Multidrug Transporter MdfA as a Target for High-resolution Structural Studies

A Thesis Submitted to the College of
Graduate Studies and Research
in Fulfillment of the Requirements
for the Degree of Master of Science
in the Department of Biochemistry
University of Saskatchewan
Saskatoon

By
Christopher O'Grady

Copyright Note

Abstract

The MdfA is a 410 amino acid-long integral membrane protein, which belongs to the Major Facilitator superfamily of multidrug transporters. It is predicted to consist of 12 transmembrane helices. MdfA uses the energy of the transmembrane proton gradient to pump a variety of toxic compounds out of *E. coli* cells. No high resolution structure of MdfA is available. The goals of this research project were to develop a practical method for purification of MdfA, to evaluate the feasibility of structure determination by Nuclear Magnetic Resonance (NMR) and X-ray crystallography, and to develop an activity assay for purified MdfA. To this end, MdfA, with a hexa-histidine tag attached to facilitate protein purification, was successfully expressed and incorporated into the cell membrane using an *E. coli* expression system. MdfA was extracted from the cell membrane with the detergents 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine (DHPC), *n*-dodecyl-β-D-maltoside (DDM), and 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)] (LMPG) and purified by affinity chromatography on nickel-nitrilotriacetic acid agarose. Pure protein was found to be monodisperse in DHPC, DDM and LMPG micelles. To achieve simple amino acid selective isotope labeling for high-resolution NMR studies, MdfA was expressed in a cell-free translation system. To determine if the purified protein was properly folded, ^{19}F NMR experiments were carried out on 5-fluoro-tryptophan-labeled MdfA while titrating the MdfA substrates ethidium bromide and chloramphenicol into the fluoro-tryptophan-labeled MdfA sample. An activity assay was developed for MdfA incorporated into liposomes using the fluorescent dye 9-amino-6-chloro-2-methoxyacridine (ACMA) to detect proton translocation coupled to substrate transport. Results from both the ^{19}F NMR and the transport activity assay indicated that the purified MdfA was properly folded and functional. NMR experiments with pure MdfA yielded spectra of insufficient quality for high-resolution structure determination but did indicate that structural studies of MdfA by NMR are feasible. Crystallization trials yielded crystals that are likely to contain protein and will serve as a starting point for further optimization of crystallization conditions for X-ray structure determination.

Acknowledgements

I would like to thank Dr. Oleg Dmitriev for his guidance and endless support throughout my Masters project. I would like to thank everyone in the Dmitriev lab for their support and encouragement including but not limited to Eva Uhlemann, Hannah Pierson, Natasha Dolgova, Igor Moshynskyy, Sergiy Nokhrin, Ben Remple, and Tahereh Haji. I would also like to thank Carla Protsko from the Delbaere lab for conducting some of the crystallization trials and for assisting with the crystallization data analysis. I would like to thank my committee for their guidance and support throughout my project. I would also like to thank all my friends and especially my wife Lara for their support and encouragement.

Dedication

I would like to dedicate this thesis to my wife and my family, all of whom gave me tremendous support throughout my research and thesis writing. Thanks for always acting interested even when you didn't know what I was talking about.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABC	Adenosine triphosphate-binding cassette family
ACMA	9-amino-6-chloro-2-methoxyacridine
AIDS	Acquired immunodeficiency syndrome
cAMP	Cyclic AMP
CBD	Chitin binding domain
CHAPS	3-[3-(Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CMC	Critical micelle concentrations
Cryo-Em	Cryo-electron microscopy
DDM	*n*-dodecyl-β-D-maltoside
DHPC	1,2-Diheptanoyl-*sn*-Glycero-3-Phosphocholine
E. coli	*Escherichia coli*
FCCP	Carbonyl cyanide 4-(trifluoromethoxy)-phenylhydrazone
GST	Glutathione S-transferase
HSQC	Heteronuclear single quantum coherence
LB	Luria Broth
LMPG	1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)]
MATE	Multidrug and toxic compound extrusion family
MFS	Major Facilitator superfamily
Ni-NTA	Nickel nitrilotriacetic acid agarose
RM	Casamino acid medium
RMSD	Root mean square difference
RND	Resistance-nodulation-cell division family
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SMR	Small multidrug resistance family
SSSC	Saskatchewan Structural Sciences Center
TMDG	50 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$ 1 mM dithiothreitol, 10% glycerol (v/v) and 1 mM phenylmethylsulfonyl fluoride
TROSY	Transverse relaxation optimized spectroscopy

