UNIVERSITY OF CALIFORNIA, SAN DIEGO

Think Positively: The Structural Basis of Cation-binding and Coupling of the Multidrug and Toxic-compound Extrusion (MATE) Transporter Family

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

in

Biology

by

Xiao He

Committee in charge:

Professor Kit Pogliano, Chair
Professor Geoffrey Chang
Professor Partho Ghosh
Professor Sanjay Nigam
Professor Milton Saier
Professor Susan Taylor

2010

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

_____

_____

_____

_____

_____

_____,
Chair

University of California, San Diego

2010

EPIGRAPH

In search of knowledge…

TABLE OF CONTENTS

List of Abbreviations

ASU—Asymmetric Unit

β-DDM—n-dodecyl-β-D-maltopyranoside

BDM—Baker's Dimercurial

Bp—Base pair

CCCP—Carbonyl Cyanide m-Chlorophenylhydrazone

C-ter—C-terminal

EtBr—Ethidium Bromide

Fo—Observed structure factors

Fc—Calculated structure factors

M—Molar

MATE—Multidrug and Toxic-compound Extrustion

MCS—Multiple Cloning Site

MFS—Major Facilitator Superfamily

MIC—Minimal Inhibition Concentration

MPP—1-methyl-4-phenylpyridinium

NMN—N^1-methylnicotinamide

TMS—Transmembrane Segment

NCS—Non-crystallographic Symmetry

N-ter—N-terminal

OCs—Organic Cations

OCT—Organic Cation Transporter

PAGE—Polyacrylamide Gel Electrophoresis

PEG—Polyethylene Glycol

PVP—Polyvinylpyrrolidone

RBS—Ribosome Binding Site

RND—Resistance, Nodulation, and Cell-division family

$R_{cryst} = \Sigma | Fo-Fc | / \Sigma | Fo |$, where Fo and Fc are observed and calculated structure factors, respectively.

$R_{free} = \Sigma | Fo-Fc | / \Sigma | Fo |$, where a subset of data (10%) is omitted from the refinement.

Rmsd—Root Mean Square Deviation

SA—Simulated Annealing

SDS—Sodium Dodecyl Sulfate

SNP—Single Nucleotide Polymorphism

TEA—Tetraethylammonium

TMA—Tetramethylammonium

List of Figures

List of Tables

Acknowledgements

I would like to thank Professor Kit Pogliano for her support as the chair of my committee. Through many sentimental discussions and multiple drafts, her guidance has been invaluable for this work. I would like to thank Professor Geoffrey Chang at TSRI, for his support and guidance in my training in the field of transporter X-ray crystallography. I would also like to acknowledge the MATE project members in Chang Lab. It is all these supports that helped me in an immeasurable way.

Chapter 1 and 2, in part, have been accepted for publication of the material as it may appear in the journal Nature, 2010, He, Xiao; Szewczyk, Paul; Karyanki, Andrey; Evin, Riah; Hong, Wen-Xue; Zhang, Qinghai; and Chang, Geoffrey, 2010. The dissertation author is the primary investigator and author of this paper.

Chapter 2, in part, and chapter 3 are currently being prepared for submission for publication of the material by He, Xiao; Chang, Geoffrey; and Pogliano, Kit. The dissertation author is the primary investigator and author of this material.

VITA

EDUCATIONS

2001 Bachelor of Science, Shandong University, Jinan, Shandong, PRC

2003 Master of Science, University of California, San Diego

2004-2006 Research assistant, Department of Molecular Biology, The Scripps Research Institute

2010 Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

"Structure of the Multidrug Transporter EmrD from Escherichia coli" Science, 312:741-744, May 2006

"Structure of A Cation-bound Multidrug and Toxic Compound Extrusion (MATE) Family Transporter" Nature, in press.

FIELD OF STUDY

Major Field: X-ray Crystallogarphy and Membrane Protein Structure

Studies in Membrane Protein X-ray Crystallography and Membrane Protein Biology

Professors Geoffrey Chang and Kit Pogliano

ABSTRACT OF THE DISSERTATION

Think Positively: The Structural Basis of Cation-binding and Coupling of the Multidrug and Toxic-compound Extrusion (MATE) Transporter Family

by

Xiao He

Doctor of Philosophy in Biology

University of California, San Diego, 2010

Professor Kit Pogliano, Chair

Transporters of the multidrug and toxic compound extrusion (MATE) family play vital roles in plant metabolite transport as well as bacterial and mammalian multidrug resistance (MDR). As secondary transporters, MATEs couple substrate transport to electrochemical gradients and are the last remaining class of MDR transporters whose structure has not been determined. The X-ray structure of the MATE

transporter NorM from *Vibrio cholera* (VC) determined to 3.65 Å reveals an outward-facing conformation with two portals open to the outer leaflet of the lipid bilayer and has a unique topology distinct from other known MDR transporters. A discrete cation-binding site in close proximity to residues critical for transport as reported previously is identified at the C-terminal half of the transporter. This conformation likely represents a stage of the transport cycle with high-affinity to monovalent cations and low-affinity to substrates. Homology modeling of other distantly related multidrug MATE transporters indicates the conservation of this cation-binding site, suggesting a similar functional attribute.

A comparison between the two prevalent secondary transporter families, MATE and MFS, revealed the similarities and differences between the structure and function of the two transporter families. The MATE and MFS families have distinct protein-folds and topologies, and use residues from different transmembrane (TM) helices for cation binding and, likely, substrate recognition. These findings suggest that the cation binding and the coupling mechanism to substrate efflux in MATE and MFS families likely arose independently. With the most conserved residues still facing the internal cavity in the low-substrate-

affinity conformation, the MATEs likely transport substrates/cations using a rocker-switch model similar to the MFS. In this alternating access model, the access and binding of substrates/cations from the two sides of the lipid membrane are mediated by rigid-body movements between the N- and C-terminal halves of the transporter. The substrate and cation binding affinity of each conformational state depends on the spatial rearrangement of a similar set of residues within the internal cavity of the transporter.

Introduction

The export of xenobiotics and natural products from the cell is a fundamental life process. This process is primarily achieved through the activities of integral membrane proteins that are classified in two major categories based on the energy source that powers their transport process: the ATP-cassette-binding (ABC) transporters and Solute carriers (SLC). Unlike the ABC transporters, which are powered by ATP binding and hydrolysis, the SLCs comprise of a large collection of structurally distinct families of membrane proteins that uses the electrochemical gradient across the lipid membrane to drive substrate transport.

Four SLC families encode multidrug type of transporters: the small multidrug resistance (SMR) family, the resistance, nodulation, and cell-division (RND) family, the MDR transporters in the major facilitator superfamily (MFS) and the multidrug and toxic-compound extrusion (MATE) families. Among these, the MATE family is less understood in terms of transport mechanisms and structure. Although never formally classified together, the MFS and the MATE transporters are similar in certain aspects. Comparable molecular mass and hydrophobicity propensity with 12-14 predicted transmembrane (TM) helices and a pseudo 2-fold symmetry between the N- and C-terminal halves even from primary amino acid

sequences' point of views are among the most readily observed features. Unlike MFS transporter, which has been extensively studied for decades, both functionally and structurally, research on MATE transporters is more recent.[1] MATE family is only MDR transporter family whose structure has yet to be determined[2,3,4,5,6,7].

As SLCs, MATE transporters use proton gradient across the lipid membrane to drive the efflux of a variety of structurally unrelated organic cations. Along with the MFS[8], MATE transporters are among the most abundant SLC types and play vital roles in eukaryotes. Functionally speaking, some MATE and MFS transporters even share similar substrate spectra, which, in some cases, is required for their synergistic action to eliminate and/or sequestrate toxic compounds [9]. A detailed understanding of the structural basis for substrate binding and transport mechanism is important for the understanding and modulation of SLC functions. The structural study of MATE family of transporters is highly anticipated as it enables a comparative method to explore the similarities and differences, both structurally and functionally, between the MDR transporter families.

MATEs are involved in various biologically important processes in all cellular organisms, as manifested by their universal presence in all eukaryotes, eubacteria, and archaea. MATE transporters can be sub-divided into three clades. These three clades are the eubacterial NorM-like MDR transporters, the eukaryotic MATE transporters, and the eubacterial/archeal VmrA-like MDR transporters.

The prototypical bacterial NorM (norfloxacin-resistance gene M) is one of the earliest identified and one of the best characterized MATE transporters [1,10]. Clade I comprises of mainly the NorM orthologs in Gram-negative bacteria. These orthologs are either proton (H^+) or sodium ion (Na^+)-driven multidrug resistance (MDR) efflux pumps[11]. The second clade comprises MATE transporters found in eukaryotes which can be further divided into four subgroups: (1) fungi, (2) plants, (3) animals, and (4) protozoa[12]. Although not all MATE transporters from this clade have been fully characterized, it is believe that they are H^+-gradient driven antiporters. The third clade includes the group of VmrA-like transporters that are relatively more distantly related to clade I or II and only found in archaea and certain eubacteria. Gram-negative bacteria often have MATEs from both clade I and clade III, while Gram-positive only have those belonging to clade III.

Bacterial MATE transporters function as multidrug resistance (MDR) efflux pumps. As the earliest identified MATE homologs involved in MDR, bacterial MATE transporters function to increase resistance towards a range of xenobiotic substrates and many DNA intercalators[11], posing an increasing threat in the treatment of infectious disease[12]. NorM, for instance, is the acronym for the norfloxacin-resistance gene M and causes increased minimal inhibition concentration (MIC) of a variety of substrates, especially fluoroquinolone class of antibiotics and many organic cationic dyes, in expression hosts [11]. In Gram-negative human pathogens such as Haemophilus influenzae[13] and *V. cholerae*[14], clade I MATE transporters have been shown to confer resistance primarily to the fluoroquinolone class of antibiotics[13]. In phytopathogen Erwinia amylovara, NorM ortholog enhance the pathogen's resistance to xenobiotics excreted from other microbes present in the host environment[15]. MATE transporters from clade III also function primarily as xenobiotic efflux pumps and can, for example, increase resistance to tigecycline, a new glycylcycline class antibiotic developed to overcome methicillian- and vancomycin-resistant *Staphylococcus aureus*[16,17]. These MDR MATE and transporters are expressed in the bacterial plasma membrane, directly mediating the

efflux of xenobiotics from the bacterial cytosol in a H^+ or Na^+ gradient-dependent manner[11].

In plants, MATE transporters are known to move metabolites between plant tissues[18], secrete organic compounds[19], as well as efflux xenobiotics in roots[20]. Plants posses the largest collection of MATE transporters, with 58 paralogs existing in *Arabidopsis thaliana*[21] alone, secreting a diverse range of metabolites as a defense against herbivores/microbial pathogens and protectants against harsh environment[22,23,24]. In plant, MATE transporters with highly specific substrate are often localized in the plasma membrane of the source cell and the vacuole membranes of the sink cell [23,25,26,27]. Some MATE transporter has been indicated to reside at the xylem tissues and transport minerals critical for crop growth[28]. And the multidrug MATE transporters are often expressed in the root tissue to limit toxic compounds entering from soil and promote root growth[20].

A. thaliana TT12 is a MATE transporter that transport glycosylated flavan-3-ols from the cytosol to vacuole in proanthocyanidin-accumulating cells of the seed coat[25]. In grapevine Vitis vinifera, two other MATE transporters anthoMATE1 (AM1) and AM3 specifically transport

acylated anthocyanin, a precursor of red/blue/purple pigments in grape skin, in the fruit vacuole[27]. The tobacco plant MATE homolog Nt-JAT1 is the primary efflux pump that is responsible for nicotine accumulation in leaf vacuole in *Nicotiana tabacum*[29].

Plant MATEs also play an important role in phytotoxic aluminum tolerance in acidic soils, a major limitation of crop production in 50% of the world's arable lands[19,30]. In *Sorghum bicolor*, SbMATE is a final effector gene in the aluminum tolerance response that excrete primary metabolite citric acid in root tissue to chelate aluminum ion (Al^{3+}) and alleviate further toxicity[31]. In rye, barley, maize, rice, and wheat, MATE transporters are critical for aluminum tolerance response in a similar fashion[28,32,33,34,35]. Another closely related paralog of these aluminum tolerance effectors, the FRD3 in *A. thaliana*, OsFRDL1 in *Oryza sativa*, and ScFRDL1 in *Secale cereale*, were thought to efflux the same substrate citrate into the xylem instead, for Fe translocation from roots to shoots[28,33,36].

Plants also use multidrug MATE transporters to extrude toxic compounds from root tissues and promote root growth as well. Diener et. al. reported a MDR MATE ALF5 in *A. thaliana* is essential for excretion of a several toxic organic cations and media contaminants[20]. AtDXT1 is

another MDR MATE efflux pumps in *A. thaliana* that extrude plant-derived alkaloids, antibiotics, and other exogenous toxic compounds, including cadmium ion (Cd^{2+})[37].

Mammalian MATE transporters guard the final excretion step of natural wastes and toxic xenobiotics in kidney and liver[12], a process that influences the plasma concentrations of many pharmaceutically valuable drugs, including metformin, a widely prescribed drug for Type 2 diabetes, thereby mitigating therapeutic efficacy[38,39]. Compared to plants, animals have far fewer MATE homologs in their genomes. Single nucleotide polymorphism (SNP) in human (h) MATE1 was also linked to altered drug disposition and drug response of several other clinical drugs[40]. In rats, rMATE1 on kidney luminal plasma membrane was indicated as the key transporter for tubular secretion of cationic drug cimetidine[41]. Compared to mammalian MATEs, the function of protozoa and fungi MATE transporters is currently unknown despite multiple copies of MATE homologs exist in their genomes. In some other parasitic genera like *Plasmodium*, MATEs seem to be absent from the genome[12].

All MATE proteins share ~40% protein sequence similarity[22], suggesting an overall conserved structure and transport function. MATE

transporters use either H^+ or Na^+ gradients across the membrane to drive substrate export [11], however, the coupling mechanism is not well understood. Although all MATE transporters characterized to date are efflux antiporters, their substrate specificities can vary considerably, ranging from very narrow to highly promiscuous. MATE transporters in clade I and clade II have been known to carry out both substrate efflux as well as uptake. The direction of substrate transport is opposite to the direction of the cation gradient[42,43].

To provide a basis for understanding the transport mechanisms of MATE transporters, the prototypical Na+-driven MDR transporter, NorM from *V. cholera* (NorM-VC) was determined by x-ray crystallography. The X-ray structures of NorM-VC with and without the heavier Na^+ analog rubidium ion (Rb^+), revealed a structurally discrete cation-binding site inthe C-ter half of the internal cavity of the transporter composed of residues from TM 7, 8, 10, 11, and 12. The corresponding residues near the cation site in NorM-VC was also found critical for transport in a closely related ortholog NorM from *V. parahaemolyticus* (NorM-VP)[42]. Substitution mutants of a relatively conserved residues in the cation-binding site abolish the cation binding in the outward-facing conformation of NorM-VC. These results correlate with loss of transport phenotypes reported in

the equivalent mutants in NorM-VP[42], suggesting the cation-binding site found in NorM-VC structure is likely the directly coupling site of substrate and cation transport in these MATE transporters. Homology models of several distantly related multidrug MATE transporters also suggest that this cation-binding site is likely to be structurally and functionally conserved among all multidrug type of MATE transporters.

Structural comparisons between NorM-VC and the multidrug MFS transporter, EmrD from *E. coli* (EmrD-EC)[7], provide insight into the structural and functional differences between MATE and MFS, two prominent families of SLC transporters in eukaryotes. Structural elements that are important for the transport functions are located at where the major topological differences are found between the two transport families.

1

Structural Determination of NorM-VC

1.1 Molecular Cloning and Expression of Recombinant MATE Transporters

We used a species ortholog approach for membrane protein X-ray crystallography to maximize the probability of a successful structural determination of a MATE transporter. This approach scans multiple bacterial MATE homologs for expression and purification. In brief, MATE orthologs were cloned into either T7 or pBAD promoter-driven expression vectors. Then each expression construct was screened at several conditions to select candidates that can be stably and highly expressed in *E. coli* hosts for later large-scale purification and crystallization screening.

1.1.1 Molecular Cloning of MATE Homologs

Eleven MATE homologs were cloned into pET19b vector (Novagen, Inc.) for expression in *E. coli BL21 (DE3)* (Invitrogen, Inc.) using an IPTG-inducible T7 system. The pET19b vector bears an N-terminus His10-tag followed by an enterokinase cleavage site upstream of the multiple cloning sites. Two NorM orthologs from the genus *Nesseria* were cloned

into pBADTOPOTA vector (Invitrogen, Inc.), which was modified with an additional ribosome binding site (RBS) placed before the multiple cloning site (MCS). Both of these two Nesseria homologs were also cloned into a pBADThioTA vector (Invitrogen, Inc.) that is a translational fusion vector producing target proteins with N-terminal thioredoxin tag for enhanced expression. These expression constructs of MATE homolog are listed in Table 1.1.1.

Table 1.1.1 List of Bacterial MATE Homologs Cloned into Various Expression Vectors for Large-scale Expression in Host Strains *E. coli* BL21 DE3 or TOP10

Construct ID	**MATE Ortholog and Source Organisms**
NorM_VC wildtype-p19b_+	NorM from *V. cholera*
NorM_VC-Nt4AAshort-p19b#1_+	NorM from *V. cholera* with 4 residues truncated at the N-termus
NorM_VA-p19b#2_+	NorM from *V. alginolyticus*
NorM_SM-pET19b#3_+	NorM from *Streptococcus mutans*
NorM_TM-pET19b#1_+	NorM from *Thermotoga maritime*
NORM_BC-pET19b#1_+	NORM from *Bacillus cereus*
VmrA_VV_p19b#1_+	VmrA from *V. vulnificus*
NorM_XC-6-p19b_+	NorM from *Xanthomonas campestris*
XH093FR_aMate2_HM-p19b#2_+	aMate2 from *Haloarcula marismortui*
XH094FR_aMate3_HM-p19b#3_+	aMate3 from *H. marismortui*
XH097FR_aMate3_MS-p19b#1_+	aMate3 from *Methanobrevibacter smithii*
xh321_NorM_NG-pBADTOPOTA+RBS#2_+	NorM from *N. gonorrhoeae*
xh322_NorM_NM-pBADTOPOTA+RBS#2_+	NorM from *N. meningitidis*
NorM_NM-pBADThioTA#1_+	NorM from *N. meningitidis*
NorM_NG-pBADThioTA#1_+	NorM from *Neisseria gonorrhoeae*

In particular, NorM-VC (NCBI accession number AE003852.1 REGION: 1652328..1653713) was cloned from *V. cholera* genomic DNA (ATCC 39315) into the expression vector pET19b between the *NdeI* and *XhoI* restriction sites of the multiple cloning site.

1.1.2 Screening of High Expresser of MATE Homologs for Structure Determination

The expression level of MATE homologs cloned into the pET19b vector was checked by culturing the *E. coli BL21 (DE3)* transformants of each expression constructs in 1 L batches at 37 ˚C to OD600=0.65-0.7 in LB media before induction. Three induction temperatures (25, 30, and 37 ˚C) and two IPTG levels (0.2 and 0.4 g/L) in combination were tested for each expression construct. For MATE homologs cloned into the pBAD vectors, the expression constructs were transformed into *E. coli TOP10* (Invitrogen, Inc.) and grown to OD_{600}=0.65-0.7 in 1 L LB media at 37 ˚C before induction. Two temperatures (25, and 30˚C) and five arabinose levels with ten-fold increment starting from 0.00002% to 2% were tested for best induction conditions.

After overnight induction, cell pellets from each induction condition were solubilized with 1% SDS in Solubilization buffer (20 mM Tris-HCL pH8.0, 100 mM NaCl, 30 mM Imidazole) and the solubilized portion were assessed by Commassie blue-stained SDS-PAGE and Western blot using anti His-tag antibodies (Sigma, Inc.) for the presence of recombinant MATE homologs. From these expression trials, the NorM ortholog from *V. cholera* (VC)

champions with the best combination of expression level and protein stability (Figure 1.1.2).

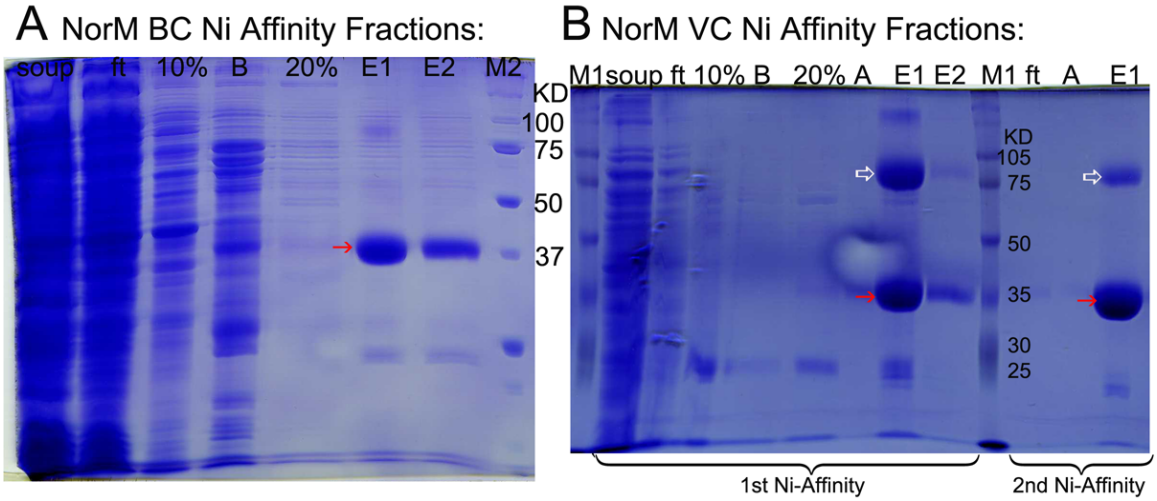


Figure 1.1.2 Coomassie-Stained PAGE Gels of Ni Affinity Purified NorM-BC and VC. (A) NorM-BC and (B) NorM-VC Ni-affinity chromatography fractions. Soup: supernatants of 1% β-DDM solubilized lysate from expressing cells; ft: flowthrough of Ni-affinity resin; A: 20mM Tris-HCl and 20mM NaCl; 10%: 30mM imidazole in A buffer; B: 1M NaCl in A buffer; 20%: 60mM imidazole in A buffer; E1/E2: 300mM imidazole in A buffer. Red closed arrows: eluted NorM ortholog mononers with His-tag. NorM-VC appeared to have the highest purity and yield from the same cell mass. All buffers contain 0.012% β-DDM, except for Soup. White open arrows: possible dimmers of NorM ortholog.

A subsequent detergent extraction screening for NorM-VC was performed to optimize the purification efficiency and protein stability for later crystallization trials. A variety of mild detergents designed for membrane protein extraction and purification was tested on their ability to extract and maintain the soluble state of NorM-VC during and after small-scale Ni-affinity purification. Detergents with maltopyranoside and

glucopyranoside meioties and hydrocarbon chain length between 9 and 12 appear to extract and stabilize NorM-VC quite well.

1.1.3 Optimized Expression and Purification Conditions of NorM-VC

Freshly transformed E. coli BL21 DE3 by *pET19b-norM-VC* plasmid were cultured in 100 L batches in a bioreactor (BioFlow 5000, New Brunswick Scientific) from an initial 1 L seed culture. Cells were grown to OD600=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 expressing NorM-VC 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 Soluliblization buffer with additional 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 HiPrepTM 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 16oC 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 A280 using an extinction coefficient ε_{280} = 56.4 x 10^{-3} $M^{-1}cm^{-1}$. NorM-VC mutants were expressed and purified in the same manner.

1.1.4 *In vivo* and *in vitro* Functionality of NorM-VC

NorM-VC retains transport activity in the E. coli expression hosts and purified NorM-C protein is also active and binds to a known MATE substrate[44]. NorM-VC transport activity was tested using a cell-based accumulation assay with a known substrate ethidium bromide (EtBr). The accumulation assay was performed as previously described elsewhere[45] with some modifications. Host strain *E. coli BL21 DE3 ΔacrAB* was used to reduce the background efflux due to endogenous MDR transporters. Briefly, these cells were transformed with empty vector *pET19b* or *pET19b-norM-VC* construct, grown to mid-log phase (OD_{600}=0.7) at 200 rpm 37°C, and then induced with 0.4 g/L IPTG at 80 rpm 30°C overnight. The cell

cultures were normalized to the same OD and volume before 0.1 μM EtBr treatment for 1 hr at 200 rpm 37°C. The treated cultures were then pelleted and washed with 100 mM Tris-HCl pH 7.0 for three times. Each cell pellet was then resuspended using 6ml of Tris-HCl buffer and divided into three aliquots. Each aliquot was pelleted again and lysed directly with 250 μl of 5% Trichloroacetic acid (TCA) and vigorously vortexed. Two hundred μl of the supernatant from each TCA extracted aliquot was used to quantify EtBr top fluorescence 96-well black plates using a Beckman DTX 880 fluorimeter with excitation at 485 nm and emission at 595 nm. Cellular content of EtBr represented by EtBr top fluorescence counts in cell extractions was obtained by averaging between replicates and a representative run is shown in Figure 1.1.4A.

The substrate binding activity of NorM-VC was assessed by fluorescence-based polarization assay with a well known MATE substrate, ethyl eosin[44], to characterize the binding activity of purified detergent solubilized NorM-VC. Measurement of fluorescence polarization was conducted similar to that described previously, using a microplate reader equipped Beckman DTX 880 fluorimeter and Nunc™ 384-well plates. Protein samples were two-fold serial-diluted into 20 μl of desalting buffer with 0.23% β-NG (Desalting-NG) to obtain conditions similar to those in

crystallization. Equal amounts of substrate were then added into each well. The reaction mixture was incubated at 25˚C for 10 min before fluorescence was measured. The polarization of ethyl eosin in the presence of NorM-VC as an indication of specific binding was carried out in at least duplicate and experiments with each substrate were repeated at least three times using protein samples from different preparations (Fig. S3A, B and Table S1).

