarily at the transition state step. *J Biol Chem, 275*(32), 25031-25038.

Urbatsch, I. L., Sankaran, B., Bhagat, S., & Senior, A. E. (1995). Both P-glycoprotein nucleotide-binding sites are catalytically active. *J Biol Chem, 270*(45), 26956-26961.

Urbatsch, I. L., Sankaran, B., Weber, J., & Senior, A. E. (1995). P-glycoprotein is stably inhibited by vanadate-induced trapping of nucleotide at a single catalytic site. *J Biol Chem, 270*(33), 19383-19390.

Vichai, V., & Kirtikara, K. (2006). Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nature protocols, 1*(3), 1112-1116.

Vogel, G., & Steinhart, R. (1976). ATPase of Escherichia coli: purification, dissociation, and reconstitution of the active complex from the isolated subunits. *Biochemistry, 15*(1), 208-216.

Voss, N. R., Gerstein, M., Steitz, T. A., & Moore, P. B. (2006). The geometry of the ribosomal polypeptide exit tunnel. *Journal of Molecular Biology, 360*(4), 893-906.

Ward, A., Reyes, C. L., Yu, J., Roth, C. B., & Chang, G. (2007). Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci U S A, 104*(48), 19005-19010.

Ward, A., Reyes, C. L., Yu, J., Roth, C. B., & Chang, G. (2007). Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proceedings of the National Academy of Sciences, 104*(48), 19005-19010.

Ward, A. B., Szewczyk, P., Grimard, V., Lee, C. W., Martinez, L., Doshi, R., . . . Chang, G. (2013). Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain. *Proceedings of the National Academy of Sciences of the United States of America, 110*(33), 13386-13391.

Watkins, R. E., Wisely, G. B., Moore, L. B., Collins, J. L., Lambert, M. H., Williams, S. P., . . . Redinbo, M. R. (2001). The human nuclear xenobiotic receptor PXR: Structural determinants of directed promiscuity. *Science, 292*(5525), 2329-2333.

Wen, P. C., Verhalen, B., Wilkens, S., McHaourab, H., & Tajkhorshid, E. (2013). On the Origin of Large Flexibility of P-glycoprotein in the Inward-Facing State. *J Biol Chem*.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., . . . Wilson, K. S. (2011a). Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography, 67*(Pt 4), 235-242.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., . . . Wilson, K. S. (2011b). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr, 67*(Pt 4), 235-242.

Woebking, B., Velamakanni, S., Federici, L., Seeger, M. A., Murakami, S., & van Veen, H. W. (2008). Functional role of transmembrane helix 6 in drug binding and transport by the ABC transporter MsbA. *Biochemistry, 47*(41), 10904-10914.

Yin, Y., He, X., Szewczyk, P., Nguyen, T., & Chang, G. (2006). Structure of the multidrug transporter EmrD from Escherichia coli. *Science, 312*(5774), 741-744.

Zhang, Q. H., Ma, X. Q., Ward, A., Hong, W. X., Jaakola, V. P., Stevens, R. C., . . . Chang, G. (2007). Designing facial amphiphiles for the stabilization of integral membrane proteins. *Angewandte Chemie-International Edition, 46*(37), 7023-7025.

Zheleznova, E. E., Markham, P. N., Neyfakh, A. A., & Brennan, R. G. (1999). Structural basis of multidrug recognition by BmrR, a transcription activator of a multidrug transporter. *Cell, 96*(3), 353-362.

Zou, P., Bortolus, M., & McHaourab, H. S. (2009). Conformational cycle of the ABC transporter MsbA in liposomes: detailed analysis using double electron-electron resonance spectroscopy. *Journal of Molecular Biology, 393*(3), 586-597.