1. Introduction

The ability of an organism to withstand the effects of a variety of toxic compounds is known as multidrug resistance. Multidrug resistance can be the result of over-expression of multidrug transporting proteins. Multidrug transporters are a family of membrane proteins that recognize a broad range of substrates and facilitate their movement out of the cell. Multidrug transporters are found in both prokaryotic and eukaryotic organisms.

Multidrug resistance due to over-expression of multidrug transporters is becoming a major problem in modern medicine. Bacterial infections that were once easy to treat with antibiotics are now more resilient due to the emergence of multidrug resistance. *Staphylococcus aureus* is one such bacterium that was previously treatable with penicillin, but is now resistant not only to penicillin, but to many other antibiotics as well, including methicillin and vancomycin (Jevons, 1961; Hiramatsu *et al.,* 1997; Howe *et al.,* 1998; Ploy *et al.,* 1998). *Staphylococcus aureus* is commonly found on the skin, but if internalized through skin abrasions, it can cause skin boils, pressure sores, blood poisoning (septicemia), pneumonia, or bone, joint, and heart valve infections. In 2005 in the United States, 477,927 people were hospitalized with *Staphylococcus aureus* infections and there were 11,406 *Staphylococcus aureus*-related deaths (Klein *et al.*, 2007). NorA is an example of a multidrug transporter found in *Staphylococcus aureus* which confers resistance to the quinolones norfloxacin, enoxacin, ofloxacin, ciprofloxacin, nalidixic acid, oxolinic acid, and sparfloxacin (Yoshida *et al.,* 1990).

Mycobacterium tuberculosis is another example of a bacterium that has developed resistance to several different antibiotics. Drug-resistant tuberculosis is one of the most deadly and common infectious diseases, claiming millions of lives per year worldwide (World Health Organization Report, 2008). Drug-resistance in tuberculosis can be acquired by several mechanisms, one being drug extrusion by multidrug transporters. DrrAB is an example of a multidrug transporter expressed in *Mycobacterium tuberculosis* that has been found to confer resistance to a broad range of clinically relevant antibiotics including tetracycline, erythromycin, ethambutol, norfloxacin, streptomycin and chloramphenicol (Choudhuri *et al.,* 2002). These are just two examples of bacterial infections that, due to multidrug resistance, are no longer as easily treatable using antibiotics as they once were.

Human diseases such as cancer, epilepsy, and acquired immunodeficiency syndrome (AIDS) are becoming very hard to treat due to the over-expression of the human multidrug transporter P-glycoprotein. P-glycoprotein binds and expels many chemotherapeutic drugs used in the treatment of cancers and AIDS (Juliano and Ung, 1976; Chen *et al*., 1986; Ueda *et al*., 1987; Endicott and Ling, 1989; Higgins, 1992; Gottesman and Pastan, 1993; Lee *et al*., 1998; Krishna and Mayer, 2000; Gottesman *et al*., 2002), as well as antiepileptic drugs (Loscher and Potschka, 2002; Siddiqui, 2003). Not only do multidrug transporters interfere with the treatment of bacterial infections, but also with the treatment of non-bacterial diseases.

A simple experimental model is required to better understand how multidrug transport proteins work. For this study, we chose the multidrug transport protein MdfA, which is found in *Escherichia coli* (*E. coli*). The ultimate goal of this research project is to elucidate the molecular mechanism of drug extrusion catalyzed by MdfA and similar transport proteins. To this end, we expressed and purified MdfA for structural studies and evaluated the feasibility of solving the three-dimensional structure of MdfA by Nuclear Magnetic Resonance (NMR) and X-ray crystallography. A high-resolution structure of MdfA could aid the development of novel means of antimicrobial therapy, and overcome drug resistance in bacterial infections and cancer.