Experimental polarization values were calculated according to the equation P = 1000 x (IV - G x IH)/(IV + G x IH), where P is the polarization value in each well, IV is the fluorescence intensity polarized parallel to the excitation light, and IH 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 = IHV/IHH. To determine the dissociation constant, KD, 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 = Po + (Pmax - Po)((NorM-VC)/((NorM-VC) – KD)), where Po is the minimum fluorescence polarization value of free substrate, Pmax 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 Po, Pmax, and KD treated as variable during the calculations to achieve the least fitting error. Experimental data were calculated and plotted in Excel. The non-linear fitting of data was performed in Origin 7 (Origin Lab, Inc.).

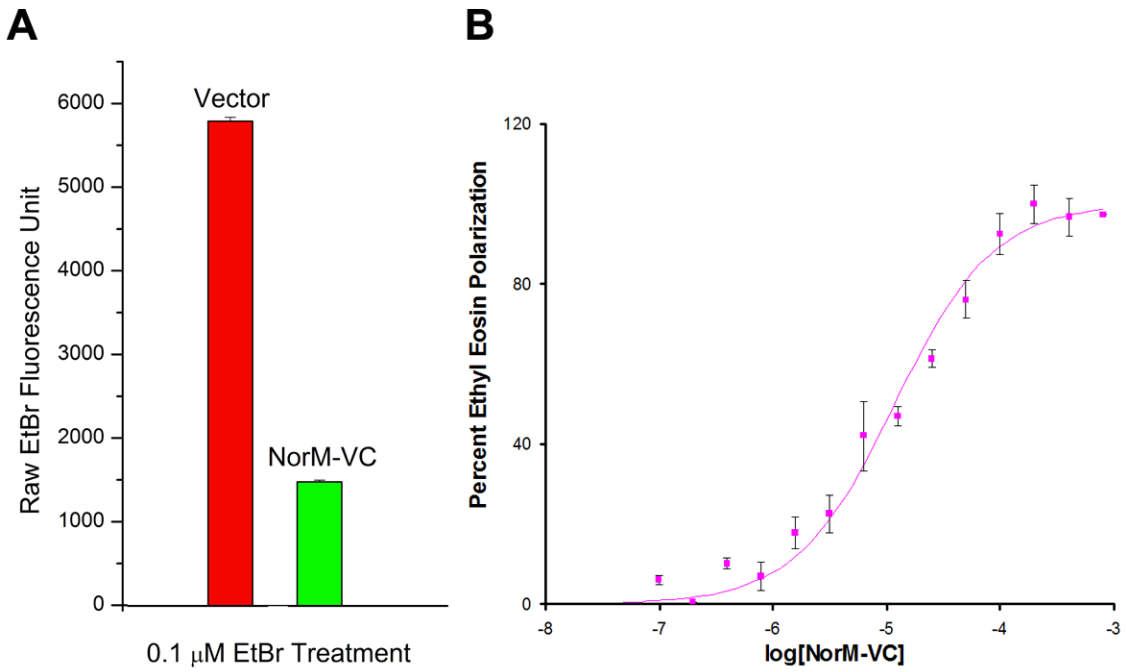


Figure 1.1.4 In vivo and in vitro characterization of recombinant NorM-VC protein. (A) Effect of recombinant NorM-VC on EtBr uptake by *E. coli* BL21 DE3 *ΔacrAB* as measured by EtBr accumulation. The cells were transformed with pET19b vector alone as a control (red column) or with pET19b-*norM-VC* (green column). Fluorescence intensity values reflect EtBr accumulation between cell samples of equal mass. (B) NorM-VC binds to ethyl eosin (▪). Data is plotted as percent polarization of the measured fluorescent substrate versus the logarithm of competitive substrate concentration (molar). The fitted curve models single-site binding. Dissociation, excitation, and emission wavelengths are presented in Table 1.1.4. 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.

Table 1.1.4
Affinity of NorM-VC to its substrates determined by fluorescent polarization

Compound	K_d, μM	λ_{Ex}, nm	λ_{Em}, nm
Ethyl Eosin	11.73 ± 1.12	485	535

1.2 Crystallization

NorM-VC protein purified in permissive detergents were used in crystallization trials with a variety of pH buffering agents covering buffering range from pH 3.0 to pH10.6, Polyethlyene glycocol (PEG) precipiants, and salts. The optimized crystallization conditions of NorM-VC requires protein sample solubilized and purified in β-DDM, with a final concentration of 0.23% n-nonyl- β -glucopyranoside (β -NG) supplemented at the end of purification. The range of the optimized conditions are the following: 50mM Tris-HCl pH 7.2-8.6, 87 mM $(NH4)_2SO_4$, 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. The addition of 1X CMC of a novel facial amphiphile, 3α-hydroxy-7 α,12 α -bis((β-D-maltopyranosyl)ethyloxy) cholane (FA-231)[46] into the final protein samples

appears to improve the reproducibility of the crystals and lower the background precipitation. It is noteworthy that these crystals were distinct from those previously reported[47]. NorM-VC mutants (Table 1.4.2) were crystallized in the same fashion (Figure 1.2.1).

The highest diffracting crystal, Crystal1 (3.65 Å; a=142.8 Å, b=240.8 Å, c=45.7 Å, $\alpha=\beta=\gamma=90^{\circ}$) and the Rb+-bound crystal, Crystal2 (4.2 Å; a=159.6 Å, b=241.7 Å, c=46.2 Å, $\alpha=\beta=\gamma=90^{\circ}$), used for the refinement of the two structures of NorM-VC (Table 1.5.1), were grown in the presence of 0.02% FA-231. In addition, the two cation-site mutants, D371A and D371N (Table 1.4.2), were also crystallized in the presence of the FA-231. We attribute the higher diffraction quality of Crystal1 mostly to serendipity as the crystal growth appears to be heavily dependent on the preparation.

1.3 Derivatization and Cryoprotection

Derivatization of native and single cysteine mutant crystals for phasing and/or 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 in D_2O, supplemented with 0.009% β-NG, and 0.016% β -DDM as cryo-protectant for periods ranging between 0.5-1 hour. Rb and Cesium (Cs) derivatization of native and cation-binding site

mutant crystals was achieved by soaking with 200 mM RbCl, or 200 mM CsCl in mother liquor. RbCl, CsCl, and NaCl (control) soaks all elicit rapid crystal decade, likely due to disruption of lattice contacts caused by conformational changes, which precludes derivatization beyond 1 minute.

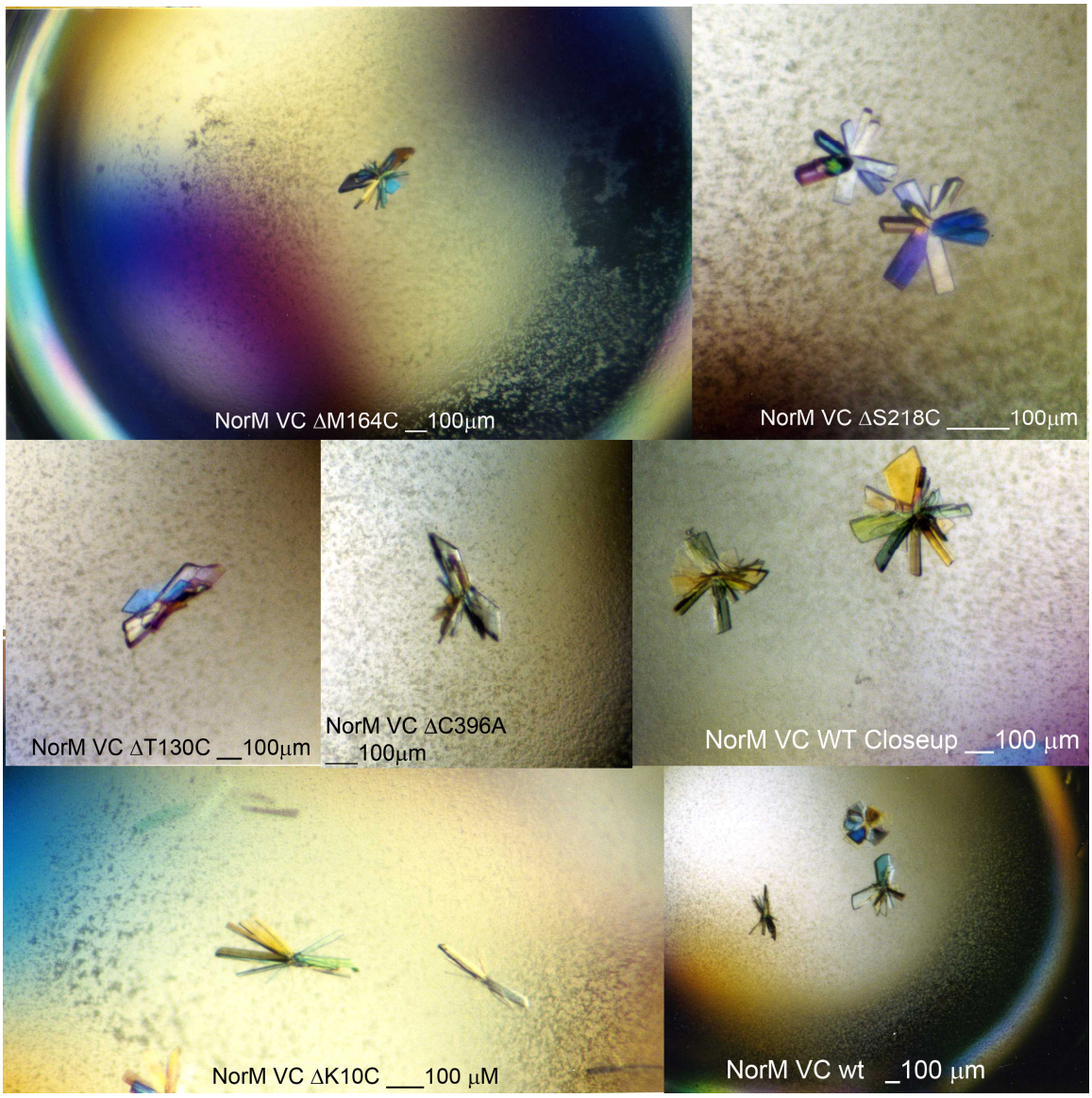


Figure 1.2.1 Crystals of NorM-VC. Crystallization conditions are the same as those for wild-type NorM_VC crystals.

1.4 Diffraction Data Acquisition and Reduction

1.4.1 Synchrotron Data Acquisition

Diffraction data used for structure refinement, topology verification, and cation-binding site characterization was collected at the Hg L-III absorption edge at various synchrotron sources (Table 1.4.1). The data sets from the three Hg compound-derivatized crystals used for MAD

Table 1.4.1 Data Processing of NorM-VC Mutants and/or Additional Data Sets Used for Topology and Caition-binding Site Verification.

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

Notes:
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.
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.

phasing was also collected at the Hg L-III absorption edge, as well as an additional set at the inflection point of Hg L-III for each crystal (Table 1.5.1). All energy edges were determined by XAFS on site.

1.4.2 Diffraction Data Reduction

All diffraction data was processed with the programs HKL2000 and MOSFLM [48,49].

1.5 Phasing the Experimental Map

1.5.1 MAD Phasing by Phase Combination

The initial experimental electron density map and structure of NorM-VC was determined by multiple anomalous dispersion (MAD) phasing using the phase combination technique. The phase information was obtained from data sets obtained from three crystals: a ΔT130C crystal derivatized by $Hg(OAc)_2$, and two wildtype crystals derivatized by $Hg(CN)_2$ and BDM, respectively (Table 1.5.1).

Anomalous Patterson Functions and direct methods (SHELX) were used to identify the heavy atom sites necessary for obtaining protein phases and experimental electron density maps. The top five sites with the highest |F| over the background from each crystal were used (six sites in

each ASU exist). Protein phasing, solvent flattening, non-crystallographic symmetry (NCS) averaging, and phase extension were performed using the programs PHASES, SOLVE/RESOLVE, and CNS v1.2 [50,51,52,53,54].

1.5.2 Experimental Density Map Calculation and Map Averaging

The experimental electron density maps were calculated using the program suite CNS v1.2 and PHASES. Two NorM-VC molecules present in the asymmetric unit (NorM1 and NorM2) (Fig. 1.5.2). The starting electron density correlation coefficient from the application of NCS averaging between NorM1 and NorM2 was 60% (Table 1.5.1).

Table 1.5.1 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

Notes:

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

1.6 Composite Omit Density Maps-guided Model Building and Refinement

Fo-Fc and 2Fo-Fc difference electron density maps are both useful tools to improve model building and refinement. Both methods were used to validate and facilitate the model building of NorM-VC structures. The composite Fo-Fc maps of NorM-VC have revealed certain structural details even at the modest resolution range 3.5-4.5 Å. An example of such is the observation of densities of bound solvent at multiple internal cavity residues in the composite (Fo-Fc) omit density maps of the NorM-VC apo-structure (Figure 1.6.1 and Figure 2.1.7 H). Although both Fo-Fc and 2Fo-Fc methods were used throughout the model building processes, we emphasize the composite (Fo-Fc) difference density maps in great details in the following chapter.

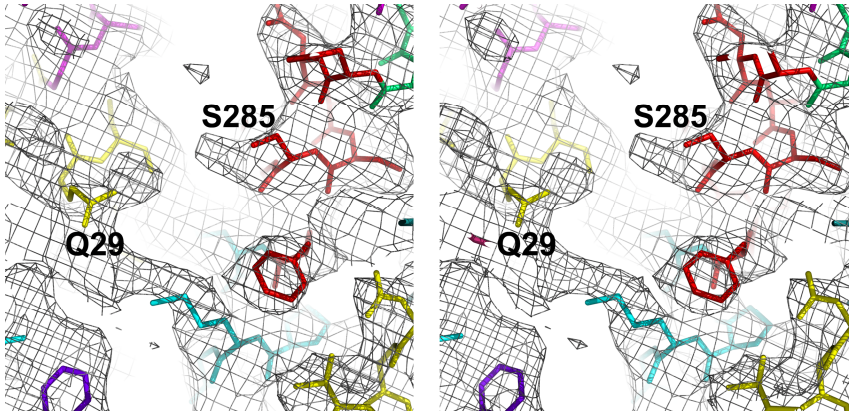


Figure 1.6.1 Composite (Fo-Fc) omit density maps of the NorM-VC apo-structure revealing bound solvent density near side chains of residue Q29 and S285. NorM-VC apo model is rendered in sticks. The omit density maps were calculated with every 8 consecutive residues (highlighted as same the color) omitted together.

1.6.1 Model Building

Models were built using the programs CHAIN, Coot, and PyMOL [55,56,57]. Stereo figures were generated using PyMOL and Adobe Photoshop 7.0.

1.6.2 Refinement

Crystallographic refinement using data from Crystals 1 and 2 was accomplished using the maximum likelihood target with amplitudes (mlf) using CNS v1.2[54]. 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 (apo) and with R_{cryst} of 30.9% and R_{free} of 34.2% for Crystal 2 (Rb^+-bound) (Table S3). 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.

Table 1.6.2 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

Notes:
[b] $R_{cryst} = \Sigma|F_o-F_c|/\Sigma|F_o|$, 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.
R_{shrink}, shrink factor value in bulk solvent model grid search.
R_{probe}, solvent radius value in bulk solvent model grid search.
K, a scale factor that defines the mean electron density in the solvent region
B_{sol}, thermal factor of the solvent.

1.6.3 Omit Density Maps-Guided Model Building and Refinement

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 (Figure 2.1.2). The Fo-Fc maps were calculated using CNS v1.2[54] 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. The refined structures of NorM-VC are the products of stepwise rebuilding and re-refining the NorM-VC model according to the omit density maps. An example

showing how such method helped the model building and refinement at the moderate resolution range (3.5-4.5 Å) is shown in Figure 1.6.3.

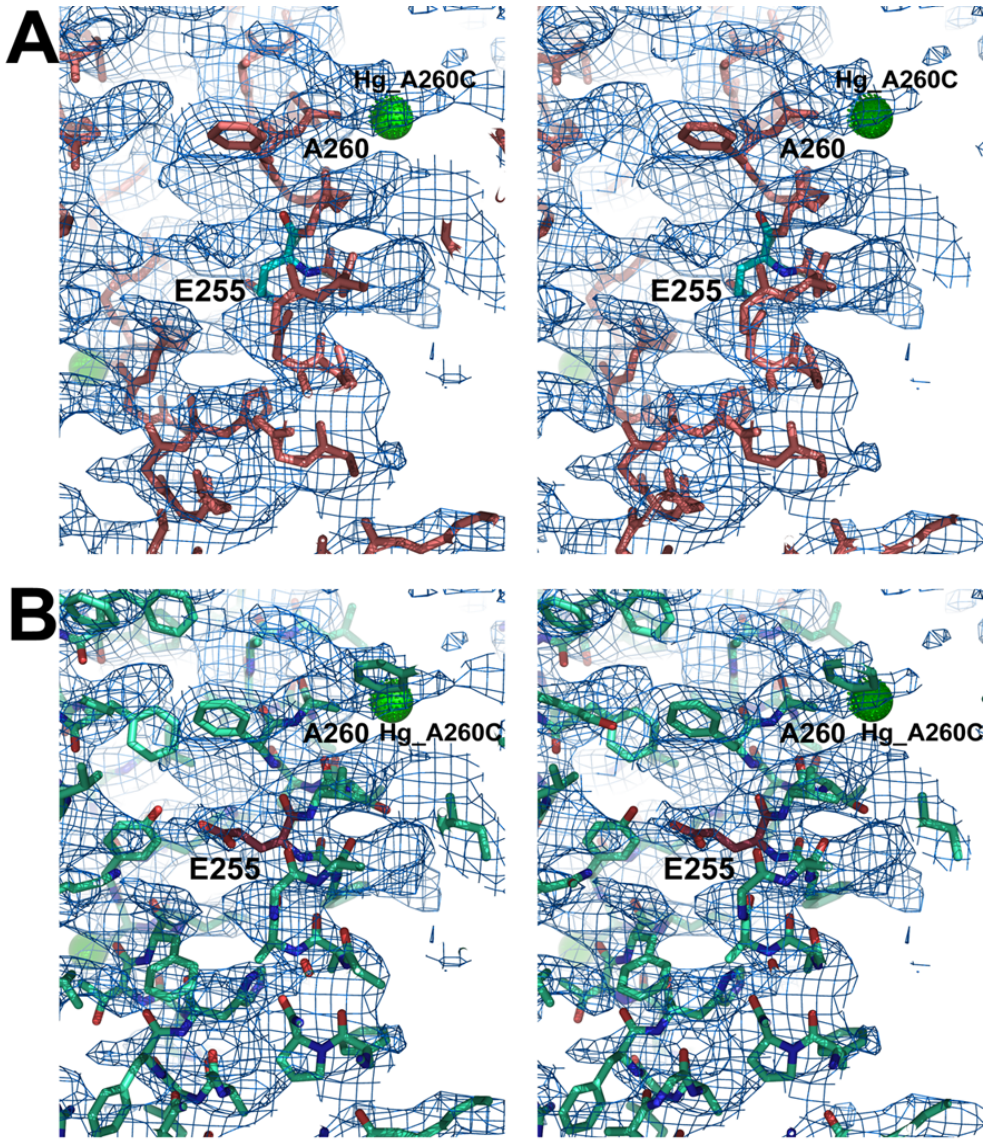


Figure 1.6.3 Example of Composite Fo-Fc Omit Density Map-Aided Model Building. (A) Before manual rebuild at residual E255 (cyan stick model) versus the rest of the model (salmon pink stick model). The NorM-VC apo model superimposed with multiple composite SA omit Fo-Fc maps aound the E255 region showed an incorrect build at residue E255 against the SA omit maps with r=0.3417 and r-free=0.3787. (B) After maual rebuild of residual E255(dark pink), the NorM-VC model (gren) fits better with the corresponding Fo-Fc omit map and has R=0.3379 and R-free=0.3770.

1.7 Mutagenesis and Topology Verification for the First Representative of the MATE Fold

1.7.1 Cysteine Substitution Mutants and Mercury-labeling of the Mutated Residues

To verify the topology and structure of NorM-VC, one hundred and thirty-six single cysteine substitution mutants were generated from the pET19b-*norM-VC-wt* construct using the QuickChangeTM 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, Inc. The complete collection of single cysteine substitution mutants whose sequences have been confirmed is listed in Table 1.7.1. All these single cysteine mutants were crystallized, derivatized, and processed as described above.

Table 1.7.1 List of NorM-VC Mutants Generated for Topology Verification and Characterization of the Cation-binding Site

Clone/Mutant ID	Cate.	**Clone/Mutant ID**	Cate.
XH194_NorM_VC_D121C-p19b#2	a	XH259_NorM_VC_L346C-p19b#1	a
XH195_NorM_VC_E123C_p19b#1	a	XH261_NorM_VC_T348C-p19b#1	c
XH196_NorM_VC_D124C-p19b#1	a	XH262_NorM_VC_Q351C-p19b#1	a
XH197_NorM_VC_M164C-p19b#1	f	XH263_NorM_VC_V353C-p19b#1	a
XH198_NorM_VC_V182C_p19b#1	f	XH264_NorM_VC_A355C-p19b#1	a
XH199_NorM_VC_G197C_p19b#1	a	XH265_NorM_VC_L356C-p19b#1	a
XH200_NorM_VC_A199C_p19b#1	d	XH268_NorM_VC_I343C-p19b#1&3	a
XH201_NorM_VC_I215C_p19b#1	d	XH269_NorM_VC_A344C-p19b#1	a
XH202_NorM_VC_V216C_p19b#1	f	XH271_NorM_VC_Q448C-p19b#1	a
XH203_NorM_VC_H223C_p19b#1	b	XH274_NorM_VC_A459C-p19b#2	a
XH204_NorM_VC_V224C_p19b#1	e	NorM_VC_C369A-p19b#1	g
XH205_NorM_VC_T229C_p19b#1	a	XH260_Y347C-p19b#1	a
XH185_NorM_VC_V43C_p19b#1	e	XH272_Q450C-p19b#2	a
XH186_NorM_VC_S44C_p19b#1	a	XH277_L322C-p19b#1	a
XH187_NorM_VC_N80C_p19b#1	c	XH278_A357C-p19b#1	a
XH188_NorM_VC_A82C_p19b#1	c	XH279_K449C-p19b#3	a
XH189_NorM_VC_G83C_p19b#1	a	XH281_L456C-p19b#1	a
XH190_NorM_VC_H86C_p19b#2	b	XH282_K461C-p19b#1	a
XH191_NorM_VC_E91C_p19b#1	f	XH126_NorM_VC_T20C-p19b#1	d
XH192_NorM_VC_H93C_p19b#1	a	XH127_NorM_VC_S26C-p19b#1	f
XH193_NorM_VC_Q94C_p19b#1	d	XH128_NorM_VC_S56C-p19b#1	d
XH207_NorM_VC_P235C-p19b#2	a	XH129_NorM_VC_S61C-p19b#1	d
XH208_NorM_VC_E255C-p19b#1	a	XH130_NorM_VC_A99C-p19b#1	d
XH209_NorM_VC_H277C-p19b#2	a	XH131_NorM_VC_S103C-p19b#1	f
XH210_NorM_VC_H307C-p19b#2	a	XH132_NorM_VC_T130C-p19b#1	e
XH211_NorM_VC_D310C_p19b#2	a	XH133_NorM_VC_A149C-p19b#1	f
XH212_NorM_VC_G313C_p19b#1	a	XH134_NorM_VC_T160C-p19b#1	c
XH213_NorM_VC_A315C_p19b#1	a	XH135_NorM_VC_A163C-p19b#1	e
XH214_NorM_VC_E349C_p19b#1	a	XH136_NorM_VC_S218C-p19b#1	d
XH215_NorM_VC_N350C_p19b#1	a	XH137_NorM_VC_A260C-p19b#1	f
XH216_NorM_VC_H392C_p19b#1	b	XH138_NorM_VC_A263C-p19b#1	e
XH217_NorM_VC_H457C_p19b#1	b	XH139_NorM_VC_S299C-p19b#1	c
XH218_NorM_VC_M289A-p19b#1	e	XH140_NorM_VC_G303C-p19b#1	a
XH219_NorM_VC_M289C-p19b#1	d	XH141_NorM_VC_T336C-p19b#1	a
XH220_NorM_VC_F290C-p19b#1	d	XH142_NorM_VC_T354C-p19b#1	a
XH221_NorM_VC_M292A-p19b#1	e	XH143_NorM_VC_S397C-p19b#1	f
XH222_NorM_VC_M292C-p19b#1	e	XH144_NorM_VC_Q418C-p19b#1	a
XH223_NorM_VC_M39A-p19b#1	a	XH145_NorM_VC_L438C-p19b#1	d
XH224_NorM_VC_M39C-p19b#1	a	CSA_p19b#1	a
XH225_NorM_VC_H86A-p19b#1	c	SCA_p19b#1	a
XH226_NorM_VC_H93A-p19b#1	a	XH151_NorM_VC_D36C_p19b#1	d
XH227_NorM_VC_H135A-p19b#1	a	XH153_NorM_VC_D47C_p19b#1	a
XH228_NorM_VC_H223A-p19b#1	b	XH155_NorM_VC_A54C_p19b#1	b
XH229_NorM_VC_H231A-p19b#1	b	XH156_NorM_VC_V76C_p19b#1	f

Table 1.7.1 List of NorM-VC Mutants Generated for Topology Verification and Characterization of the Cation-binding Site, Continued.