2. Background and State of the Problem

2.1. Mechanisms of Antibiotic Resistance

The four main mechanisms by which bacteria can acquire drug resistance are: drug inactivation; alteration of drug target; alteration of the targeted metabolic pathway; and reduced drug accumulation (Tenover, 2006).

The first mechanism involves chemical modification of the drug molecule by an enzymatic reaction resulting in drug inactivation. An example of this mechanism is inactivation of penicillins by β-lactamase. Penicillins achieve antibiotic activity through their interactions with transpeptidases, also known as penicillin-binding proteins (Wise and Park, 1965). Transpeptidases are responsible for the formation of cross-links in the peptidoglycan layer. The lactam ring of penicillins binds the active sites of transpeptidases, inactivating the enzyme and preventing cross-linking of the peptidoglycan, which disrupts synthesis of the peptidoglycan layer. First discovered in 1940 (Abraham and Chain, 1940), β-lactamase hydrolyzes the lactam ring of penicillins. Hydrolysis of the lactam ring abrogates its antimicrobial activity.

The second mechanism is the alteration of the antibiotic target site in such a way that the drug no longer interacts with the target, thereby losing its effectiveness. An example of this mechanism is found in *Staphylococcus pneumoniae* which developed altered forms of penicillin-binding proteins with decreased affinity for penicillins (Laible and Hakenbeck, 1991). This allows transpeptidation of the peptidoglycan layer to proceed even in the presence of penicillin.

The third mechanism occurs when bacteria develop an alternative metabolic pathway that bypasses the target protein of the drug. *E. coli* infections, as well as other infections caused by Gram-negative bacteria, have been found to exhibit resistance to trimethoprim by bypassing the folic acid metabolic pathway (King *et al.*, 1983). Trimethoprim achieves its antibiotic activity by competitively inhibiting dihydrofolate reductase, an enzyme, which is critical for the biosynthesis of folic acid. *E. coli* has been found to be resistant to trimethoprim by becoming dependent on an external supply of thymine, which is normally synthesized from folic acid (King *et al.*, 1983).

The last mechanism, reduced drug accumulation, is achieved by either limiting drug entrance into the cell, or by actively pumping the drug out of the cell. As an example, tetracycline resistance in *E. coli* is conferred by actively pumping the drug out of the cell. Tetracycline achieves antibiotic activity through the binding of 16S rRNA which prevents the

attachment of aminoacyl-tRNA to the ribosomal acceptor site, thus inhibiting protein synthesis. TetA is a metal-tetracycline/proton antiporter, which actively pumps tetracyclines out of the cell leading to *E.coli* resistance to tetracycline (Yamaguchi *et al.*, 1990). All multidrug transporters use this fourth mechanism to achieve antibiotic resistance.

2.2. Bacterial Multidrug Transporters

Multidrug transporters are a family of membrane proteins found in many different organisms that recognize a broad range of substrates and facilitate their subsequent movement across the cell membrane. Multidrug transporters are divided into five families: adenosine triphosphate (ATP)-binding cassette (ABC) superfamily (Fath and Kolter 1993; Higgins, 2001), multidrug and toxic compound extrusion (MATE) family (Brown *et al.*, 1999), small multidrug resistance (SMR) family (Paulsen *et al.*, 1996b), resistance-nodulation-cell division (RND) family (Saier *et al.*, 1994), and major facilitator superfamily (MFS) (Pao *et al.*, 1998) (Figure 2.1).

All of these multidrug transporters can be divided into two main groups, primary and secondary active transporters, depending on the driving force of the substrate transport. Primary active transporters use the energy released by ATP hydrolysis to transport their substrates, whereas secondary active transporters use electrochemical ion gradients to drive their substrate transport. The ABC transporters is the only family that is classified as a primary active transporter type, whereas the other families, MFS, RND, MATE, and SMR, are classified as secondary active transporters.