XH230_NorM_VC_H277A-p19b#1	a	XH157_NorM_VC_A77C_p19b#1	a
XH232_NorM_VC_H392A-p19b#1	a	XH158_NorM_VC_L101C_p19b#1	f
XH233_NorM_VC_H457A-p19b#1	a	XH159_NorM_VC_L150C_p19b#1	a
XH190+XH228_NorM_VC_H86C+H223A_p19b#1	c	XH160_NorM_VC_D155C_p19b#1	a
XH190+XH233_NorM_VC_H86C+H457A_p19b#2	b	XH161_NorM_VC_G184C_p19b#1	a
XH225+XH203_NorM_VC_H86A+H223C_p19b#1	b	XH162_NorM_VC_E190C_p19b#1	d
XH225+XH203_NorM_VC_H86A+H223C_p19b#2	b	XH163_NorM_VC_G195C_p19b#1	a
XH225+XH217_NorM_VC_H86A+H457C_p19b#1	b	XH164_NorM_VC_E228C_p19b#1	a
XH225+XH228_NorM_VC_H86A+H223A_p19b#1	b	XH165_E237C_p19b#2	a
XH225+XH233_NorM_VC_H86A+H457A_p19b#1	b	XH166_F242C_p19b#1	a
XH244_NorM_VC_M39A&V43C-p19b#1	c	XH169_E308C_p19b#1	a
XH245_NorM_VC_H223A&V224C-p19b#1	b	XH171_E341C_p19b#1	a
XH246_NorM_VC_F139C-p19b#2	a	XH172_V352C_p19b#1	a
XH247_NorM_VC_V256C-p19b#1	a	XH173_M370C_p19b#1	d
XH248_NorM_VC_F64C-p19b#1	a	XH174_D371C_p19b#1	d
XH249_NorM_VC_Q29C-p19b#1	a	XH177_E417C_p19b#1	a
XH250_NorM_VC_A296C-p19b#1	f	XH178_G424C_p19b#1	a
XH251_NorM_VC_I300C-p19b#1	c	NorM_VC_D453C-p19b#1	a
XH252_NorM_VC_L68C-p19b#1	c	NorM_VC_F339C-p19b#1	d
XH255_NorM_VC_F153C-p19b#1	b	XH099_NorM_VC_ACA-p19b#1	b
XH256_NorM_VC_M323C-p19b#1	f	XH100_NorM_VC_CAA-p19b#1	b
XH257a_NorM_VC_Q148C-p19b#1	a	XH327_NorM_VC_Y367C-p19b#1	f
XH257b_NorM_VC_N139C-p19b#1	a	PRS_NorM_VC_K10C	f
XH258_NorM_VC_L335C-p19b#2	a	AK_NorM_VC_F429C	f

Notes:

Categories: a, Expression level is insufficient (less than 50% of that of the wild-type) for pursuing structural information; b, Purified protein is unstable; c, Null crystallization at the conditions close to those of the wild-type; d, Crystal diffraction quality is unsatisfactory to reliably conclude whether there is a Hg site or the accurate position of such a site; e, No Hg site was observed to confirm neither the position nor the identity of the residue being mutated, however, the data sets per se was of satisfactory quality to conclude unsuccessful labeling due to accessibility or occupancy as the cause; f, Two NCS-related Hg sites were observed within each ASU, with each site corresponding to the same residue in each NorM-VC monomer (these are reported in Table 1.4.2); g: Naturally occurring cysteine replacement mutants that reveal the loss of Hg site at the corresponding position but were not reported in reference[58].

1.7.2 Individual Difference Fourier Maps of the Mercury-derivatized Cysteine Mutants

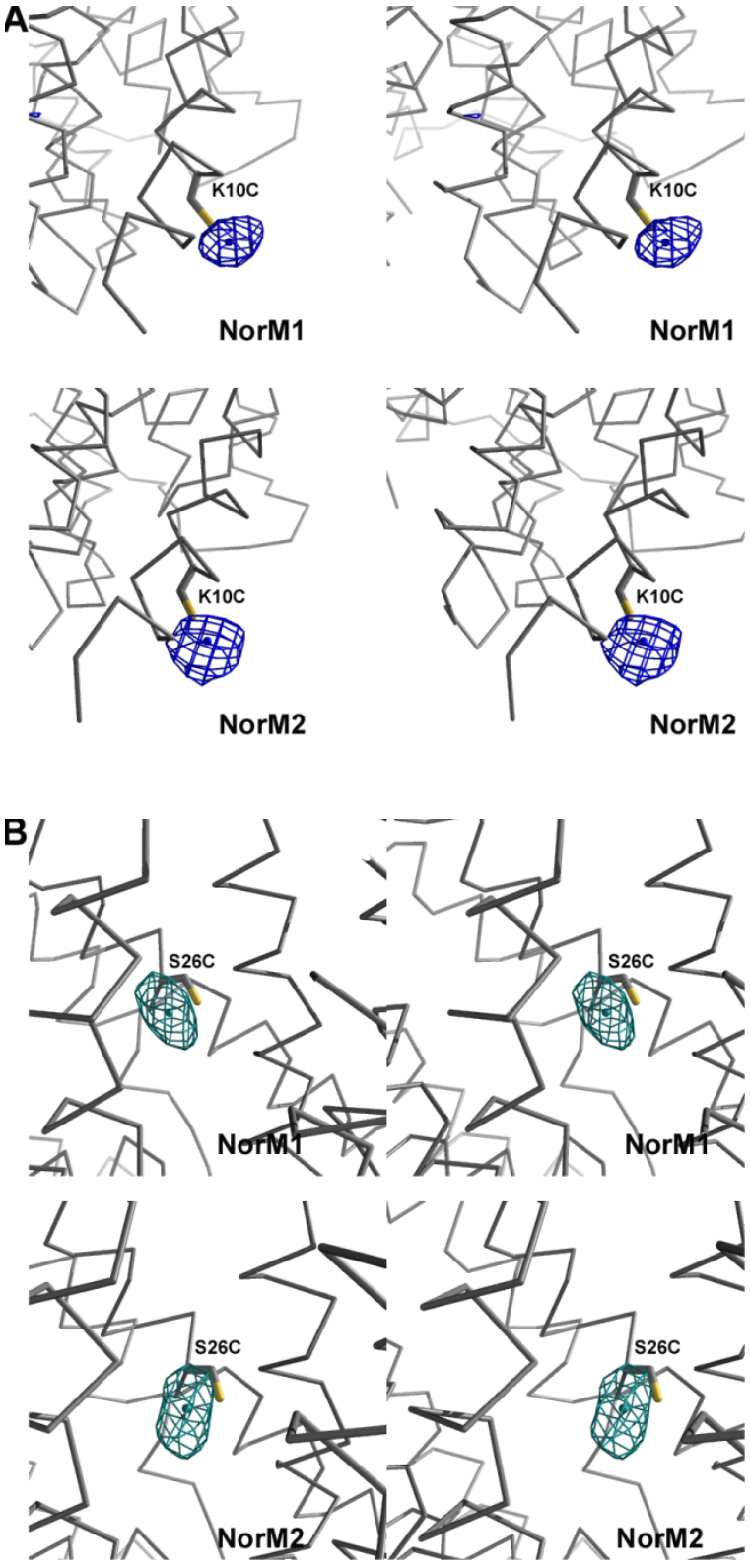
Each of these cysteine substitution mutants were tested for expression and those that can yield substantial amount for crystallization were then grown in 100 L batches and purified as described for the wild-type NorM-VC. About half of all the cysteine mutants generated do not express nearly as the wild-type using similar bacterial growth/induction conditions (category a in Table 1.7.1). Among those that express at high enough level, about 20% become unstable in the same aqueous solutions used for wild-type NorM-VC purification at various steps (category b in Table 1.7.1). Of those that are stable, about 20% did not crystallized at condition range (Tris-HCl pH 7.0-8.8, PEG250DME 16-24% in D_2O, $(NH_4)_2SO_4$ 63 mM) close to that of the wild-type (category c in Table 1.7.1). Nearly one third of the mutants that can crystallize as the wild-type does in similar conditions do not diffract sufficiently high enough resolution such that the presence (or at least one) of the Hg sites cannot be resolved consistently (category d in Table 1.7.1). The remaining 29 mutants yield at least one reliable data set that can be used to calculate isomorphous difference Fourier synthesis along with another matching data set. The peaks of height at 3.5 σ or higher from these difference Fourier maps confirms the presence or absence of NCS-related Hg sites within each ASU due to the

presence of a corresponding cysteine residue in each monomer. Eight of these 29 distinct cysteine substitution mutants did not yield additional Hg sites as a result of the replacement of the corresponding residue by a cysteine, probably due to accessibility and/or occupancy of the site in the crystallographic context (category e in Table 1.7.1). Sixteen mutants revealed unambiguous Hg sites, where density peaks of 3.5 σ or above over the background were observed in isomorphous difference Fourier maps calculated among a matching pair of data sets (category f in Table 1.7.1; Table 1.4.1). Moreover, a naturally-occurring cysteine residue substitution mutant ΔC396A demonstrated the loss of the corresponding Hg sites when isomorphous difference Fourier map was calculated between a mercury (II) acetate-derivatized such crystal and a similarly derivatized wild-type crystal.

The isomorphous difference Fourier synthesis for each mutant data set was generated with the CNS v1.2 software suite[54] using the model from Crystal 2 with the Rb^+ removed. For each of these calculations, model phases were generated after rigid-body refinement and gradient minimization, using this Rb+-removed model and the structure factors of the corresponding mutant data (Table 1.4.1). 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 (Table 1.4.1 and Figure 1.7.2 Continued).

Figure 1.7.2 Close up views of mercury sites resolved from Hg-soaked crystals of cysteine substitution mutants used for topology verification. 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 in the supplemental text, and in Table 1.4.1. The mutation positions and sigma cutoffs of the difference maps for both monomers are the following. (A) K10C, 3.5 σ for NorM1 and 4.0 □ for NorM2. (B) S26C, 4.0 σ for NorM1 and 3.5 σ for NorM2. (C) V76C, 7.0 σ. (D) E91C, 4.0 σ. (E) L101C, 4.0 σ for NorM1 and 4.5 σ for NorM2. (F) S103C, 4.5 σ for NorM1 and 4.0 σ for NorM2. (G) A149C, 5.0 σ. (H) M164C, 5.0 σ for NorM1 and 4.0 σ for NorM2. (I) V182C, 4.5 σ. (J), V216C, 6.0 σ for NorM1 and 5.0 σ for NorM2. (K) A260C, 4.0 σ for NorM1 and 3.5 σ for NorM2. (L) A296C, 4.5 σ. (M) M323C, 5.5 σ. (N) Y367C, 5.0 σ for NorM1 and 4.0 σ for NorM2. (O) S397C, 5.0 σ for NorM1 and 4.5 σ for NorM2. (P) F429C, 4.0 σ. Mesh coloring is the same as in Figure 1.7.3.


A
K10C
NorM1
K10C
NorM1
K10C
NorM2
K10C
NorM2
B
S26C
NorM1
S26C
NorM1
S26C
NorM2
S26C
NorM2

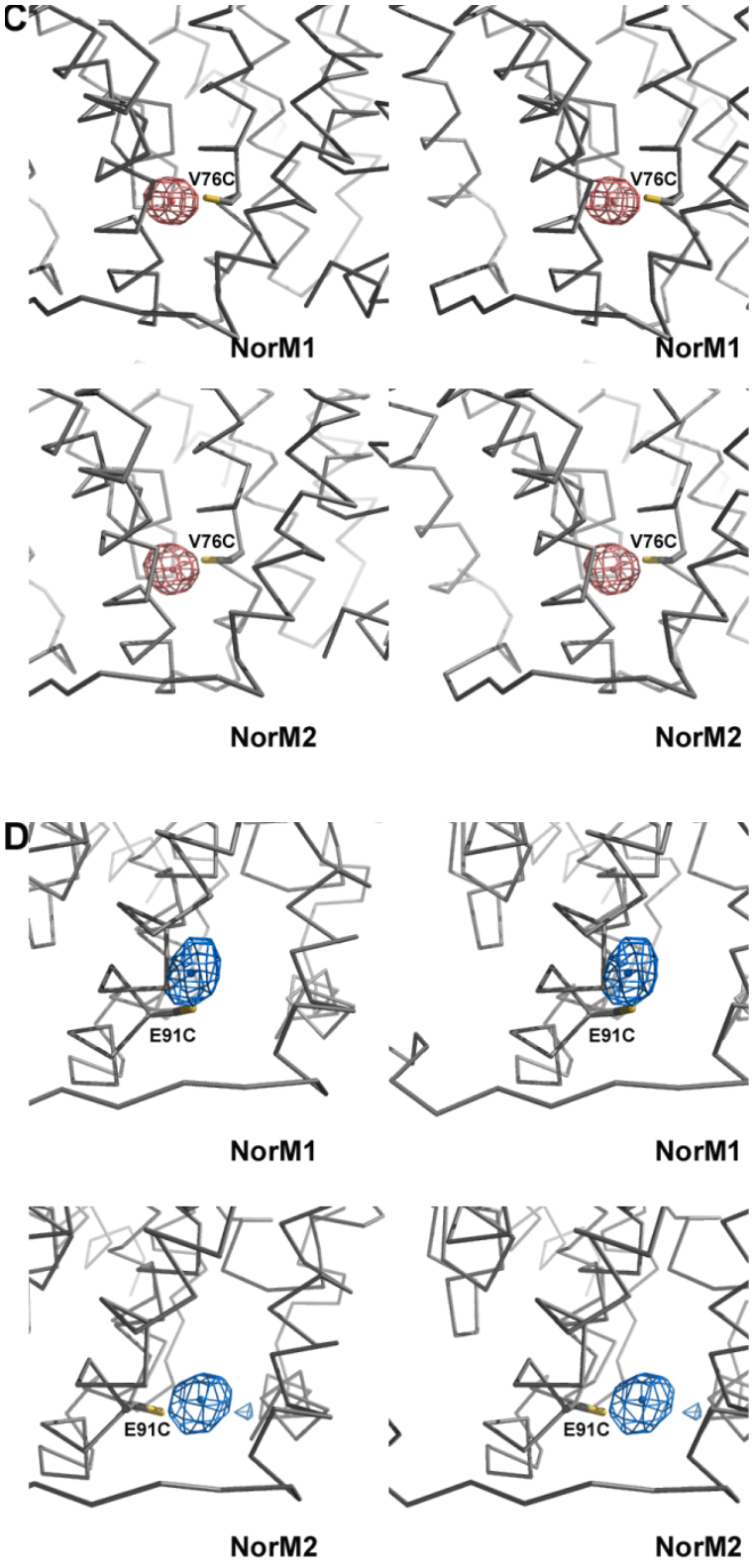


Figure 1.7.2 Continued

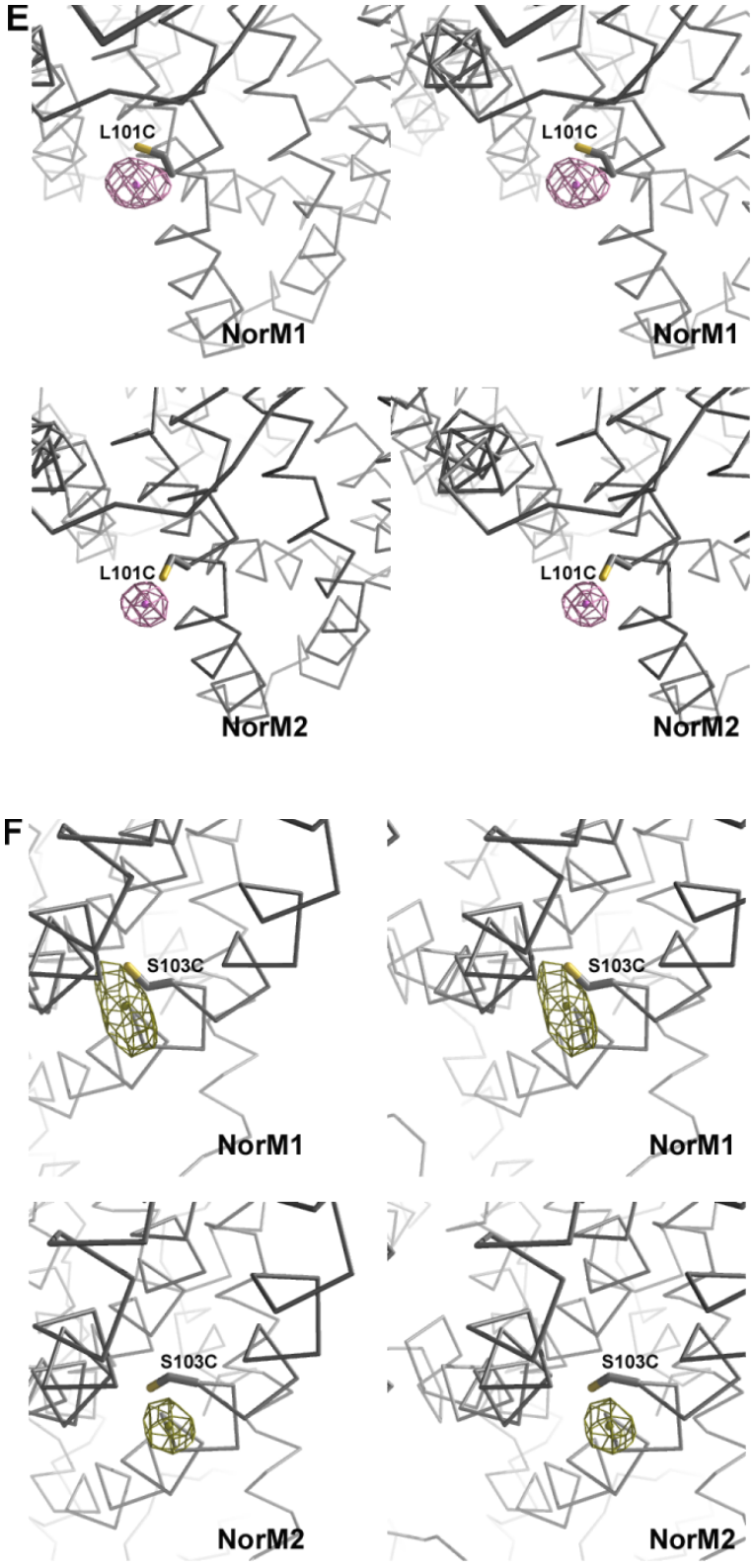


Figure 1.7.2 Continued

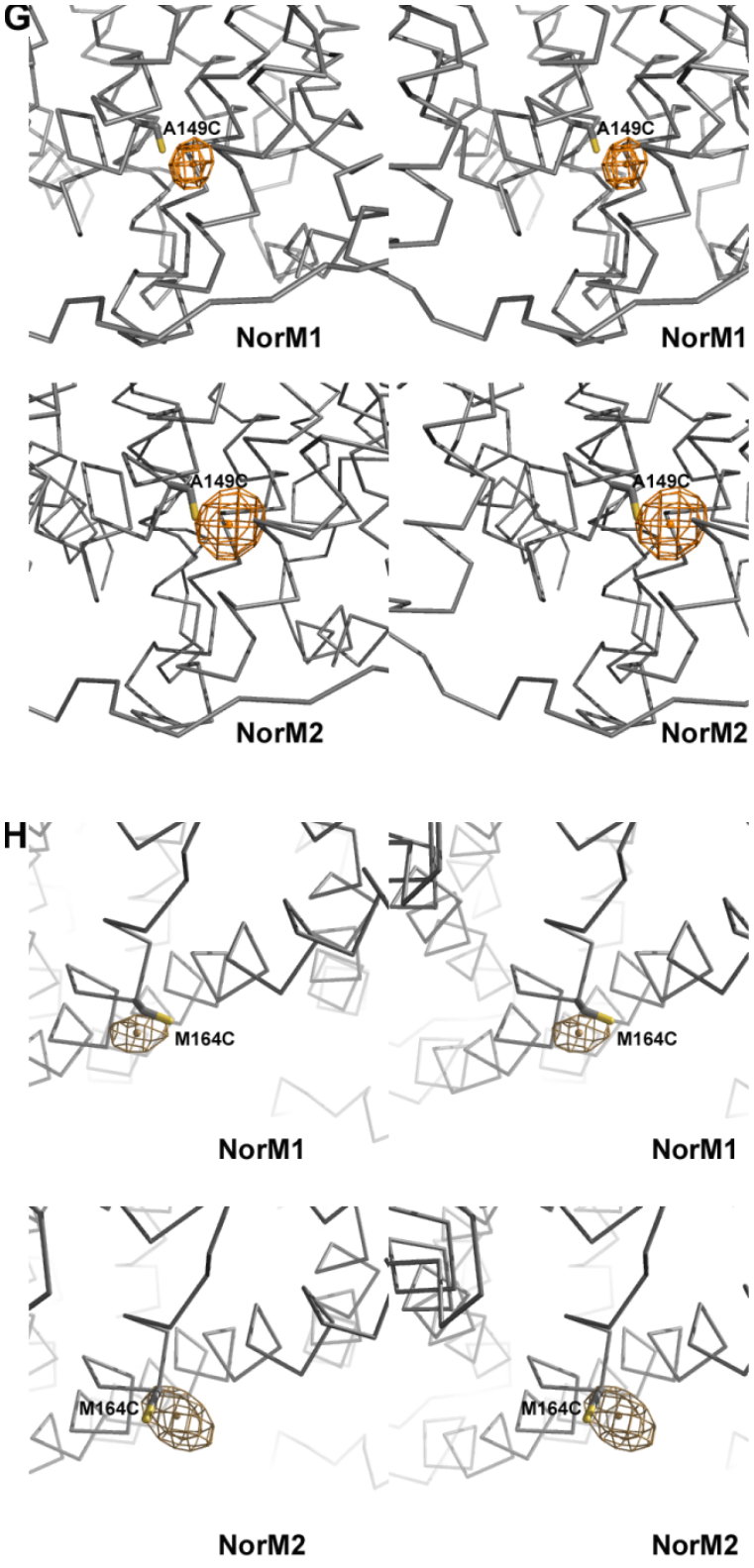


Figure 1.7.2 Continued

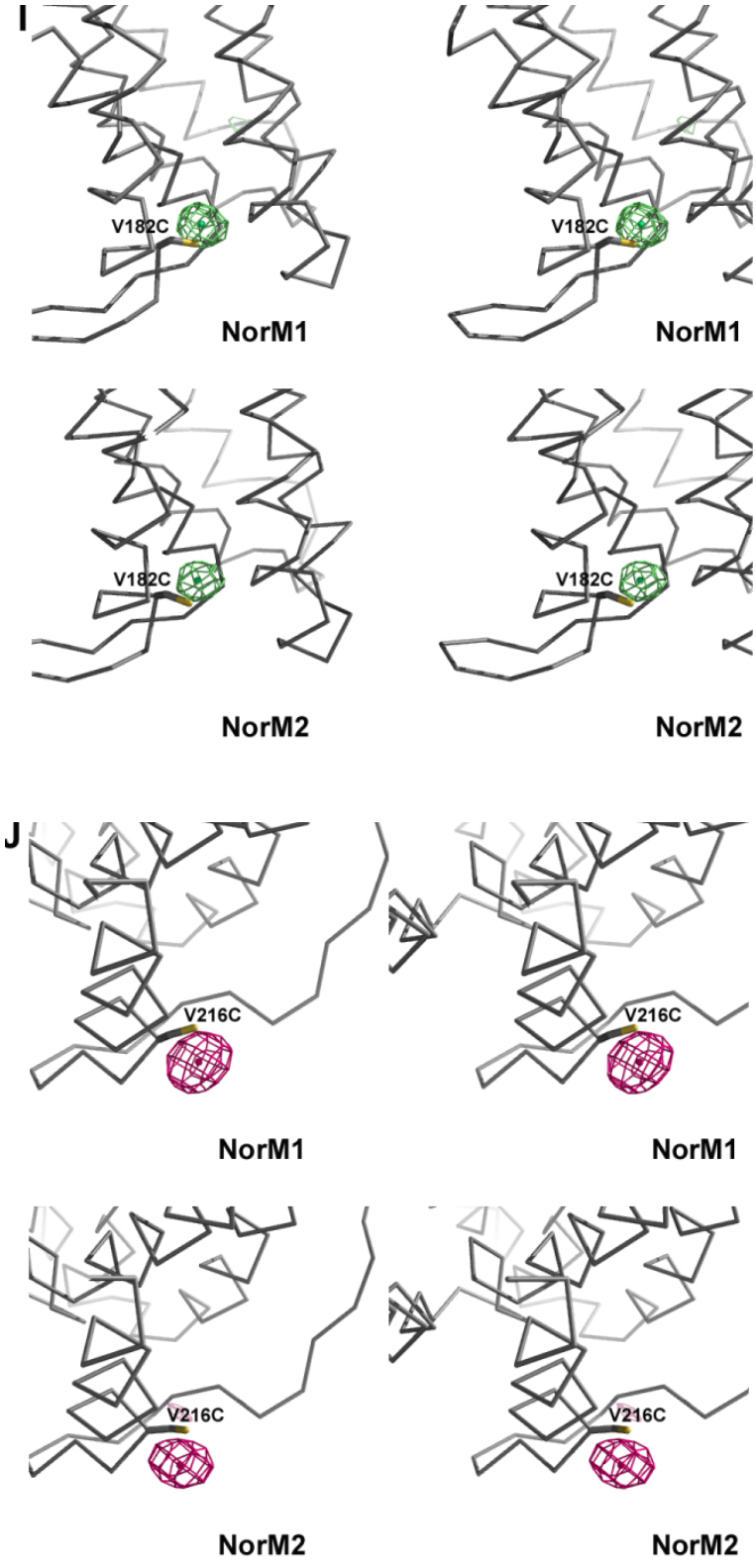


Figure 1.7.2 Continued

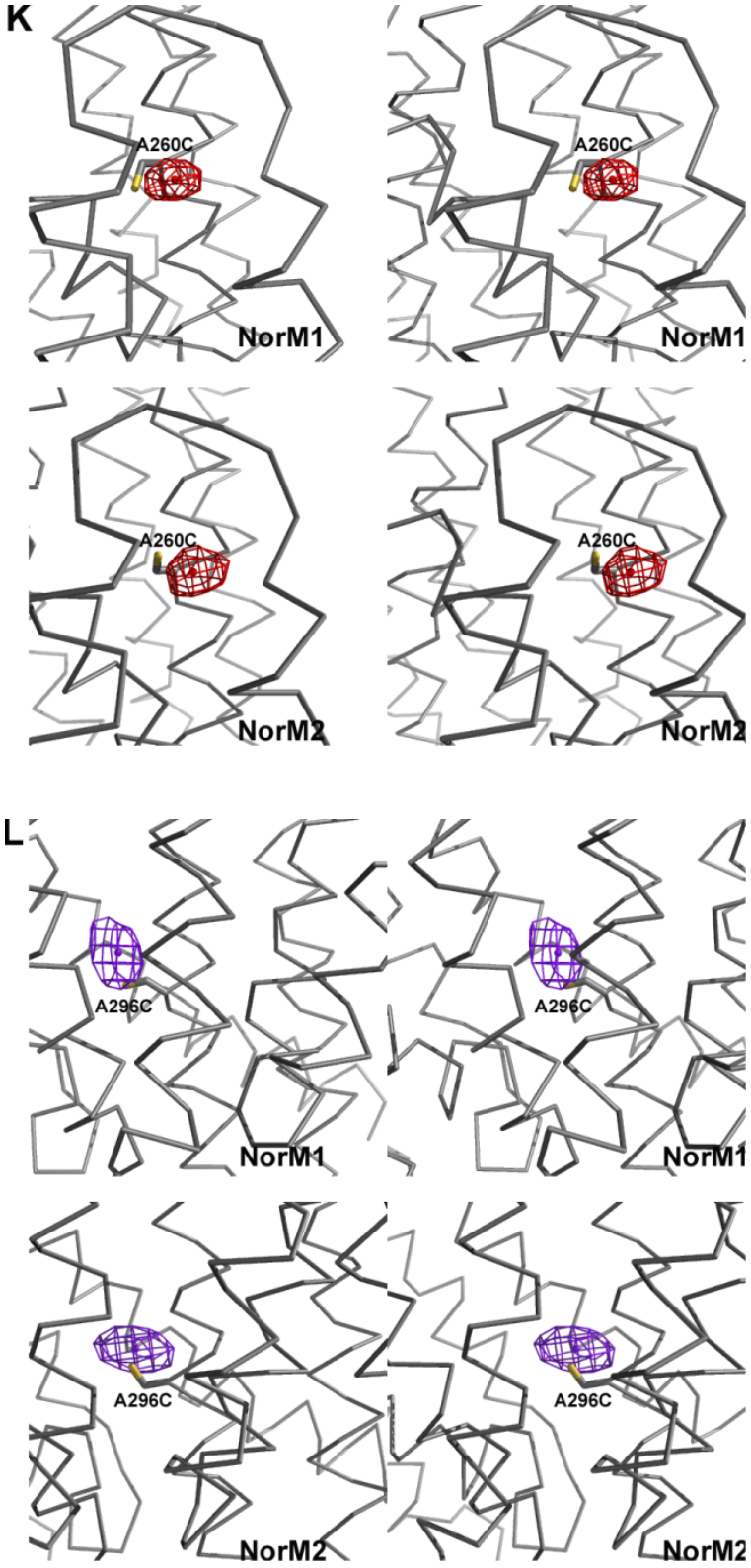


Figure 1.7.2 Continued

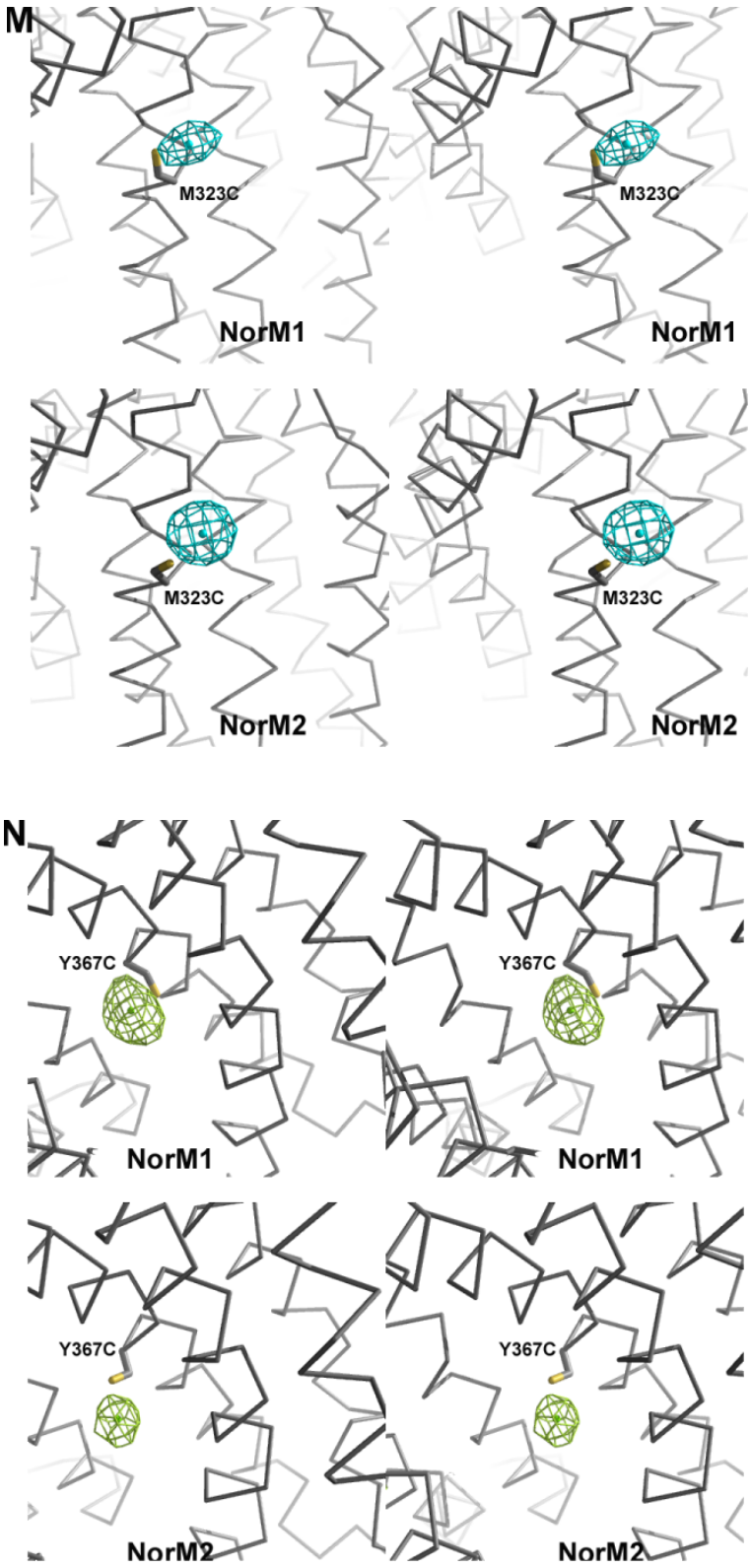


Figure 1.7.2 Continued

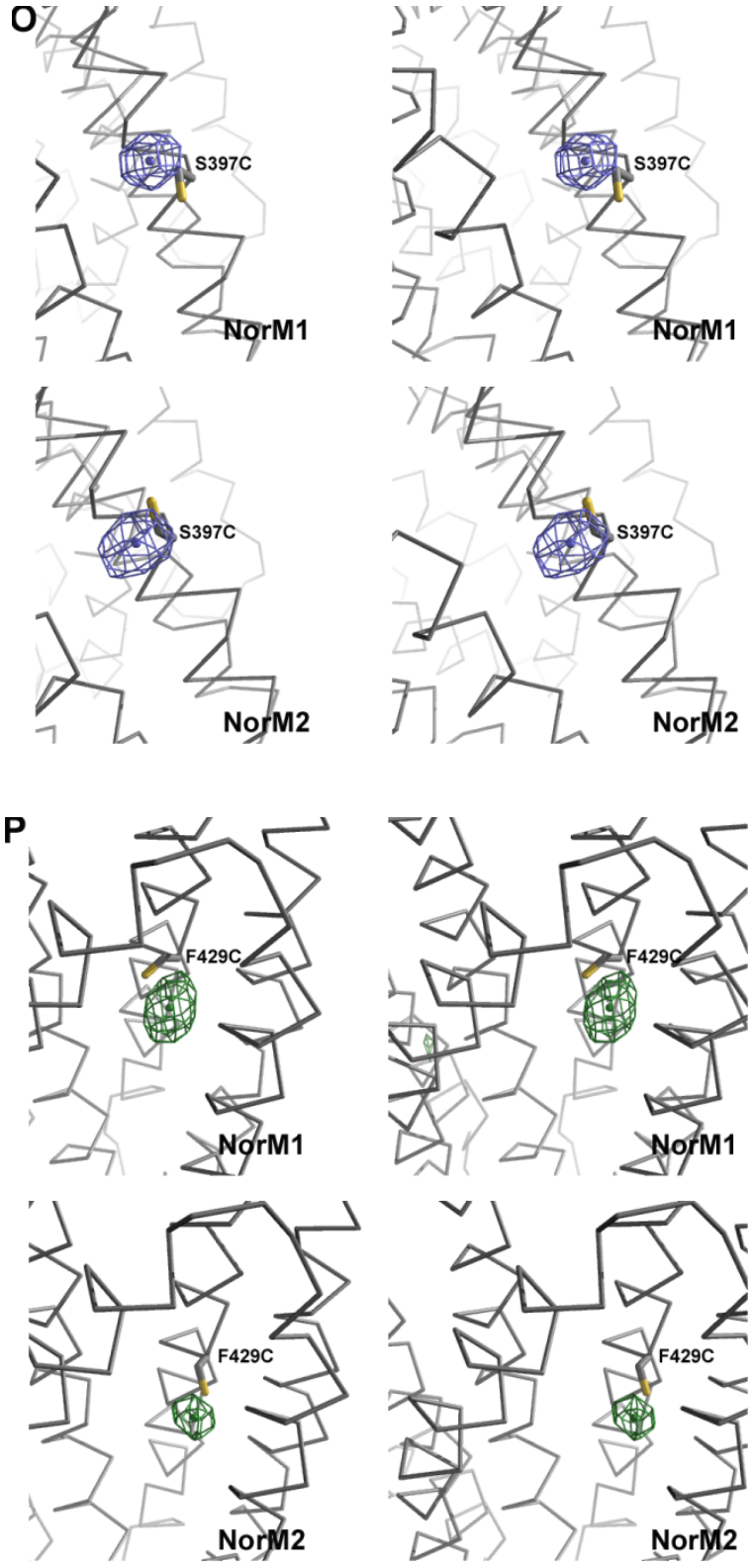


Figure 1.7.2 Continued

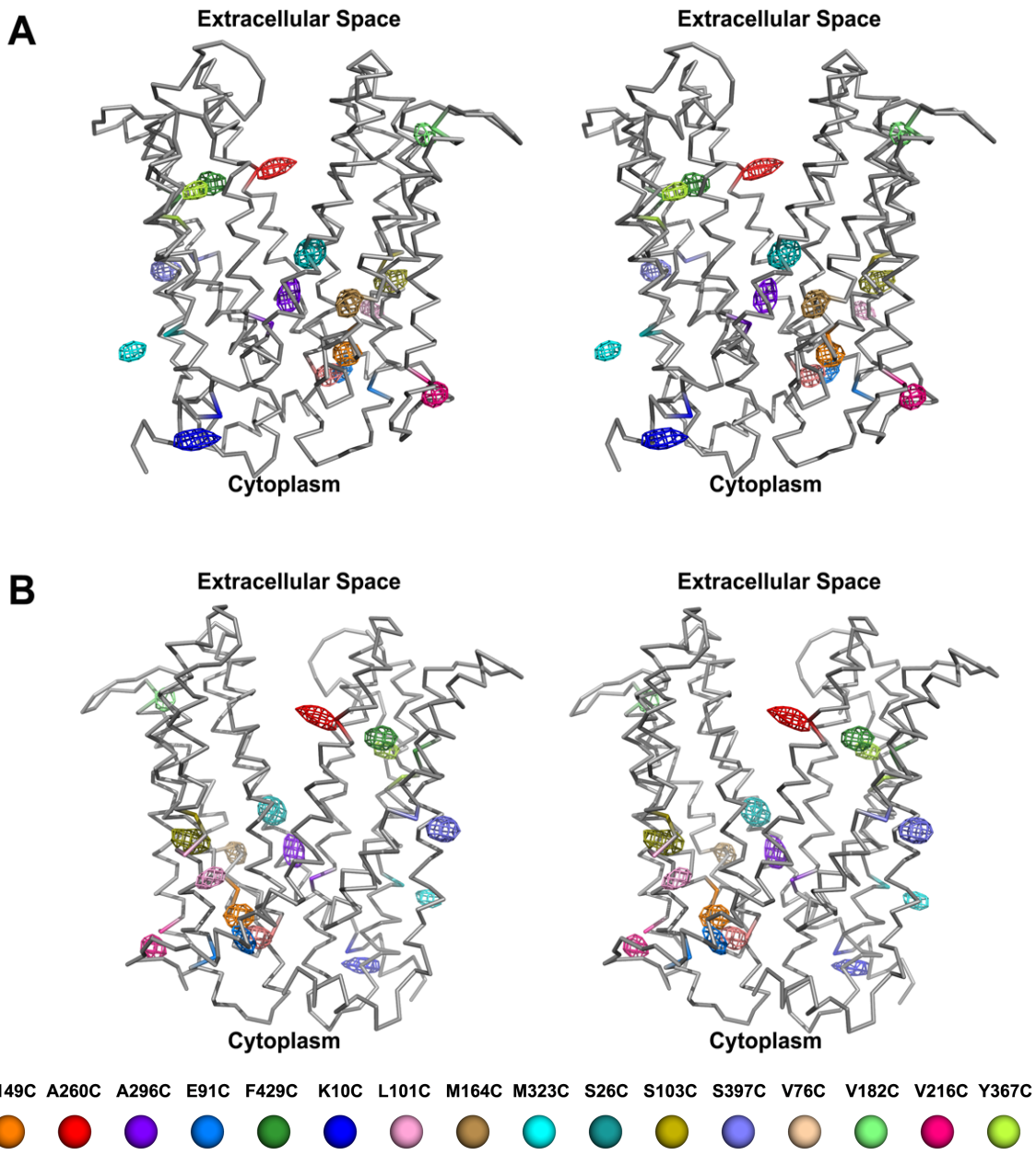
1.7.3 Composite Isomorphous Difference Fourier Maps by Real-space Transformation

A composition of all the isomorphous difference Fourier maps of the 16 cysteine mutants (Figure 1.7.2) was generated by real-space map transformation of each isomorphous difference map to the coordinates of the model from Crystal 2, and is shown in Figure 1.7.3.

Collectively, the isomorphous difference Fourier syntheses of these 16 individual cysteine substitution mutants verified the position and identity of at least one residue from each of the 12 TMS. These results confidently support the topology and the fold of the MATE family represented by the two NorM-VC structures for the first time. The following chapter will describe and discuss these structures and their functional implication in-depth.

Chapter 1, along with chapter 2 in part, have been accepted for publication of the material as it may appear in the journal Nature, 2010, He, Xiao; Szewczyk, Paul; Karyanki, Andrey; Evin, Riah; Hong, Wen-Xue; Zhang, Qinghai; and Chang, Geoffrey, 2010. The dissertation author is the primary investigator and author of this paper.

Figure 1.7.3 Topology verification of NorM-VC structure. Stereo view of NorM1 and NorM2. (A) Front and (B) back (rotated 180° about the pseudo molecular 2-fold axis relating the two halves of the molecule) views showing 16 mercury peaks corresponding to single cysteine mutations as viewed from membrane plane in NorM1 in the ASU. 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: 5.0 σ (orange); A260C: 4.0 σ (red); A296C: 4.5 σ (purple blue); E91C: 4.0 σ (marine); F429C: 4.0 σ (forest); K10C: 3.5 σ (blue); L101C: 4.0 σ (pink); M164C: 5.0 σ (sand); M323C: 5.5 σ (cyan); S26C: 4.0 σ (deep teal); S103C: 4.5 σ (olive); S397C: 5.0 σ (slate); V76C: 7.0 σ (deep salmon); V182C: 4.5 σ (lime); V216C: 6.0 σ (hotpink); Y367C: 5.0 σ (yellow green). (C) Front and (D) back views showing the 16 mercury peaks for the same set of single cysteine mutants in NorM2 of the ASU. A149C: 3.5 σ; A260C: 3.5 σ; A296C: 4.5 σ; E91C: 4.0 σ; F429C: 4.0 σ; K10C: 4.0; L101C: 4.5 σ; M164C: 4.0 σ; M323C: 5.5 σ; S26C: 3.5 σ; S103C: 4.0 σ; S397C: 4.5 σ; V76C: 7.0 σ; V182C: 4.5 σ; V216C: 5.0 σ; Y367C: 4.0 σ. Difference Fourier maps were obtained as described for in text from single cysteine mutant crystals soaked with mercury compounds (Table 1.4.1). The NorM-VC model is rendered as grey ribbon.


A
Extracellular Space
Cytoplasm
Extracellular Space
Cytoplasm
B
Extracellular Space
Cytoplasm
Extracellular Space
Cytoplasm
A149C
A260C
A296C
E91C
F429C
K10C
L101C
M164C
M323C
S26C
S103C
S397C
V76C
V182C
V216C
Y367C

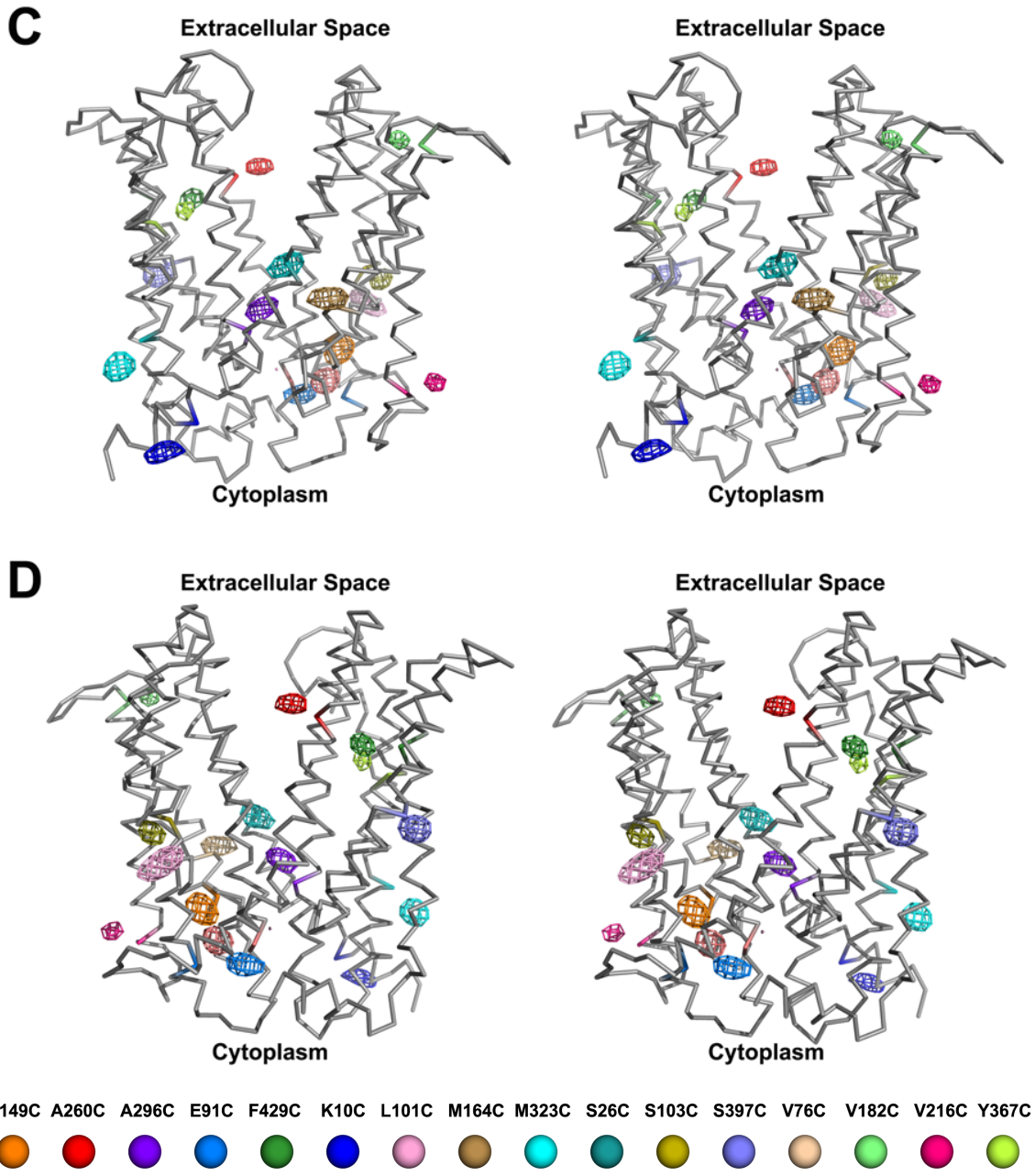


Figure 1.7.3 Continued

2

Outward-facing Conformation of NorM-VC

2.1 The NorM-VC Apo Structure

2.1.1 Structural Features of the NorM-VC Apo Struture

MATE transporters use either H^+ or Na^+ gradients across the membrane to drive substrate export. However, the coupling mechanism is not well understood. All MATE proteins share ~40% protein sequence similarity[22] (Figure 2.1.1), suggesting an overall conserved structure and transport function. To provide a basis for understanding the function of MATE transporters, the X-ray structure of NorM from *V. cholera* (NorM-VC) was solved to 3.65 Å in resolution.

This 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 each (TMs 1-6 and TMs 7-12) forming a large internal cavity open to the extracellular space (Figure 2.1.2). 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.

This structure also reveals a novel topology distinct from all other MDR transporter families[4,7,59,60,61]. 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 (Figure 2.1.3). The structural and functional similarities and differences between the two major secondary transporter families in eukaryotes, MATE and MFS, will be discussed to further extend in Chapter 3. Here, the structural features and the functional implications of the MATE representative NorM-VC will be the center of discussion.