The ABC transporters have been found to transport sugars, amino acids, ions, drugs, iron complexes, polysaccharides and proteins across the cell membrane (Higgins, 1992; Fath and Kolter, 1993; Saier *et al.*, 1999). They are the largest multidrug transporters due to their bulky ATP binding domains. The large ATP binding domain can be observed in the X-ray structure of the ABC transporter Sav1866 from *Staphylococcus aureus* in Figure 2.2. This protein functions as a homodimer where each monomer consists of 6 transmembrane helices and an ATP binding domains. The ATP binding domain is a very large and complex portion of the ABC transporters, which in turn can complicate structural studies in both NMR and X-ray crystallography. Therefore, a smaller secondary active transporter lacking the ATP binding domain, such as MdfA, would be a simpler model protein for structural studies.

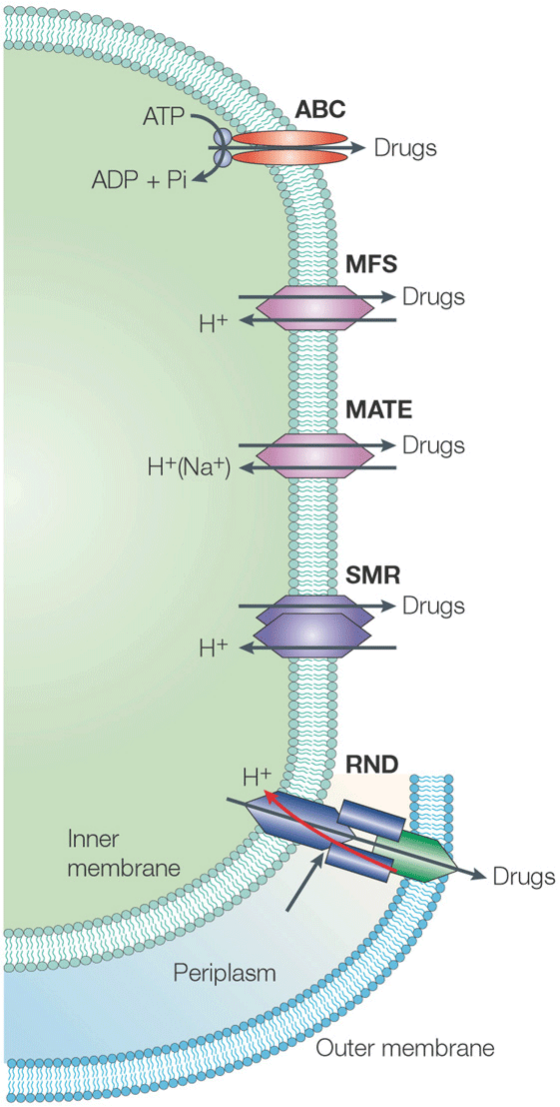


Figure 2.1. The five multidrug transporter families. Reprinted with permission from Macmillan Publishers Ltd: (Nature) (Krulwich *et al.*, 2005).

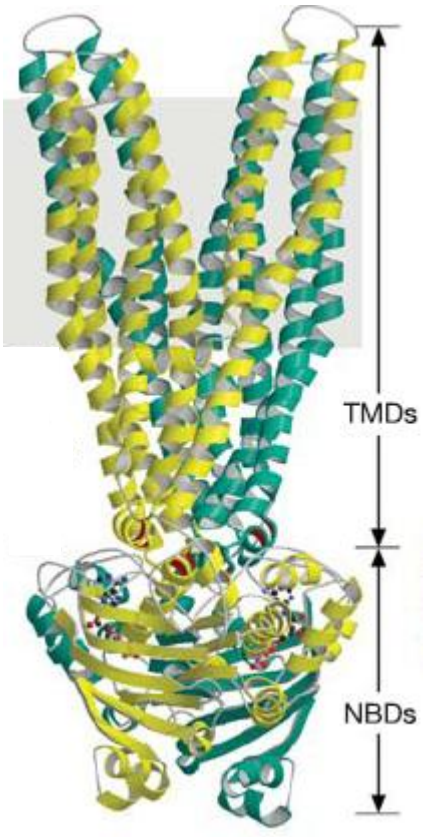


Figure 2.2. Side view of the three dimensional structure of the ABC transporter Sav1866 with the subunits colored yellow and turquoise. TMDs: transmembrane domains; NBDs: nucleotide-binding domains. Reprinted with permission from Macmillan Publishers Ltd: (Nature) (Dawson and Locher, 2006) (PDB ID 2ONJ).