In this outward-facing conformation of NorM-VC, two non-equivalent portals open within the lipid membrane, allowing the expulsion of substrates directly to the outer membrane leaflet or the extracellular space (Figure 2.1.4 A, B). These portals are formed by TMs 1 and 8 on one side and TMs 2 and 7 on the other (Figure 2.1.2). At the widest point within the lipid bilayer, the portals are ~12 Å wide and can accommodate the passage of its transport substrates (Figure 2.1.4 C, D). The volume of the internal cavity embedded within the lipid bilayer is ~4,300 Å3 (Figure 2.1.5),

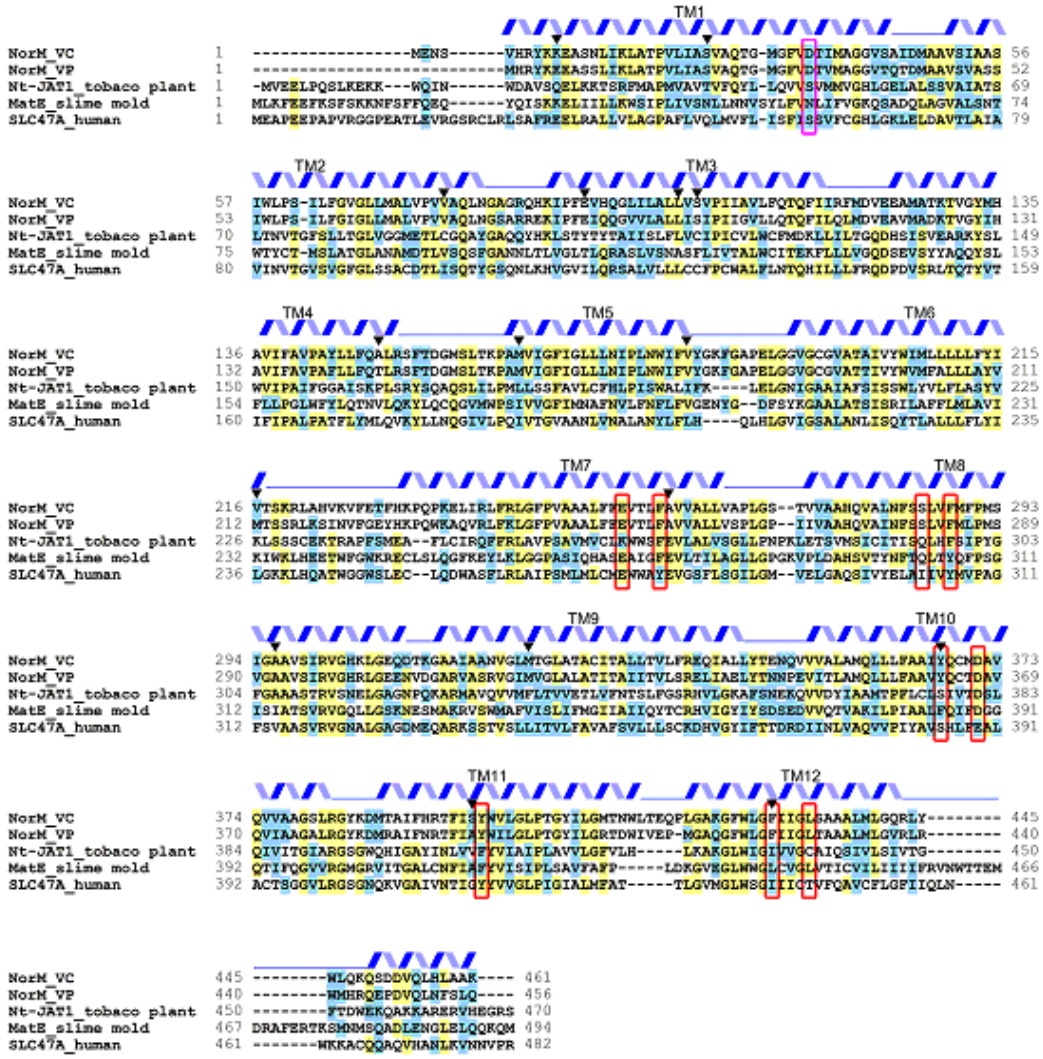
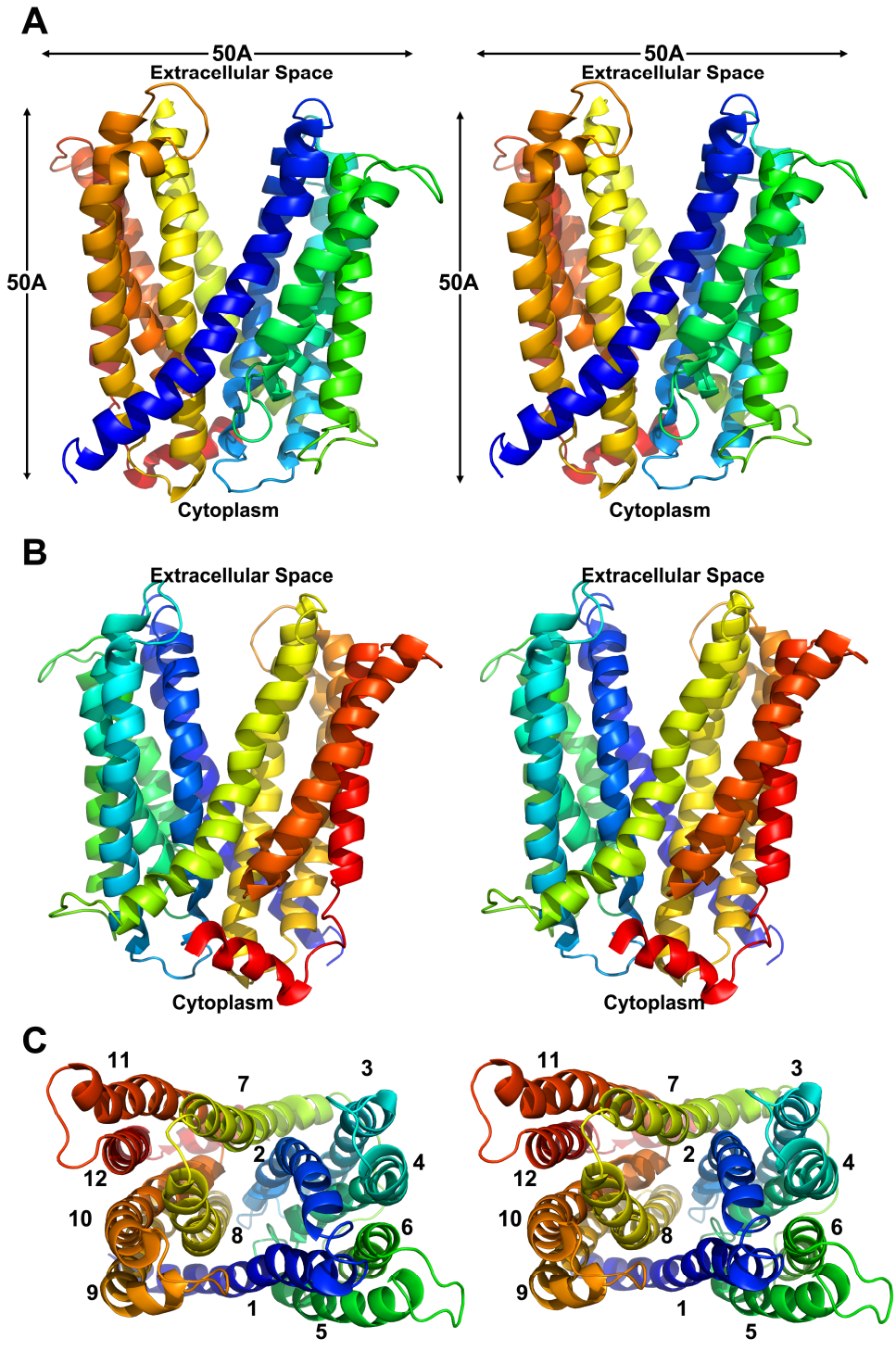


Figure 2.1.1 Amino acid sequence alignment of NorM-VC with selected orthologs. Similar residues are shaded blue, identical residues are shaded yellow. Threshold for shading is 60%. NCBI reference sequences used are: AAF94694.1 (NorM *V. cholera*), BAA31456.1 (NorM *V. parahaemolyticus*), CAQ51477.1 (Nt-JAT1 *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 (Figure 2.2.x and Table 2.2.1) and are highlighted with red rectangles. Residue D36 (highlighted with a pink rectangle) corresponds to D32 in NorM-VP.

Figure 2.1.2 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
50A
Extracellular Space
50A
Cytoplasm
50A
Extracellular Space
50A
Cytoplasm
B
Extracellular Space
Cytoplasm
Extracellular Space
Cytoplasm
C
11
7
3
2
4
12
10
8
6
9
1
5
11
7
3
2
4
12
10
8
6
9
1
5

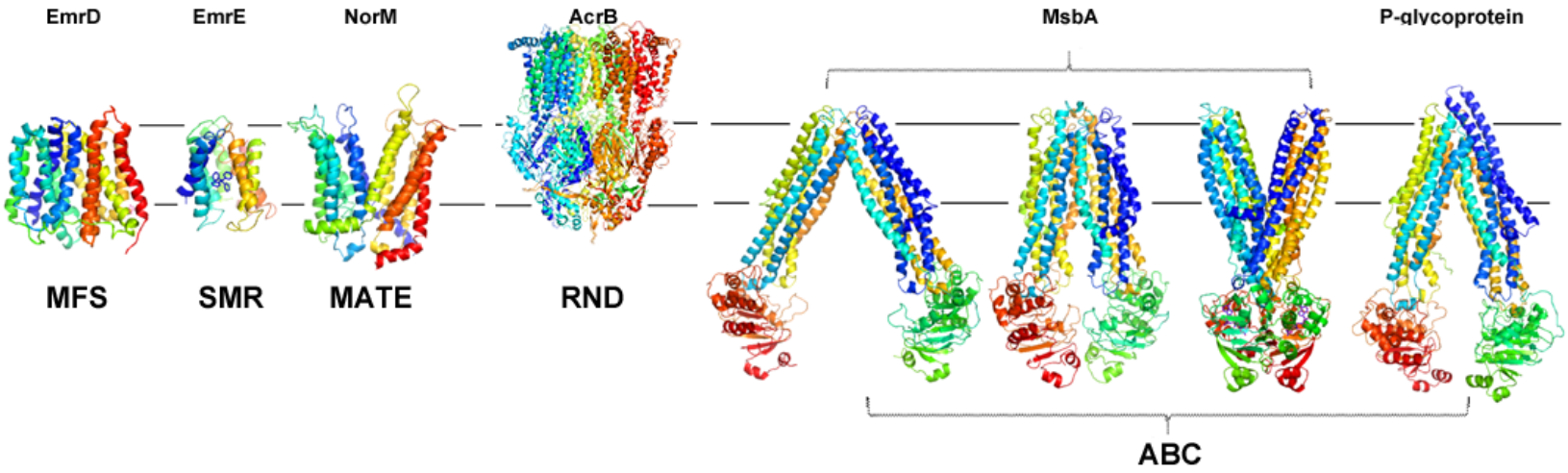
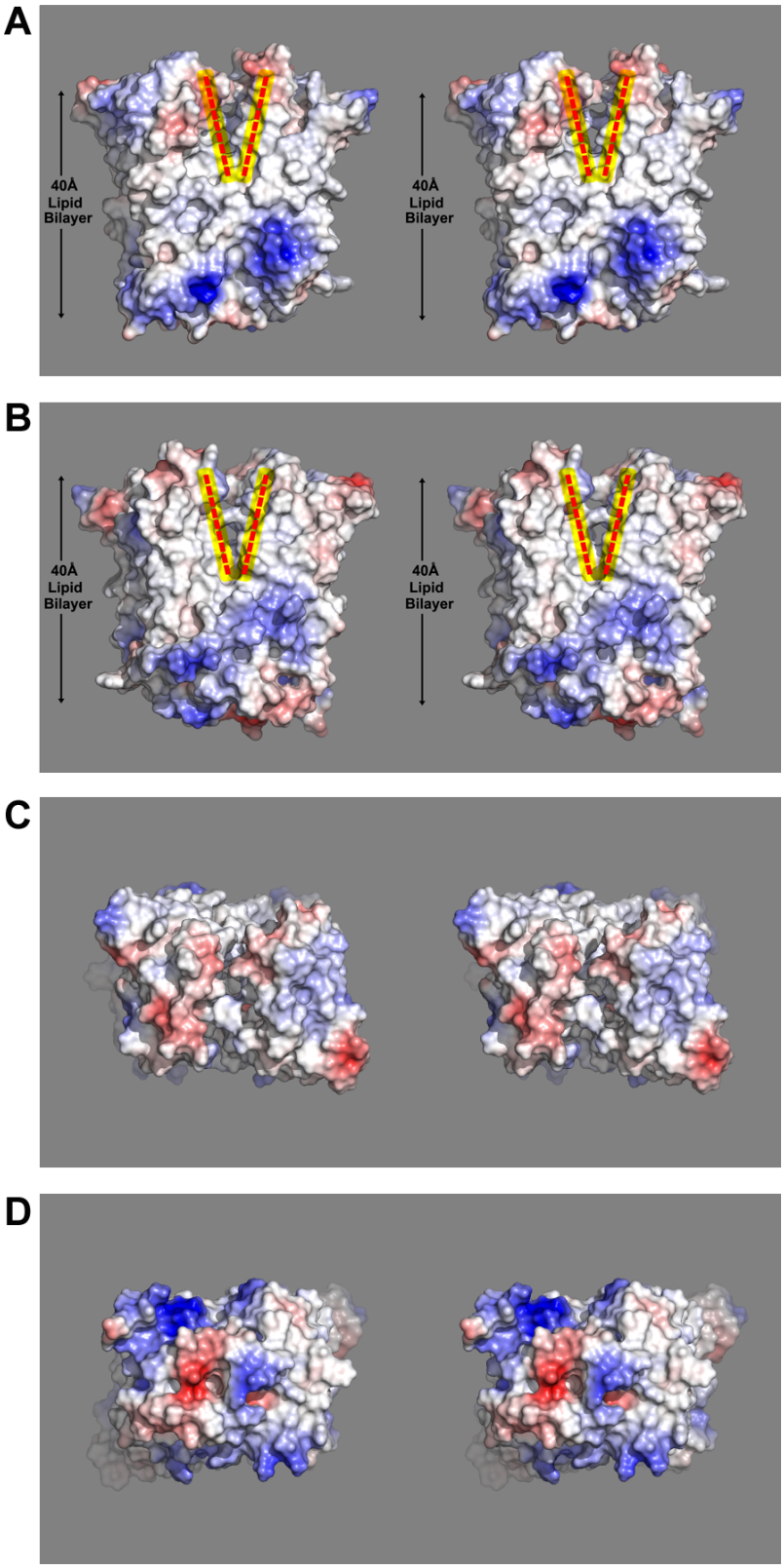


Figure 2.1.3 Structure representatives of the five families of MDR transporters. From left to right, the PDB codes for generating this presentation are: 2GFP[7], 3B5D[59], 3MKT[58], 2DHH[60], 3B5W[4], 3B5X[4], 3B60[4], and 3G61 [61]. Images were rendered in Pymol and Photoshop. The two black lines mark the two interfaces of plasma membrane and the immediate aqueous environment.

Figure 2.1.4 Electrostatic potential surface representation of NorM-VC. Stereo views of the portals formed by (A) TMs 1 and 8 and (B) by TMs 2 and 7. The views are the same as in Figure 2.1.2 A and 1B, respectively. The portals to the outer membrane leaflet are highlighted (yellow) and marked by dashed lines (red). 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) to red (negative). Hydrophobic and aromatic residues are colored in white.


A
40Å
Lipid
Bilayer
40Å
Lipid
Bilayer
B
40Å
Lipid
Bilayer
40Å
Lipid
Bilayer
C
D

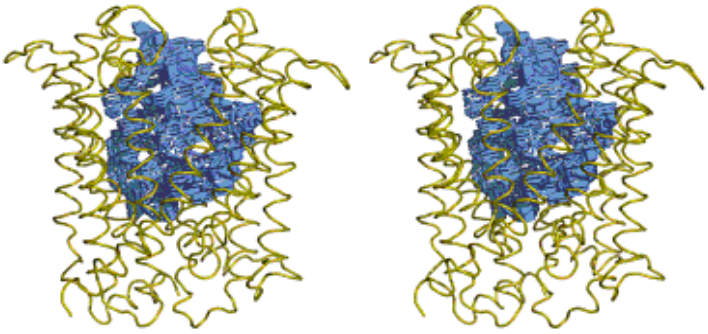


Figure 2.1.5 NorM-VC internal cavity. The stereo view of the volume of the internal cavity (blue) was generated using the program 3V Channel Extractor with a probe size of 1.6 Å [62]. The Cα trace of the protein is shown as golden ribbon.

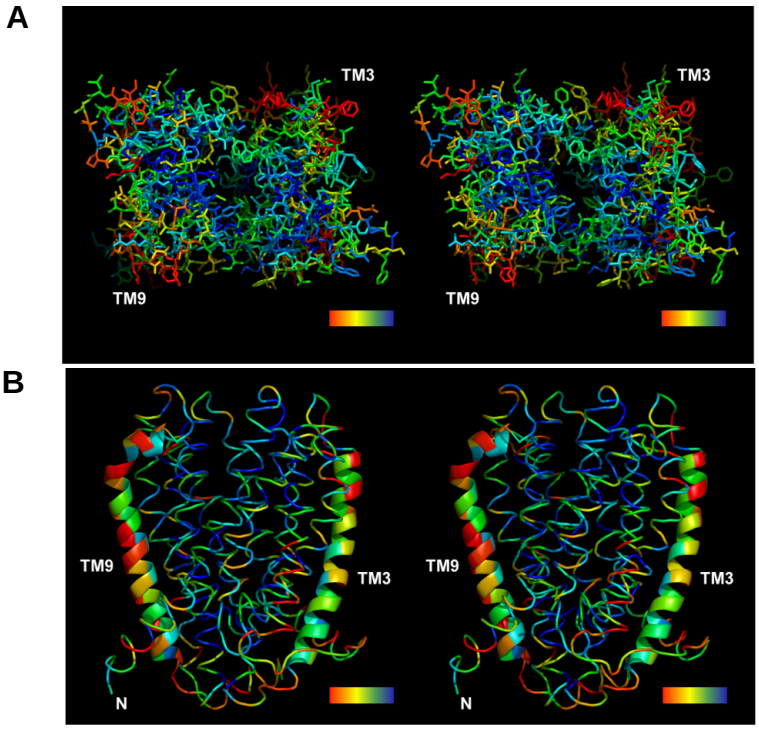


Figure 2.1.6 Conserved amino acid residues of NorM-VC. (A) Stereo view from the extracellular side showing residue conservation, colored as a rainbow going from most (blue) to least (red) conserved. The most conserved residues are located towards the internal cavity where cation and substrate binding are expected. (B) Side view with TM3 and 9 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 [63,64,65].

Table 2.1.1 Residues in the internal cavity

	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					
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

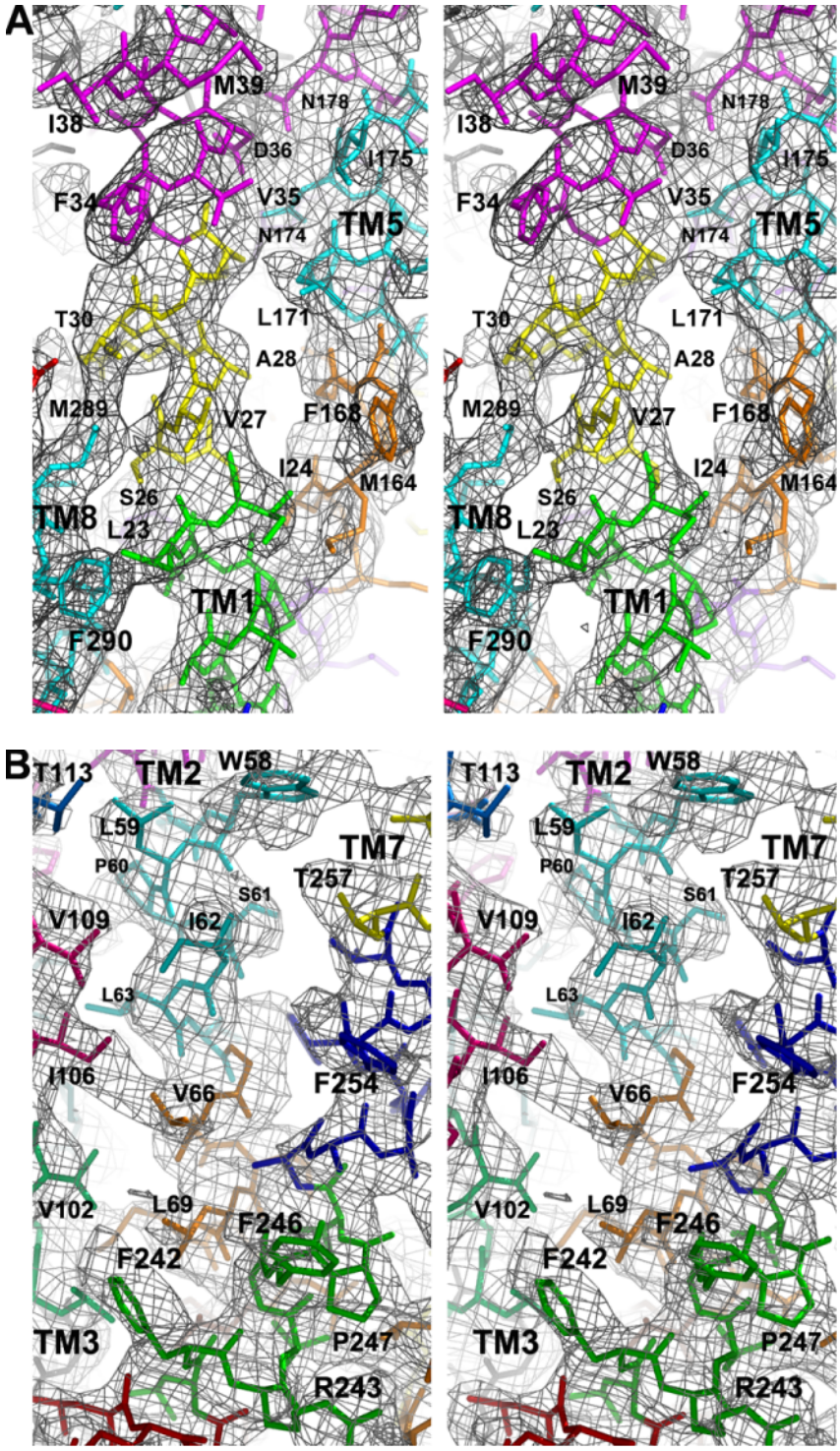
Notes:
Residues in the internal cavity, where Ala, Ile, Leu, Phe, Trp, Tyr, and Val are classified hydrophobic; Asn, Cys, Gln, Ser, Thr, Tyr Polar; Asp, His, Glu charged; and Phe, Trp, and Tyr aromatic. The Rb^+ column denotes residues bounding the cation site.

approximately four times larger than the substrate binding pockets of the transcription MDR regulators xPXR and BmrR [66,67], and nearly two thirds that of ABC exporters Pgp or MsbA [61,68]. 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 (Figures 2.1.2 and 2.1.6). 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 2.1.1).

2.1.2 Omit Density Maps for the NorM-VC Apo Struture

Multiple $F_{obs} - F_{calc}$ simulated annealing omit maps were calculated along the entire peptide chain to verify the correctness of the structure (Figures 2.1.7). Along with a series of 16 mercury-labeled single cysteine mutant crystals (Figure 1.7.2 and 1.7.3; Table 1.4.1), the topology of the strucuture and the peptide chain trace were validated throughout the entire model. 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 (Figure 1.7.2).

Figure 2.1.7 (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 4 Å and a map cushion of 2 Å. The windows of residues omitted from the model are shown in different colors. Maps B-D and F-K were contoured at 2.0 σ while A, E, and L-O were at 1.5 σ.


A
M39
N178
I38
D36
I175
F34
V35
TM5
N174
T30
L171
A28
M289
V27
F168
I24
M164
S26
TM8
L23
TM1
F290
B
T113
TM2
W58
L59
TM7
P60
T257
S61
V109
I62
L63
I106
V66
F254
V102
L69
F246
F242
TM3
P247
R243

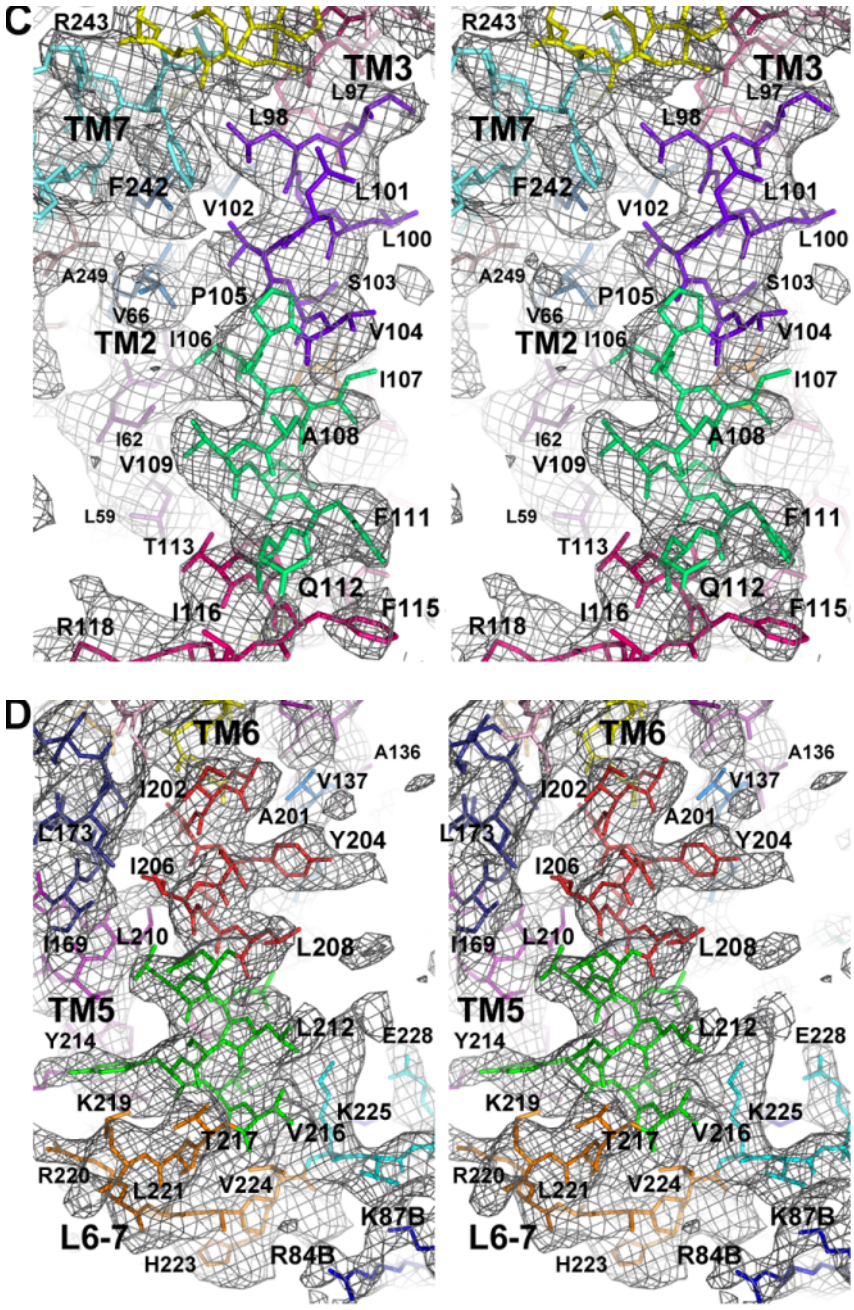


Figure 2.1.7 Continued

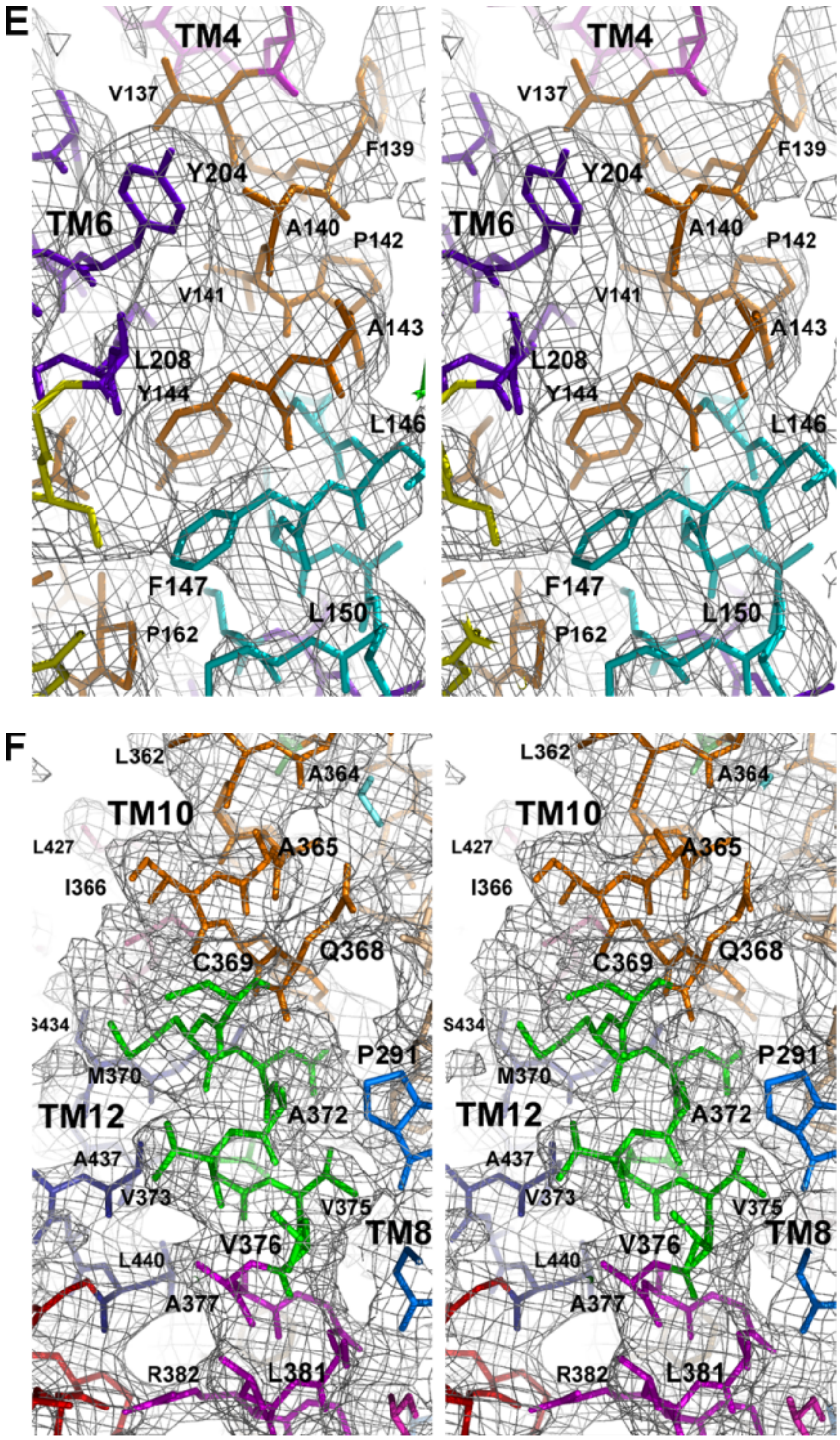


Figure 2.1.7 Continued

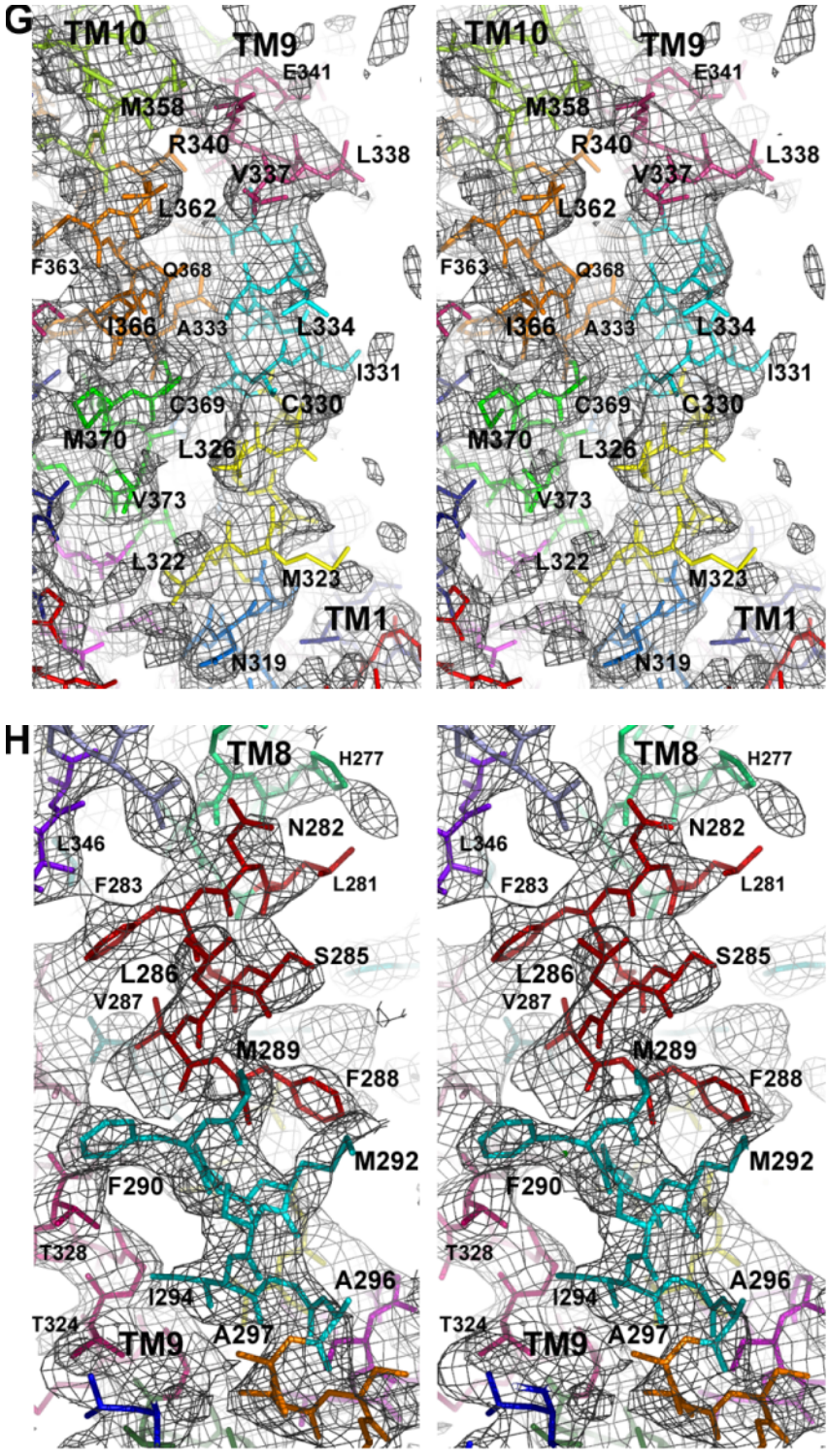


Figure 2.1.7 Continued

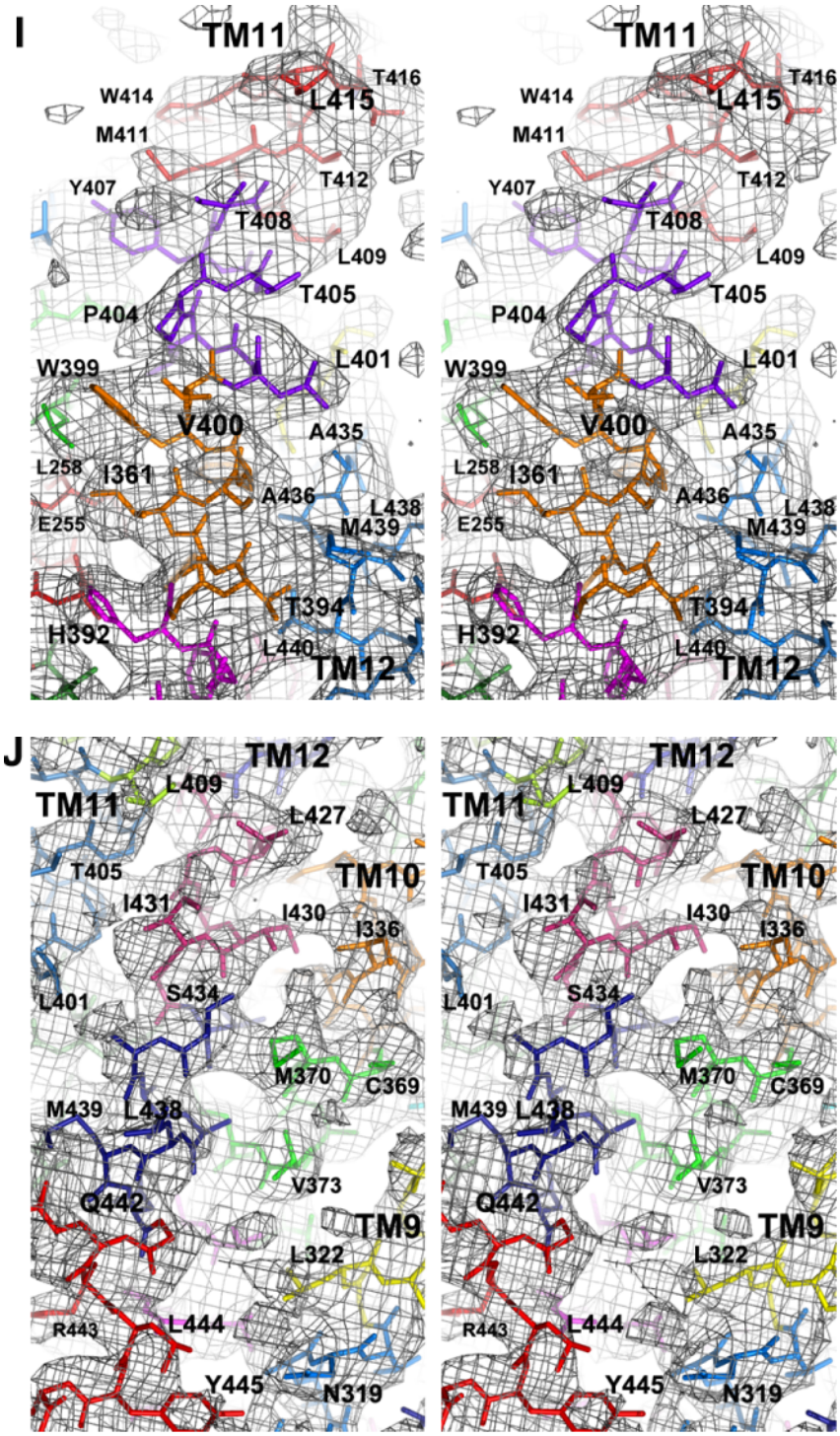


Figure 2.1.7 Continued

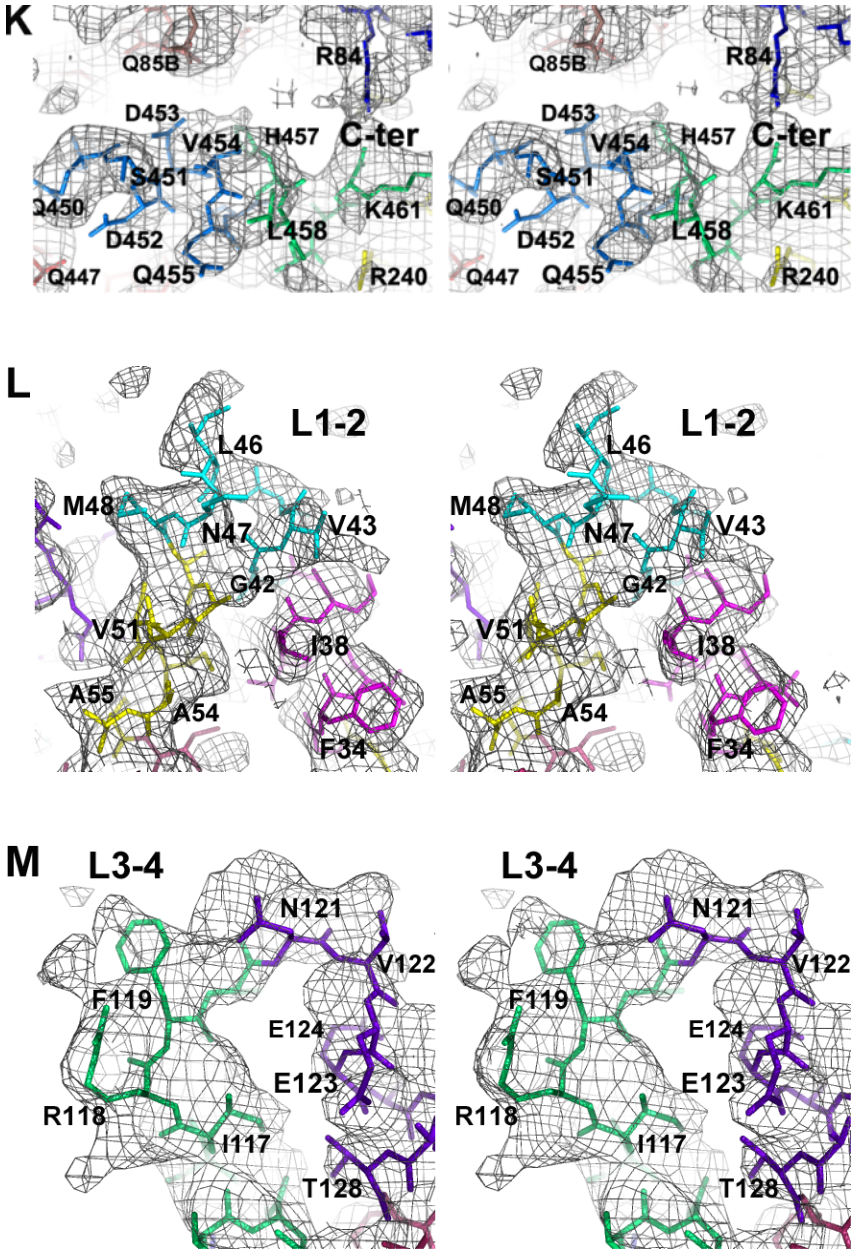


Figure 2.1.7 Continued

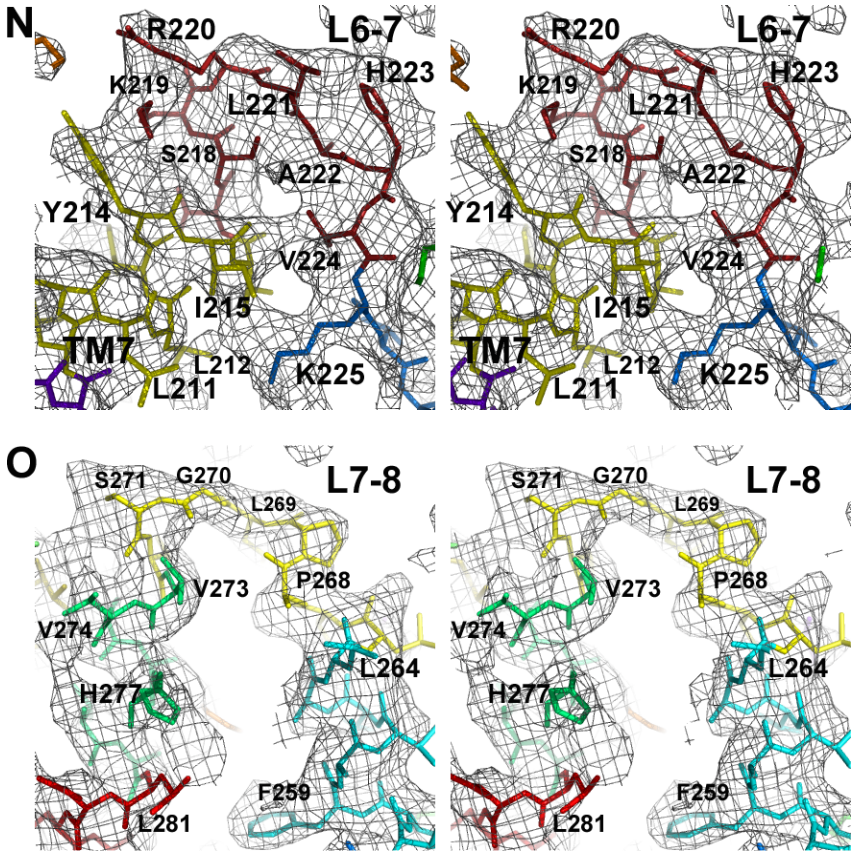


Figure 2.1.7 Continued

2.2 The Cation-bound Structure of NorM-VC

2.2.1 The Cation-binding Properties of the Outward-facing Conformation of NorM-VC

In 1998, Morita *et. al.* cloned the first NorM ortholog from *V. parahaemolyticus* (VP), which was very closely related to NorM-VC, 76% sequence identity and 86% similarity[69,70]. NorM-VP was then indicated as a Na^+-driven MDR transporter of the MATE family[10]. The NorM-VC apo-structure presented earlier is an outward-facing conformation and likely represents a substrate-releasing state, which by the premise has lower

affinity for the substrates and higher affinity for Na^+ [71]. Thus, analogs of Na+, such as Rb+ and Cs+, were used to derivatize the same NorM-VC wild-type crystal form to probe the presence of a cation-binding site. A Rb^+-bound structure was then determined to 4.2 Å and nearly identical to the 3.65 Å apo-structure. The structural model was also verified by multiple Fo-Fc SA omit density maps (Figure 2.2.1).

2.2.2 The Cation-binding Site in the Outward-facing Conformation of NorM-VC

The Rb^+-bound structure revealed a monovalent cation-binding site within the internal cavity of the transporter. The isomorphous difference Fourier maps, calculated in a similar fashion as the isomorphous difference Fourier maps showing the corresponding Hg sites as a result of single cysteine substitution mutation (Figure 1.7.2 and 1.7.3), reveal the presence of Rb^+ or Cs^+ in these cationic compound-derivatized crystals (Figure 2.2.2). In this outward-facing conformation, 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 (Figure 2.1.3; Table 2.1.1).

.

Figure 2.2.1 (A-F) Stereo views of Fo-Fc simulated annealing omit density for NorM-VC Rb+-bound strucuture with multiple density maps superimposed. The Fo-Fc density maps (grey) were generated in the same manner as Fig. S5. The windows of residues omitted from the model are shown in different colors. Maps (A-F) were contoured at 2.0 σ

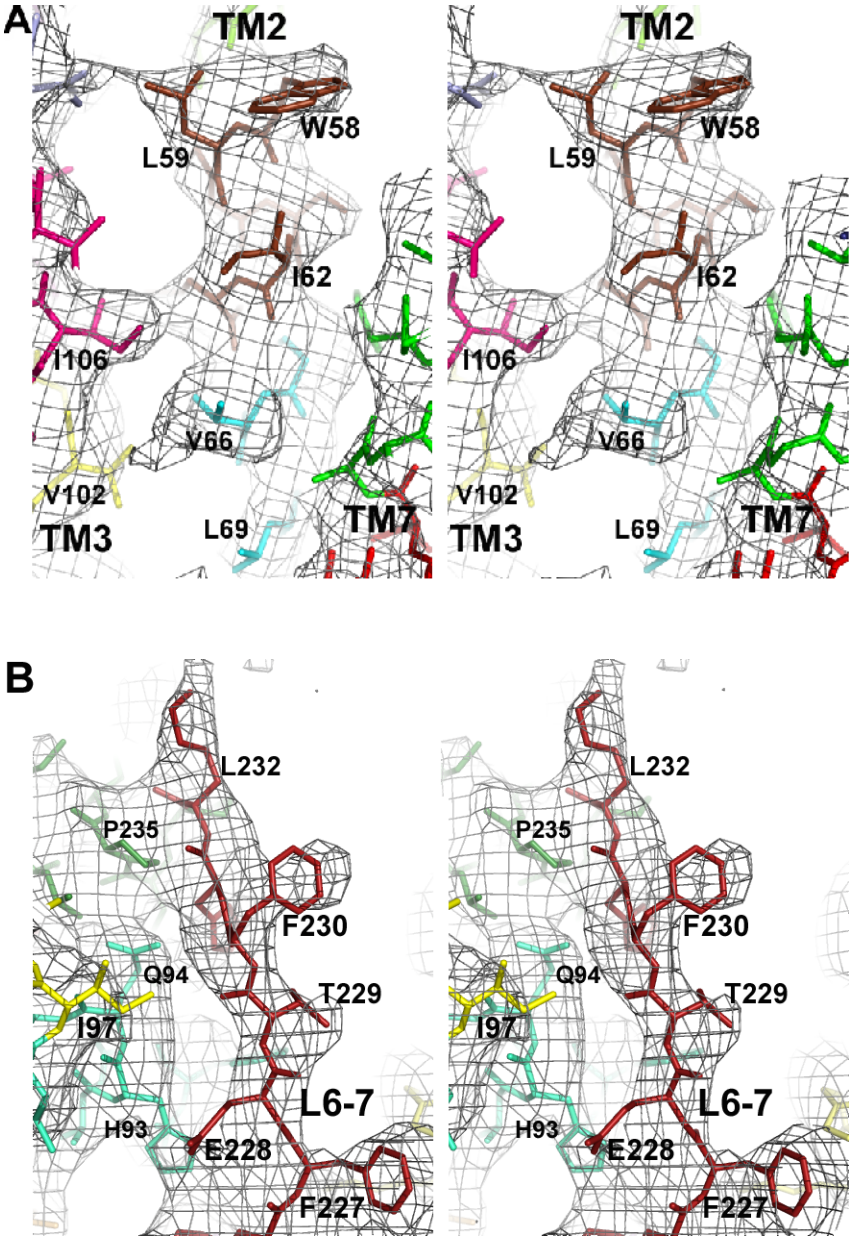


Figure 2.2.1 Continued

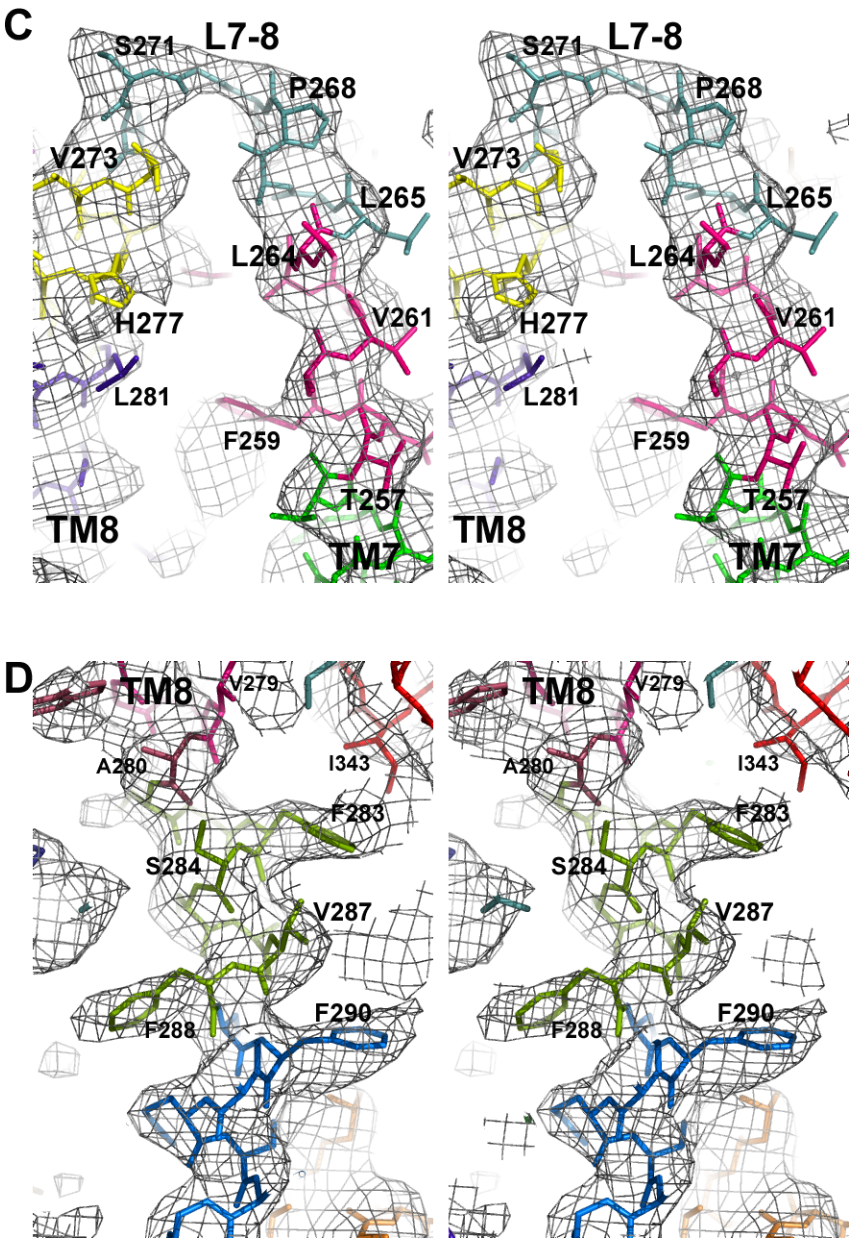


Figure 2.2.1 Continued

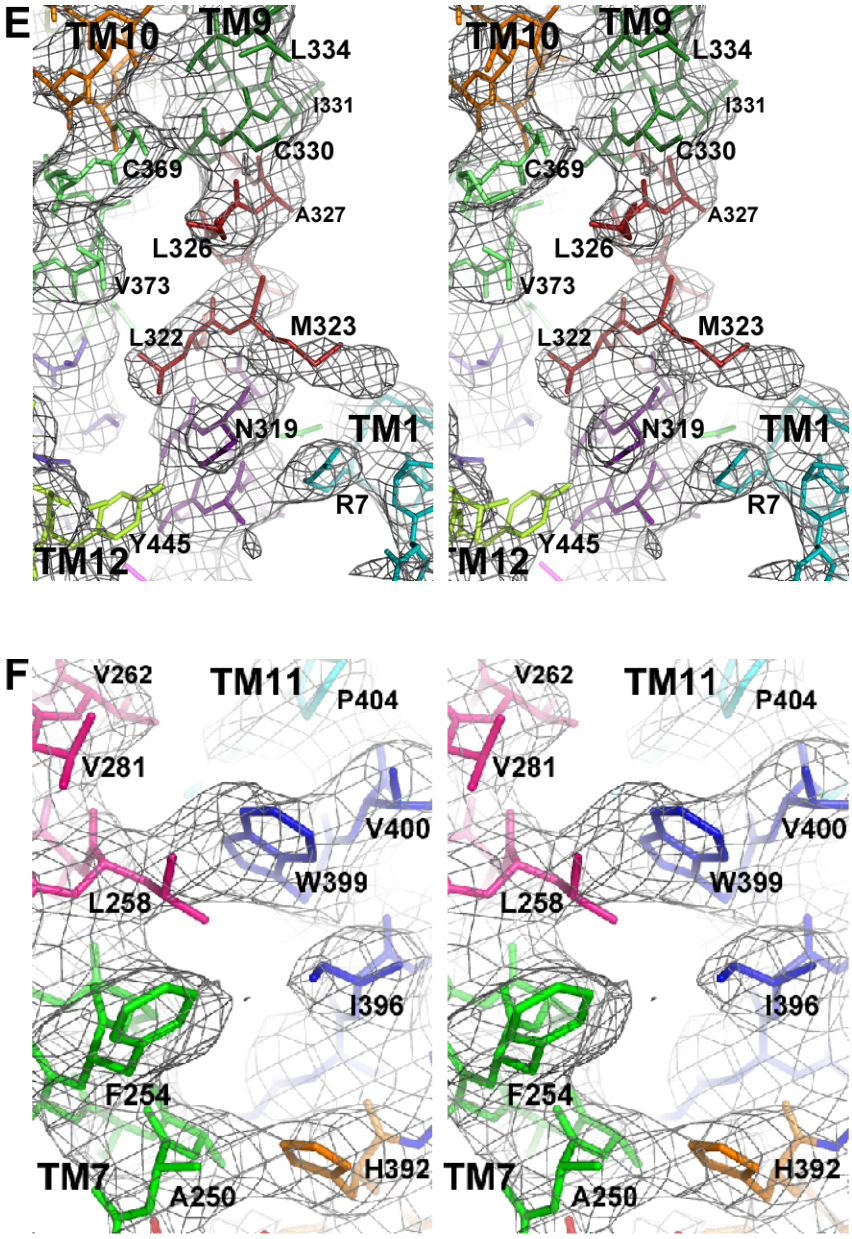


Figure 2.2.1 Continued

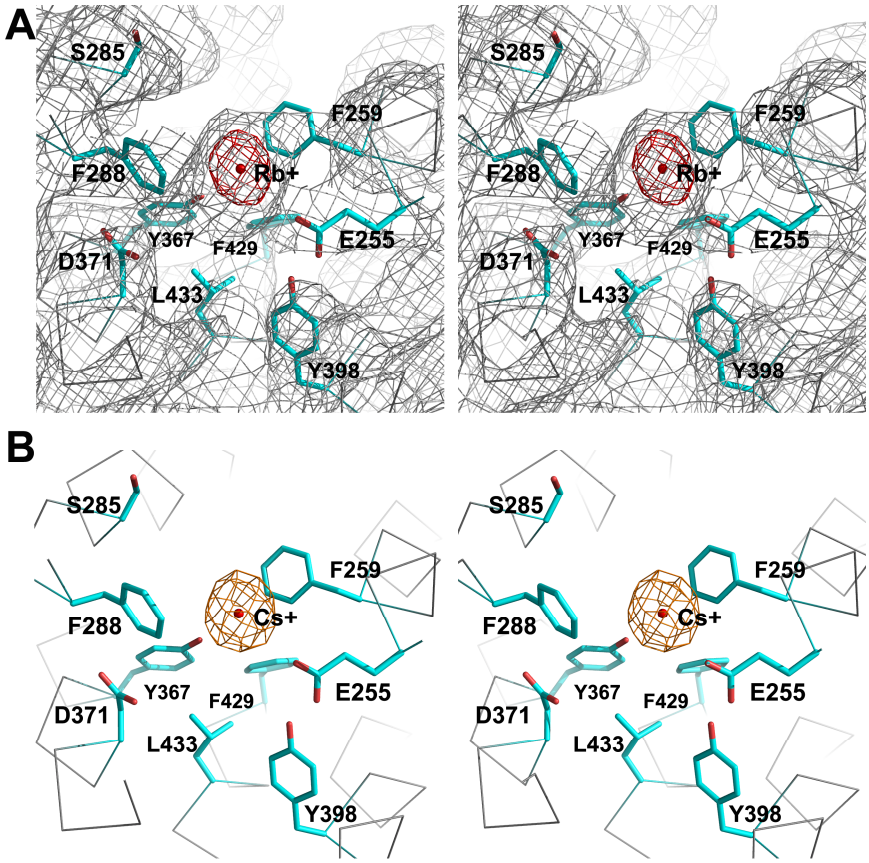


Figure 2.2.2 The cation-binding site of NorM-VC and mechanism of transport. (A) Composite Fo-Fc simulated annealing omit map (gray mesh, contoured at 2.0 σ) of the cation-binding site superimposed with an isomorphous difference Fourier map calculated between Crystal 2 (containing Rb^+) and Crystal 3 (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 2.1.1). (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) (Table 1.4.1).

2.2.3 The Residues Important for the Coupling of Electrochemical Gradient and Substrate Transport in the MDR MATE Transporter NorM

Previous studies have identified three conserved acidic residues that are critical for transport function in NorM-VP[42], a closely related

ortholog of NorM-VC. 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 (Figure 2.2.2). For NorM-VP, both of these negatively charged residues are important for transport function as certain mutations at these positions decouple substrate/Na+ antiport[42]. Like NorM-VP, mutations at residue 255 in NorM-VC affect protein folding and, as a consequence, do not express well[42].

2.3 Loss of Cation binding Due to Mutations in the Cation-binding site

2.3.1 Loss of Rb^+-binding by ΔD371A Mutant

Substitution mutation from D371 to alanine abolishes Rb^+-binding in the outward-facing conformation of NorM-VC (Figure 2.3.1 A). Mutant ΔD367A of NorM-VP (corresponds to ΔD371A of NorM-VC) disrupts transport activity in a cell-based transport assay[42]. Both Na^+-dependent efflux of substrate EtBr and the substrate uptake-evoked Na^+ efflux by this ΔD367A mutant were impaired compared to wild-type control.

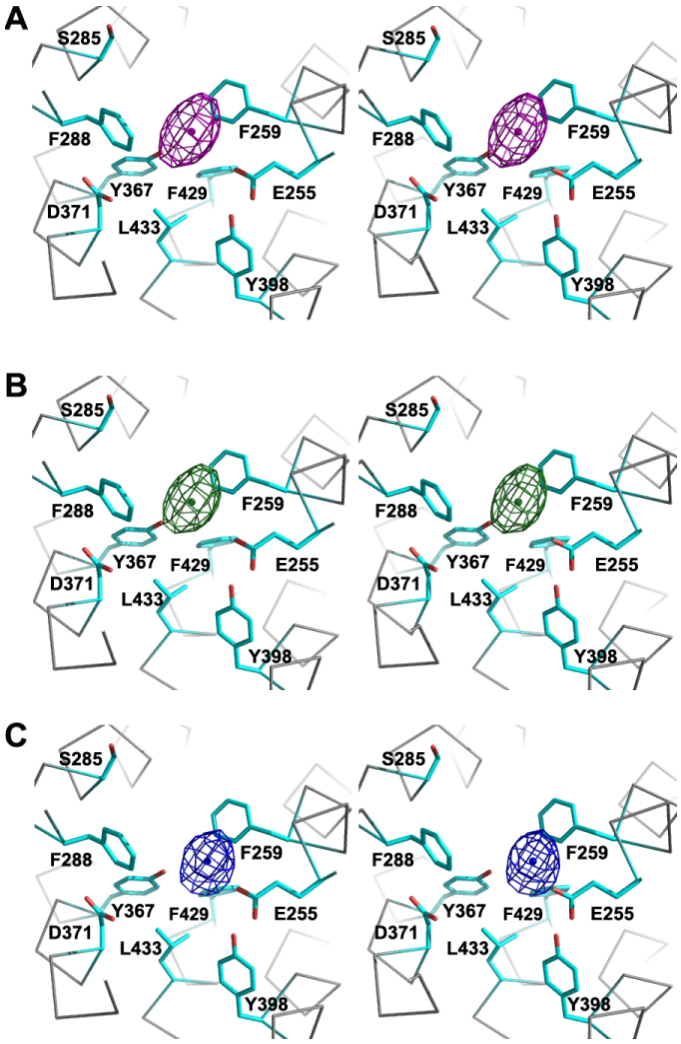


Figure 2.3.1 Loss of cation-binding as a result of ΔD371N and ΔD371A mutations. (A) Shows a loss of Rb^+-binding in the ΔD371N mutant as indicated by a -5.0 σ peak (purple mesh) in the isomorphous difference Fourier map calculated between a RbCl-soaked ΔD371N crystal and Crystal 2, which contains Rb^+. 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^+. (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 in the supplemental text. Residues that make up the cation-binding site are marked. Isomorphous difference Fourier maps with native data reveals no peaks for A-C.

2.3.2 Loss of Rb^+- and Cs^+-binding by D371N Mutant

Substitution mutation from D371 to asparagine abolish Rb^+- and Cs^+-binding in the outward-facing conformation of NorM-VC (Figure 2.3.2 B, C). Mutant ΔD367N of NorM-VP also disrupts transport activity to certain extent[42]. Only the Na^+-stimulated efflux of EtBr appeared to be deteriorated in this mutant. The uptake of three structurally unrelated substrates seemed to cause Na^+ efflux to an extent considerably close to that by the wild-type transporter.

The absence of Rb^+ or Cs^+ at the cation-binding site as a consequence of mutations D371A and D371N indicates the cation-binding site residues are critical for the binding of monovalent cations in the outward-facing conformation of NorM-VC. The disruption of transport, in a substrate- and electrochemical gradient-dependent manner due to a homologous mutation in a very closely related ortholog NorM-VP[42], suggests these residues are important, possibly mediating the direct coupling of the substrate and cation antiport

2.3.3 Other Potentially Important Residues for the Transport Funtions of NorM Orthologs

Residues D32 and E251 were also identified as critical for transport functions in NorM-VP. However, in the 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. Neither Rb^+ nor Cs^+ was observed near D36 in the wild-type structure. In NorM-VP, the homologous residue D31 is the only residue that is absolutely required for transport functions, tolerating no other substitutions, in the cell-based transport assays for the antiport of substrate and Na^+[42]. Therefore, it is very likely that this residue plays an important role in other conformations during the transport cycle.

Due to the insufficient expression and/or poor stability of a variety of NorM-VC E255 mutants, the extent to which this residue is affecting the cation-binding properties could not be pursued with success. Nevertheless, being a residue composing the cation-binding site in NorM-VC and whose homologous residue in NorM-VP is critical for transport functions [42], E255 is also likely to be a key residue coupling substrate efflux with Na^+ import. Collectively, these findings suggest that these and possibly other surrounding residues at the Rb^+ site of the NorM-VC outward-facing structure are crucial for the recognition and binding of cations of an inward gradient across the plasma membrane.

2.4 The Conservation of the Cation-binding site in the Multidrug Type of NorM-VC Homologs

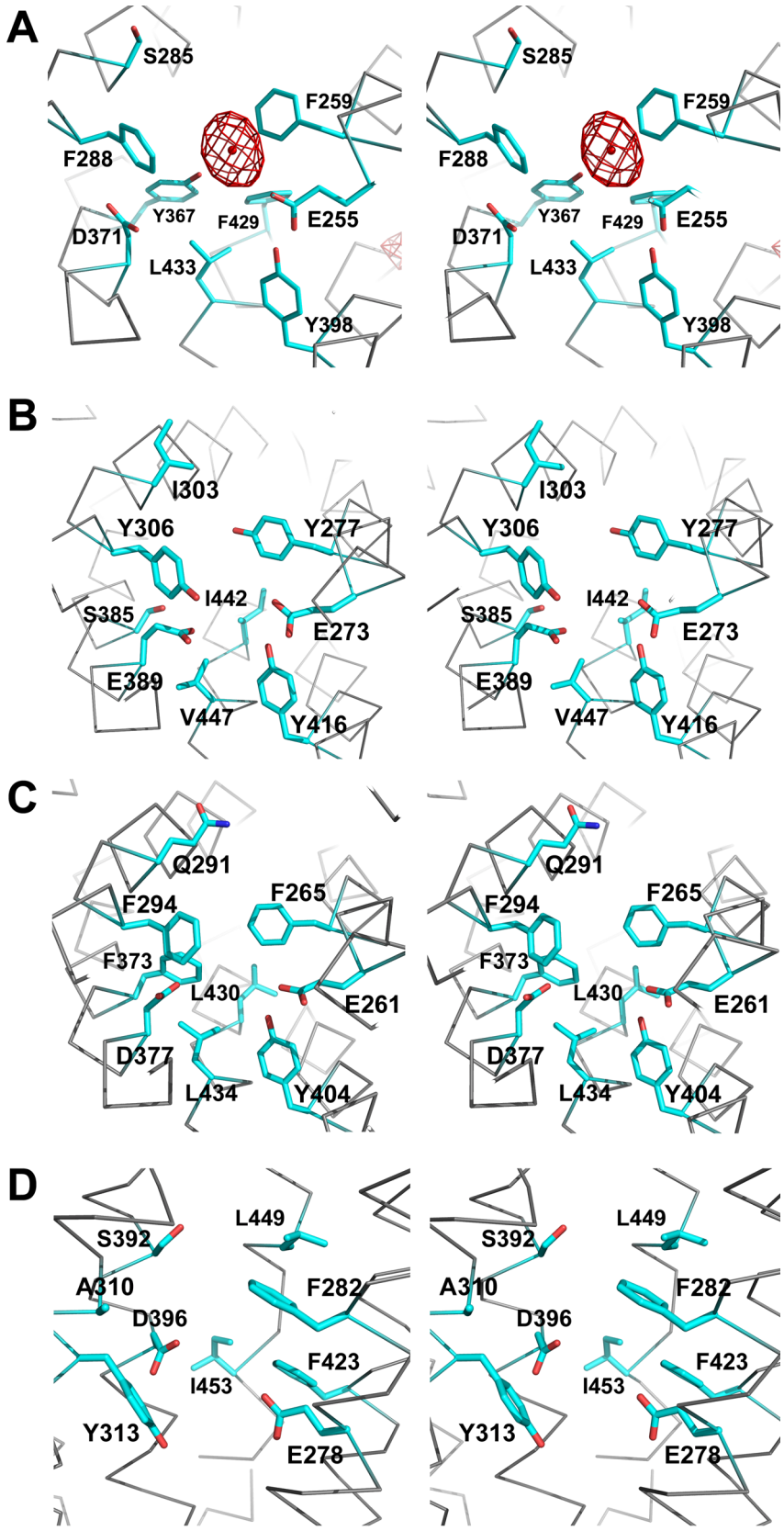
To further investigate whether such a cation-binding site in the outward-facing conformation could exist in other relatively distantly related NorM-VC orthologs, several homolog models of NorM-VC were produced computationally. The rational for this approach is that sequence homology between two orthologs should dictate the most functionally conserved features. As such, only the aligned regions in a selected ortholog sequence and that in the NorM-VC sequence is build in the homolog model, using the apo-structure and CNS v1.2 suites. Since it is not practical to produce a homology model for every MATE homolog of known sequence, one representative from each of the following MATE subgroup is chosen for the homology modeling: bacterial, mammalian, and plant MATE transporters. The selected MATE homologs were chose also due to their biological importance as indicated by previous studies.

2.4.1 Mammalian Orhologs

Human MATE homolog hMATE1 plays critical roles in waste and drug deposition and is likely the counterpart of NorM in bacterial MDR. Human has two homologs of NorM-VC. The hMATE1 or hSLC47A1 is the ortholog of the bacterial NorM proteins. The other human homolog hMATE2 or

hSLC47A2 appears to be the result of genomic duplication of the hMATE1 and has a kidney specific variant often referred as hSLC47A2K. The hMATE1 primarily expresses in the brush border of the renal proximal tubules and has been identified as the major player mediating the final excretion of metabolic wastes and various pharmaceutically valuable drugs. These include a few commonly used anti-inflammatory, anti-arrhythmic, anti-cancer, and anti-diabetic drugs[9]. Targeted gene-disruption in mouse ortholog mMATE1[72] and human genetic variations as nonsynonymous SNP in the hMATE1 locus[73] have linked the transport activity of mammalian MATE1 directly to the efficacy of a first-line type II diabetes drug metformin. Based on these functional studies, hMATE1 is selected as the mammalian representative for the homology modeling. The homolog model of hMATE1 shows a similar cation-binding site (Figure 2.4.1 B) as that of NorM-VC in the outward- facing comformation (Figure 2.4.1 A).

Figure 2.4.1 Putative cation-binding site in the NorM orthologs. A similar cation-binding site is found in the (A) NorM-VC, (B) hMATE1, (C) NorM-XC, (D) atAFL5 homology models. The side chains of the residues composing each putative cation-binding site were highlighted as cyan sticks and labeled in each ortholog's sequence. Homolog models were made by threading the homologous sequences of each orthologs along the corresponding segments on NorM-VC ourward-facing apo-model, using the CNS v1.2 suite.


A
S285
F259
F288
Y367
F429
E255
D371
L433
Y398
S285
F259
F288
Y367
F429
E255
D371
L433
Y398
B
I303
Y306
Y277
I442
S385
E273
E389
V447
Y416
I303
Y306
Y277
I442
S385
E273
E389
V447
Y416
C
Q291
F294
F265
F373
L430
E261
D377
L434
Y404
Q291
F294
F265
F373
L430
E261
D377
L434
Y404
D
L449
S392
A310
F282
D396
F423
I453
Y313
E278
L449
S392
A310
F282
D396
F423
I453
Y313
E278

2.4.2 Bacterial Ortholog

NorM from *Xanthomonas campestris* (XC) was selected to represent a relatively distantly related bacterial ortholog of NorM-VC. *X. campestris* is widely spreading plant pathogen, causing a variety of plant diseases. A nosocomial opportunistic pathogen from the same family, *Xanthomonadaceae*, has been shown to possess remarkable capacity for multidrug and heavy metal resistance[74]. In NorM-XC, a similar cation-binding site is also present in the outward-facing conformation (Figure 2.4.1 C).

2.4.3 Plant Ortholog

ALF5 is a multidrug MATE transporter expressed strongly in the root epidermis of *A. thaliana* where it effluxes toxic compounds and protects the root tissue from the inhibitory elements in soil[20]. ALF5 can transport cation tetramethylammonium (TMA), a similar compound to the mammailian and bacterial multidrug MATE substrate TEA. A deletion mutant *alf5* with 29 bp in the coding region of TM5 in ALF5 protein was found deficient in tolerating the contaminants in media made with 0.01% (w/v) commercial polyvinylpyrrolidone (PVP) and transporting several other xenobiotics[20]. The homolog model of ALF5 shows a cation-binding

site with striking resemblance to that observed in NorM-VC likely to exist in this plant MDR homolog (Figure 2.4.1 D).

From the above survey, I found functionally similar multidrug MATE homologs from distantly related taxa are likely to share a similar cation-binding site as observed in the outward-facing conformation of NorM-VC (Table 2.4.1).

Table 2.4.1 Conservation of the cation-binding site in distantly related NorM homologs.

NorM-VC	**hSLC47A1**	**NorM-XC**	**ALF5**	**Identity, %**	**Similarity,%**
E255	E273	E261	E278	100	-
F259	Y277	F265	F282	75	100
S285	I303	Q291	A310	-	50
F288	Y306	F294	Y313	50	100
Y367	S385	F373	S392	-	50
D371	E389	D377	D396	75	100
Y398	Y416	Y404	F423	75	100
F429	I442	L430	L449	50	75
L433	V447	L434	I453	50	100

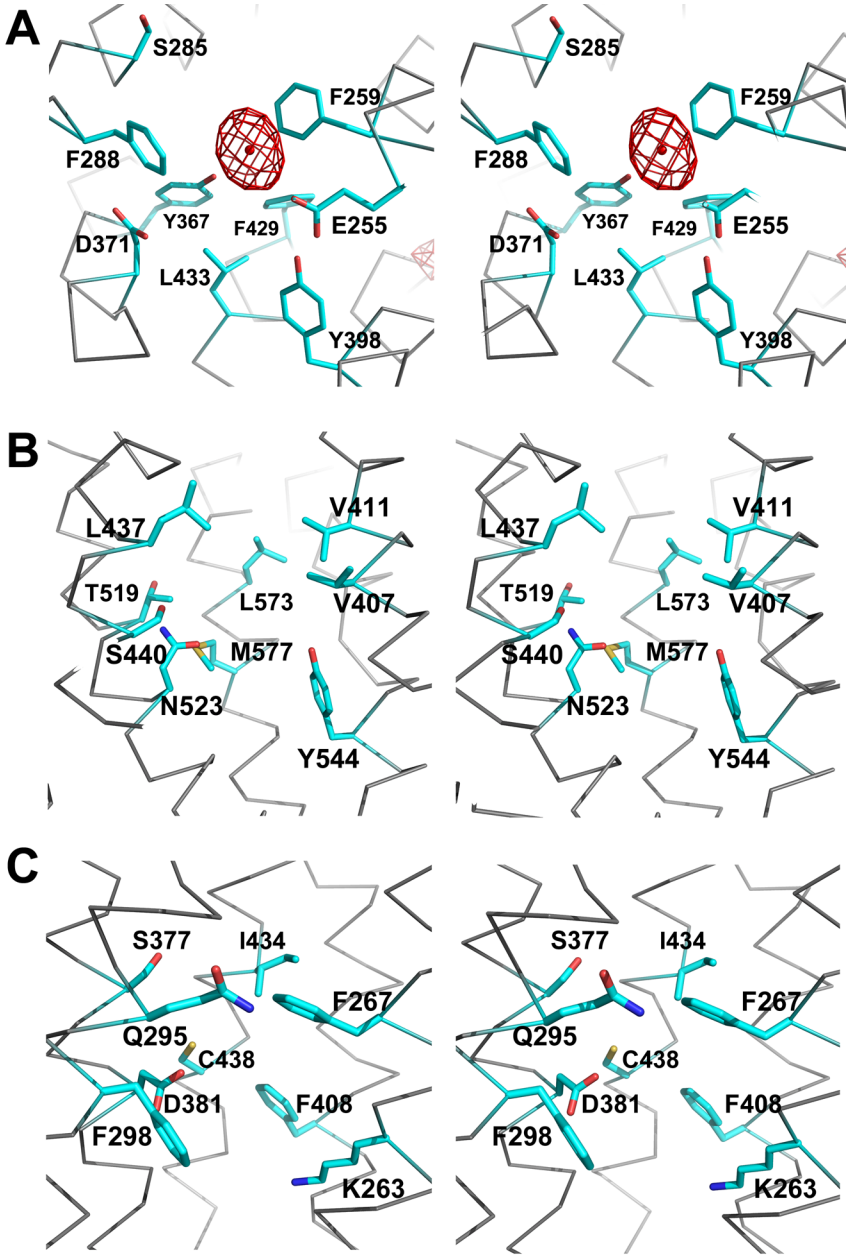


Figure 2.4.3 The corresponding location of the cationg-binding site in the sbMATE and Nt-JAT1 homolog models. (A) The cation-binding site of NorM-VC. The corresponding part of the NorM-VC cation-binding site in (B) sbMATE and (C) Nt-JAT1, respectively. Residues corresponding to the cation-binding site residues of NorM-VC are colored in cyan.

To further investigate whether such a cation binding site is likely to be present in other plant MATE transporters which have more specific transport substrates, the homolog models of a sorghum homolog known to transport primary metabolite citrate and a tabaco plant homolog Nt-JAT1 transporting secondary metabolite nicotine were examined. The sbMATE protein from *Sorghum bicolor* is a final effecter in the Al^{3+} tolerance response, through exporting primary metabolite citric acid into the extracellular space to eliminate further Al^{3+} toxicity[19]. The similar mechanisms in other domesticated plants together affect the crop production up to 50% worldwide[75]. In the sbMATE outward-facing homolog model, the corresponding locus (Figure 2.4.3 B) of the cation-binding site in NorM-VC has slightly more different composition than those of the multidrug MATE homologs (Figure 2.4.1 and Table 2.4.1).

In tobacco plant *Nicotiana tabacum*, Nt-JAT1 is responsible for the vacuole transport and accumulation of nicotine [26]. The homolog model of Nt-JAT1 was then made by threading the homologous regions of Nt-JAT1 onto the homologous regions of the NorM-VC outward-facing conformation structure as previously. The corresponding part of the NorM-VC cation-binding site in the Nt-JAT1 (Figure 2.4.3 C) homolog model is also more dissimilar than those of the multidrug MATE homologs (Figure

2.4.1 and Table 2.4.1). The biochemical features of this locus appear to be slightly different from the cation-binding site of NorM-VC. The NorM-VC D371 equivalent residue D381 is maintained in Nt-JAT1 to potentially donate and receive proton during transport cycle. However, the corresponding Nt-JAT1 locus contains positively charged residue and is slightly more polar than NorM-VC's cation-binding site.

Thus, it is likely that the plant MATEs with more stringent substrate specificity might have different cation coupling/transport mechanisms than the multidrug MATE transporters. One possibility is that the cation binding is needed for the cooperatively binding of the specific substrate in these MATEs, like the Neurotransmitter:Sodium Symporter (NSS) family transporter LeuT[76]. In this scenario, the protonation at this site is needed for an effective substrate binding in the adjacent region within the internal cavity. An alternative explanation for the biochemically dissimilar yet structurally equivalent loci of the NorM-VC cation-binding in these highly specific MATE transporters would be they evolve a completely different substrate and cation coupling mechanism. In this case, the cation-binding site is not homologous to that of the multidrug MATE.

On the other hand, the study on four other distantly related multidrug type of NorM-VC homologs indicates a relatively conserved tertiary structure of the cation-binding site exits in all these transports in an outward-facing conformation. In particular, the corresponding residues (E389 in hMATE1 and mMATE1, D377 in NorM-XC, and D396 in ALF5) of the NorM-VC D371 are highly conserved among these homology modles (Figure 2.4.1 and Table 2.4.1). As shown in NorM-VC structure (Figure 2.3.1) and the previous functional study on the closely related ortholog NorM-VP[42], residueD371 and its equivalent are critical for cation binding/transport in NorM orthologs. Therefore, the cation-binding property is likely to be the feature of the multidrug MATE transporters in clade I and II. The similar biochemical features of the corresponding cation-binding sites in the homology model indicate the cation-binding in the outward-facing conformation is likely a critical step in the transport cycle of the multidrug NorM homologs.

2.5 The Structure-based Transport Mechanism of the MDR MATE NorM-VC

The outward-facing conformation of NorM-VC, like other efflux pumps[71], represents a state in the transport cycle where the transporter has higher affinity for monovalent cations (Figure 2.5.1 step 1), and lower

affinity for substrates. Although the internal cavity is accessible to mercury compounds (Figures 1.7.3), this conformation is not amenable to co-crystallization or soaking with several well-established drugs or transport substrates [11]. Upon cation binding, it is proposed that NorM-VC undergoes an outward- to inward-facing conformational change more favorable to substrate binding (Figure 2.5.1 step 2). Herein, cation is released into the cytoplasm, and the binding of substrate induces a structural change back to the outward-facing conformation[71], where the substrate is released into the outer leaflet of the lipid bilayer and/or extracellular space (Figure 2.5.1 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[7,59,61,68]. 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 (Figure 2.5.1) 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.

Chapter 2, in part, has been accepted for publication of the material as it may appear in the journal Nature, 2010, He, Xiao; Szewczyk, Paul; Karyanki, Andrey; Evin, Riah; Hong, Wen-Xue; Zhang, Qinghai; and Chang, Geoffrey, 2010. The dissertation author is the primary investigator and author of this paper.

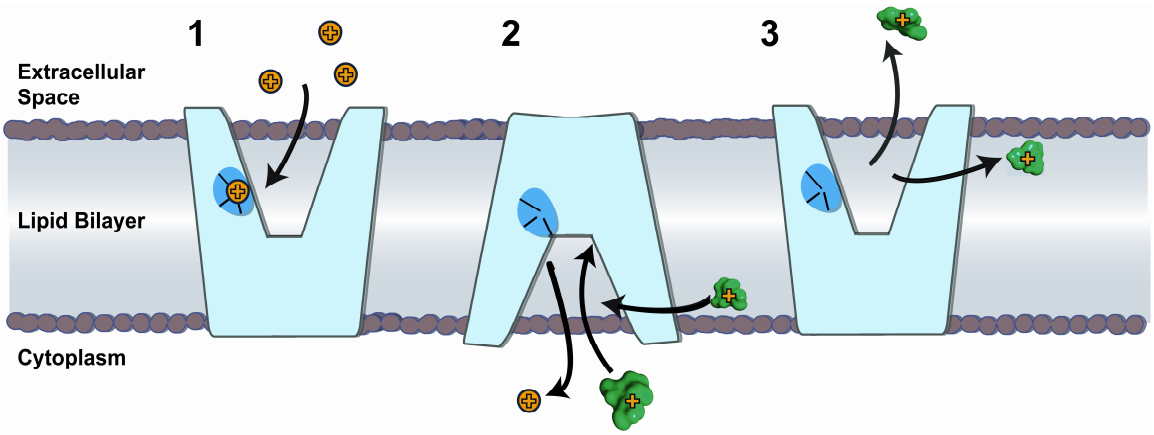


Figure 2.5.1 Transport mechanism of NorM-VC: In the substrate-free outward-facing conformation, cation (yellow) binds at a structurally and functionally conserved site (blue oval; step 1). Cation binding induces structural changes to the inward-facing conformation (step 2), which is competent to bind cationic 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.

3

Similarities and Differences between the Electrochemical Gradient-driven Transporters in the MATE and MFS Families

—A Comparative Approach to the Understanding of Transport Functions and Mechanisms

As one of the most abundant classes of SLC species in animals[8], the functions and structures of MFS has been studied extensively for decades. The structural arrangement and transport substrates across the lipid membrane through a mechanism known as the Rocker-switch model[8]. Compared to MFS, studies of MATE transporters function are more recent[1]. Nevertheless, certain similarities between the MFS and MATE families (similar mass and hydropathy, a pseudo 2-fold symmetry between the N- and C-terminal halves, and shared substrate spectrum in some case) have suggested that we may be able to learn more about the structural basis and functionality of these transporters by comparing the known features of the two families.

3.1 Topological Differences between MATE and MFS

Most MATE and MFS transporters have twelve predicted TM segments based on hydropathy analysis. A direct comparison between the apo-structure of a MDR MFS transporter EmrD from *E. coli* (EC)[7] and that of the MDR MATE transporter NorM-VC clearly demonstrate the two families of transporters have distinct protein folds and likely arose independently.

The apo-structure of a MDR MFS EmrD-EC at 3.5 Å of resolution revealed an intermediate conformation between inward-facing and outward- facing conformations[7]. In this intermediate conformation between inward- and outward-facing, a highly hydrophobic internal cavity tailored for amphipathic substrates such as uncoupler CCCP was observe. A putative selectivity filter composed of two intracellular loops and the inner leaflet side of the immediate TM segments (4-5 and 10-11) is reminiscent of the selectivity filter of certain ion channels. This selectivity filter could be a sieve through the inner leaflet side of the lipid bilayer for amphipathic substates diffusing into the membrane[7]. Residues in the homologous regions in other closely related MDR MFS transporters to EmrD, like LmrP[77], FloR[78], and MdfA[79,80], play important roles regarding substrate specificity.

A direct comparison between the outward-facing conformation of MDR MATE NorM-VC and the intermediate conformation of MDR MFS EmrD-EC demonstrates that the major topological differences between the two families lie in the local interaction between TM helices 1, 4, and 5 in the N-terminal half, and that between TM helices 7, 10, and 11 (Figure 3.1.1). In the MFS structure, TM helix 1 intersects the consecutive helical bundle of TM4, 5, and 6 in the N-ter half; likewise, TM helix 7 does so to the consecutive helical bundle of TM 10, 11, and 12 in the C-ter half. In contrast, there is no such interaction between any helix and a consecutive helical bundle in the MATE fold. These topological differences dictate the differential compositions of the internal cavities and inter-helical interactions during conformational changes in MATE and MFS families.

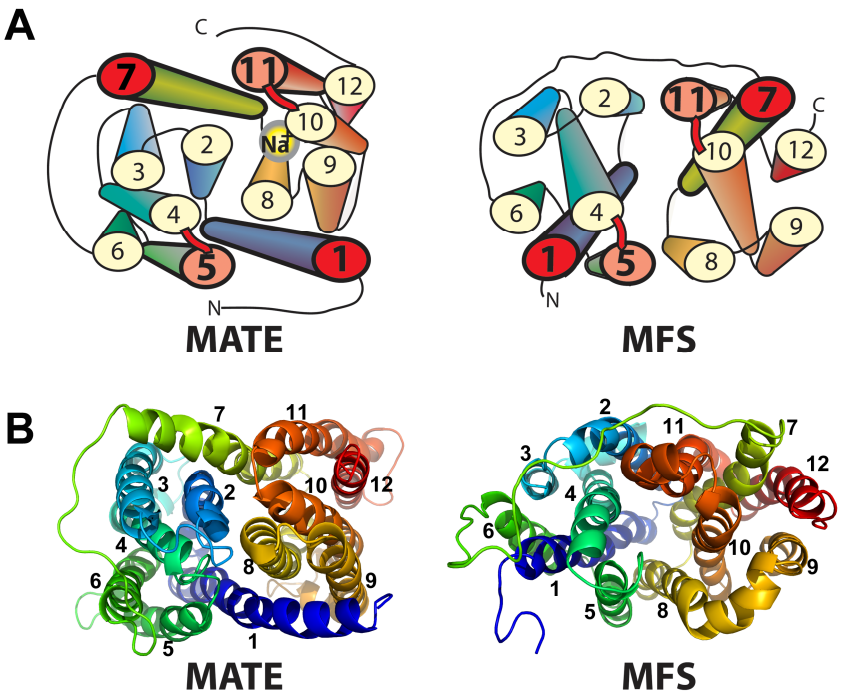


Figure 3.1.1 Topology comparison between the MATE (left) and MFS (right) families of transporters. (A) Topological representation of the MATE (left) and the MFS (right) transporters from the cytoplasmic side. The position of the cation (Na^+) observed in our MATE structure is indicated. In MATEs, TM1 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. (B) X-ray structures of NorM, and EmrD[7], as an example of the MFS. The C-terminal helix (residues 450-461) is removed for clarity.

3.2 The Functional Implications of the Structural Differences

3.2.1 The Coupling Mechanism of Substrate and Cation Transport in MDR

MATE and MDR MFS Transporters likely arose independently

Previous structural works on electrochemical gradient-driven sugar importers in the MFS suggested a Rocker-switch model for substrate transport [8,81,82]. In this transport model, the binding/transport of

electrochemical gradient causes the transporter to change from outward to inward-facing conformations, where the substrate-binding site has alternative access to the two sides of the membrane, respectively[8]. In principle, the conformational changes between outward and inward-facing states also result in altered substrate affinity at the substrate-binding site. In antiporters' case, the inward-facing conformation has lower affinity, while the outward-facing conformation higher affinity for the substrate being exported. As the transport cycle continues, the net effect is that the substrates are imported into the cell as an expense of the dissipation of the electrochemical gradient. The binding/transport of cation is thought to reset the low substrate-affinity outward-facing conformation to the high substrate-affinity inward-facing state in each transport cycle.

3.2.2 Transport Mechanism of MATE Family Is Likely Similar To the Rocker-switch Model of the MFS

The outward-facing conformation of the multidrug MATE transporter NorM-VC represents the substrate-releasing state in the transport cycle and hence has lower affinity for substrates. Yet, residues in the polyspecific substrate-binding pocket in this conformation are among the most conserved residues in NorM homologs (Figure 2.1.6). This observation suggests that residues composing the substrate-binding pocket are largely

to be a similar set of residue in the outward- and inward-facing conformations. The differential affinities for substrate at different conformations are likely the result of spatial rearrangement of the N- and C-terminal halves. In this scenario, helical rotations likely play a lesser role in altering the affinity of the transporter at different conformations/transport stages than the rigid-body movements between the two halves. Hence, a transport mechanism resembling the Rocker-switch model of the MFS family[8] might also be the case for the MATE family, albeit different substrate and cation binding/coupling mechanisms might exist independently.

3.3 Synergistic Actions of MATE and MFS Transporters in Kidney and Liver Is Critical for Mammalian Detoxification Process

In mammals, the MATE homologs MATE1 and MATE2 and a set of multidrug MFS, the organic cation transporter 1, 2, and 2 (OCT1-3), belonging to the SLC22 class (SLC22A1-3) are expressed differentially at the two opposing sides of the luminal epithelial cells in liver and kidneys. These two families of multidrug transporters function synergistically to direct the vectorial movement of metabolic wastes and xenobiotics from the plasma into the lumens for their final deposition[12]. In human, for instance, hMATE1 and 2 are primarily expressed in the brush border side of

the proximal tubule cells, while the OCT2 the basolateral side [12]. Human MATE1 is also expressed in the canaliculus side of the hepatocytes in the liver and OCT1 the sinusoidal side[83]. The mammalian OCTs use the pmf across the basolateral (kidney)/sinusoidal (liver) membrane to drive the uptake of the substrates[83], while the MATEs on the apical membrane efflux the same set of OCs in a eletroneutral manner[84]. A set of organic cations (OC) are the well known substrates for both MATE1-2 and OCT1-3. These OCs include anti-inflammatory drug cimetidine, anti-arrhythmic drug procainamide, anit-diabetic drug metformin, anti-cancer drugs in the cisplatin class, and prototypical cations such as tetraethylammonium (TEA), 1-methyl-4-phenylpyridinium (MPP), and N^1-methylnicotinamide (NMN) [9]. The clearance for these OC species from the plasma to urine/bile relies on the uptake activity of OCT1-3 at the basolateral/sinusoidal side and efflux activity of MATE1-2 at the apical side in the renal proximal tubule cells and hepatocytes, respectively.

Although the multidrug MFS OCTs and MATEs have opposite directions of substrate transport and likely dissimilar mechanisms of cation coupling due to the structure differences between the two families discussed above, their share spectrum of substrates is the very key of detoxification for multi-cellular organisms such as mammals. In rare cases,

divergent substrate specificities between the two families of multidrug transporters have been the cause of toxic compound-induce renal failure in cancer patients[9]. Smaller analogs in the cisplatin class of anti-cancer drugs such (SP-4-2)-diamminedichloridoplatinum retains in renal tubule cells and causes nephrotoxicity because it cannot be transported by mammalian MATE1 on the brush border side whilst still a substrate for OCT2 on the basolateral side[9].

Chapter 3, along with chapter 2 in part, is currently being prepared for submission for publication of the material by He, Xiao; Chang, Geoffrey; and Pogliano, Kit. The dissertation author is the primary investigator and author of this material.

Conclusion

The MATE family is a prominent SLC group that governs the transport of a large collection of nutrients/metabolites in plants and multidrug efflux in eubacteria and eukaryotes. The apo- and cation-bound X-ray structures of a prototypical multidrug MATE transporter NorM-VC revealed a protein fold unique to the MATE family and a structurally discrete cation-binding site within the internal cavity of the transporter for the first time. The single monovalent cation-binding site was found important for cation binding and transport function in NorM orthologs. Homolog modeling of distantly related MATE homologs suggests the cation-binding site is likely to be a structurally and functionally conserved tertiary element that is crucial for multidrug MATE function. The MATE transporter family likely achieves the substrate transport in a similar fashion to the Rocker-switch model of the MFS family, albeit very different coupling mechanisms could exist for cation and substrate binding/transport between the two families. Further understanding of the transport mechanisms of the MATE family awaits future studies that combine functional characteristics with structural attributes of these transporters.

References

1 Brown, M. H., Paulsen, I. T. & Skurray, R. A. The multidrug efflux protein NorM is a prototype of a new family of transporters. *Mol. Microbiol.* **31**, 394-395 (1999).

2 Saier, M. H., Jr, Tran, C. V. & Barabote, R. D. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucl. Acids Res.* **34**, D181-186 (2006).

3 Higgins, C. F. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* **446**, 749-757 (2007).

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

5 Lomovskaya, O., Zgurskaya, H. I., Totrov, M. & Watkins, W. J. Waltzing transporters and 'the dance macabre' between humans and bacteria. *Nature Reviews Drug Discovery* **6**, 56-65 (2007).

6 Chen, Y.-J. *et al.* X-ray structure of EmrE supports dual topology model. *Proceedings of the National Academy of Sciences* **104**, 18999-19004 (2007).

7 He, X., Yin, Y., Szewczyk, P., Ngyuen, T. & Chang, C. Structure of the multidrug transporter EmrD from Escherichia coli (vol 312, pg 741, 2006). *Science* **317**, 1682-1682 (2006).

8 Law, C. J., Maloney, P. C. & Wang, D.-N. Ins and Outs of Major Facilitator Superfamily Antiporters. *Annual Review of Microbiology* **62**, 289-305 (2008).

9 Terada, T. & Inui, K. I. Physiological and pharmacokinetic roles of H+/organic cation antiporters (MATE/SLC47A). *Biochem. Pharmacol.* **75**, 1689-1696 (2008).

10 Morita, Y. *et al.* NorM, a Putative Multidrug Efflux Protein, of Vibrio parahaemolyticus and Its Homolog in Escherichia coli. *Antimicrob. Agents Chemother.* **42**, 1778-1782 (1998).

11 Kuroda, T. & Tsuchiya, T. Multidrug efflux transporters in the MATE family. *BBA-Proteins Proteomics* **1794**, 763-768 (2009).

12 Moriyama, Y., Hiasa, M., Matsumoto, T. & Omote, H. Multidrug and toxic compound extrusion (MATE)-type proteins as anchor transporters for the excretion of metabolic waste products and xenobiotics. *Xenobiotica* **38**, 1107-1118 (2008).

13 Xu, X. J. *et al.* Molecular cloning and characterization of the HmrM multidrug efflux pump from Haemophilus influenzae Rd. *Microbiology and Immunology* **47**, 937-943 (2003).

14 Huda, M. N., Morita, Y., Kuroda, T., Mizushima, T. & Tsuchiya, T. Na+-driven multidrug efflux pump VcmA from Vibrio cholerae non-O1, a non-halophilic bacterium. *Fems Microbiology Letters* **203**, 235-239 (2001).

15 Burse, A., Weingart, H. & Ullrich, M. S. NorM, an Erwinia amylovora multidrug efflux pump involved in in vitro competition with other epiphytic bacteria. *Applied and Environmental Microbiology* **70**, 693-703 (2004).

16 Kaatz, G. W., McAleese, F. & Seo, S. M. Multidrug resistance in Staphylococcus aureus due to overexpression of a novel multidrug and toxin extrusion (MATE) transport protein. *Antimicrob Agents Chemother* **49**, 1857-1864 (2005).

17 McAleese, F. *et al.* A novel MATE family efflux pump contributes to the reduced susceptibility of laboratory-derived Staphylococcus aureus mutants to tigecycline. *Antimicrob Agents Chemother* **49**, 1865-1871 (2005).

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

19 Kochian, L. V., Pineros, M. A. & Hoekenga, O. A. The physiology, genetics and molecular biology of plant aluminum resistance and toxicity. *Plant Soil* **274**, 175-195 (2005).

20 Diener, A. C., Gaxiola, R. A. & Fink, G. R. Arabidopsis ALF5, a multidrug efflux transporter gene family member, confers resistance to toxins. *Plant Cell* **13**, 1625-1637 (2001).

21 Hvorup, R. N. *et al.* The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. *Eur J Biochem* **270**, 799-813 (2003).

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

23 Shitan, N. & Yazaki, K. Accumulation and membrane transport of plant alkaloids. *Current Pharmaceutical Biotechnology* **8**, 244-252 (2007).

24 Yazaki, K., Sugiyama, A., Morita, M. & Shitan, N. Secondary transport as an efficient membrane transport mechanism for plant secondary metabolites. *Phytochemistry Reviews* **7**, 513-524 (2008).

25 Marinova, K. *et al.* The Arabidopsis MATE transporter TT12 acts as a vacuolar flavonoid/H+-antiporter active in proanthocyanidin-accumulating cells of the seed coat. *Plant Cell* **19**, 2023-2038 (2007).

26 Morita, M. *et al.* Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in Nicotiana tabacum. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 2447-2452 (2009).

27 Gomez, C. *et al.* Grapevine MATE-Type Proteins Act as Vacuolar H+-Dependent Acylated Anthocyanin Transporters. *Plant Physiol.* **150**, 402-415 (2009).

28 Yokosho, K., Yamaji, N. & Ma, J. F. Isolation and characterisation of two MATE genes in rye. *Funct. Plant Biol.* **37**, 296-303 (2009).

29 Morita, M. *et al.* Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in Nicotiana tabacum. *Proc Natl Acad Sci U S A* **106**, 2447-2452 (2009).

30 Wood, S., Sebastian, K. L. & Scherr, S. J. *Pilot analysis of global ecosystems : agroecosystems*. (World Resources Institute, 2000).

31 Magalhaes, J. V. *et al.* A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nat Genet* **39**, 1156-1161, (2007).

32 Maron, L. G. *et al.* Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. *Plant J.* **61**, 728-740 (2010).

33 Yokosho, K., Yamaji, N., Ueno, D., Mitani, N. & Ma, J. F. OsFRDL1 Is a Citrate Transporter Required for Efficient Translocation of Iron in Rice. *Plant Physiol.* **149**, 297-305 (2009).

34 Furukawa, J. *et al.* An aluminum-activated citrate transporter in barley. *Plant and Cell Physiology* **48**, 1081-1091, (2007).

35 Ryan, P. R., Raman, H., Gupta, S., Horst, W. J. & Delhaize, E. A Second Mechanism for Aluminum Resistance in Wheat Relies on the Constitutive Efflux of Citrate from Roots. *Plant Physiol.* **149**, 340-351 (2009).

36 Durrett, T. P., Gassmann, W. & Rogers, E. E. The FRD3-mediated efflux of citrate into the root vasculature is necessary for efficient iron translocation. *Plant Physiol.* **144**, 197-205 (2007).

37 Li, L. G., He, Z. Y., Pandey, G. K., Tsuchiya, T. & Luan, S. Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification. *Journal of Biological Chemistry* **277**, 5360-5368 (2002).

38 Becker, M. L. *et al.* 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**, 745-749 (2009).

39 Tsuda, M. *et al.* Targeted disruption of the multidrug and toxin extrusion 1 (mate1) gene in mice reduces renal secretion of metformin. *Mol Pharmacol* **75**, 1280-1286 (2009).

40 Chen, Y. *et al.* Genetic variants in multidrug and toxic compound extrusion-1, hMATE1, alter transport function. *Pharmacogenomics Journal* **9**, 127-136 (2009).

41 Nishihara, K., Masuda, S., Ji, L., Katsura, T. & Inui, K. I. Pharmacokinetic significance of luminal multidrug and toxin extrusion 1 in chronic renal failure rats. *Biochem. Pharmacol.* **73**, 1482-1490 (2007).

42 Otsuka, M. *et al.* Identification of essential amino acid residues of the NorM Na+/multidrug antiporter in Vibrio parahaemolyticus. *J. Bacteriol.* **187**, 1552-1558, doi:10.1128/jb.187.1552-1558.2005 (2005).

43 Tsuda, M. *et al.* Oppositely directed H+ gradient functions as a driving force of rat H+/organic cation antiporter MATE1. *Am J Physiol Renal Physiol* **292**, F593-598 (2007).

44 Omote, H., Hiasa, M., Matsumoto, T., Otsuka, M. & Moriyama, Y. The MATE proteins as fundamental transporters of metabolic and

xenobiotic organic cations. *Trends in Pharmacological Sciences* **27**, 587-593 (2006).

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

46 Zhang, Q. H. *et al.* Designing facial amphiphiles for the stabilization of integral membrane proteins. *Angew. Chem.-Int. Edit.* **46**, 7023-7025 (2007).

47 Su, C. C., Long, F., McDermott, G., Shafer, W. M. & Yu, E. W. Crystallization and preliminary X-ray diffraction analysis of the multidrug efflux transporter NorM from Neisseria gonorrhoeae. *Acta Crystallogr. F-Struct. Biol. Cryst. Commun.* **64**, 289-292 (2008).

48 Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* **276**, 307-326 (1997).

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

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

51 Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55**, 849-861, doi:GR0894 (pii) (1999).

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

53 Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-921 (1998).

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

55 Sack, J. S. Chain - a Crystallographic Modeling Program. *J Mol Graphics* **6**, 224-225 (1988).

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

57 DeLano, W. L. (2009).

58 He, X. *et al.* Structure of a Cation-bound Multidrug and Toxin Compound Extrusion (MATE) Transporter. *Nature,* in press (2010).

59 Chen, Y. J. *et al.* X-ray structure of EmrE supports dual topology model. *Proc Natl Acad Sci U S A* **104**, 18999-19004 (2007).

60 Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T. & Yamaguchi, A. Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* **443**, 173-179 (2006).

61 Aller, S. G. *et al.* Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **323**, 1718-1722 (2009).

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

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

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

65 Landau, M. *et al.* ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* **33**, W299-W302 (2005).

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

67 Watkins, R. E. *et al.* The human nuclear xenobiotic receptor PXR: Structural determinants of directed promiscuity. *Science* **292**, 2329-2333 (2001).

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

69 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

70 Altschul, S. F. *et al.* Protein database searches using compositionally adjusted substitution matrices. *FEBS J* **272**, 5101-5109 (2005).

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

72 Tsuda, M. *et al.* Targeted Disruption of the Multidrug and Toxin Extrusion 1 (Mate1) Gene in Mice Reduces Renal Secretion of Metformin. *Mol. Pharmacol.* **75**, 1280-1286 (2009).

73 Becker, M. S. L. *et al.* 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**, 745-749 (2009).

74 Crossman, L. *et al.* The complete genome, comparative and functional analysis of Stenotrophomonas maltophilia reveals an

organism heavily shielded by drug resistance determinants. *Genome Biology* **9**, R74 (2008).

75 S. Wood, K. L. S., S. J. Scherr *Pilot analysis of global ecosystems : agroecosystems.* . Vol. xii 110p (World Resources Institute, Washington, D.C., 2000, 2000).

76 Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature* **437**, 215-223 (2005).

77 Mazurkiewicz, P., Konings, W. N. & Poelarends, G. J. Acidic residues in the lactococcal multidrug efflux pump LmrP play critical roles in transport of lipophilic cationic compounds. *Journal of Biological Chemistry* **277**, 26081-26088 (2002).

78 Braibant, M., Chevalier, J., Chaslus-Dancla, E., Pages, J. M. & Cloeckaert, A. Structural and functional study of the phenicol-specific efflux pump FloR belonging to the major facilitator superfamily. *Antimicrob. Agents Chemother.* **49**, 2965-2971 (2005).

79 Edgar, R. & Bibi, E. A single membrane-embedded negative charge is critical for recognizing positively charged drugs by the Escherichia coli multidrug resistance protein MdfA. *Embo Journal* **18**, 822-832 (1999).

80 Adler, J. & Bibi, E. Promiscuity in the geometry of electrostatic interactions between the Escherichia coli multidrug resistance transporter MdfA and cationic substrates. *Journal of Biological Chemistry* **280**, 2721-2729 (2005).

81 Abramson, J. *et al.* Structure and mechanism of the lactose permease of Escherichia coli. *Science* **301**, 610-615 (2003).

82 Huang, Y. F., Lemieux, M. J., Song, J. M., Auer, M. & Wang, D. N. Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. *Science* **301**, 616-620 (2003).

83 Ciarimboli, G. Organic cation transporters. *Xenobiotica* **38**, 936-971 (2008).

84 Otsuka, M. *et al.* A human transporter protein that mediates the final excretion step for toxic organic cations. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17923-17928 (2005).