The members of the MATE family of multidrug transporters have been found to mediate resistance to cationic dyes, aminoglycosides, and fluoroquinolones. Unlike the ABC transporters, MATE transporters use proton and sodium electrochemical gradients as driving forces of transport (Morita *et al.*, 1998, 2000). MATE family members are similar to MFS family members because they contain 12 transmembrane helices, but they do not exhibit sequence similarity with any members of the MFS (Morita *et al.*, 1998).

The transporters of the SMR protein family also pump drugs out of the cell in exchange for protons. SMR proteins are much smaller than other multidrug transporters, about 110 amino acid residues in length, and typically consist of four transmembrane helices. They also have a smaller substrate range, usually limited to lipophilic cations (Grinius *et al.*, 1992; Schuldiner *et al.*, 1997). EmrE, a member of the SMR family, is an asymmetric dimer where each monomer contains four transmembrane helices (Figure 2.3). The EmrE dimer

confers resistance to toxic poly-aromatic cations. The structure for EmrE purified using the detergent dodecylmaltoside was originally solved using cryo-electron microscopy (cryo-EM) (Figure 2.3, A). A subsequent X-ray structure was found to be similar to the cryo-EM structure (Figure 2.3, B), having a pairwise root mean square difference (RMSD) of 1.4 Ångstrom over the equivalent C^{α} atoms (Chen *et al*., 2007).

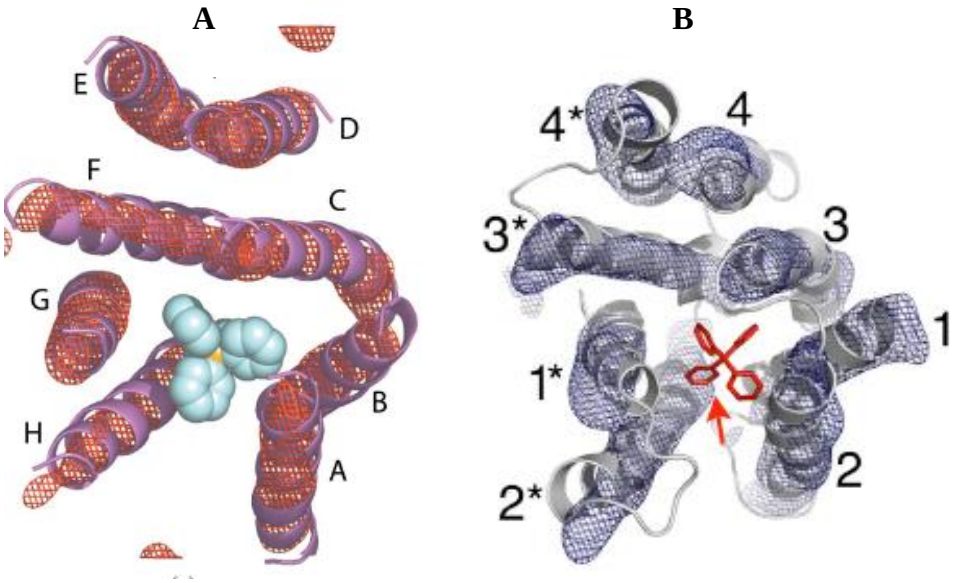


Figure 2.3. Top down view of the three dimensional structure of EmrE solved by cryo-electron microscopy (**A**) (Tate, 2006) (PDB ID 2I68) and X-ray crystallography (**B**) (Chen *et al*., 2007) (PDB ID 3B5D). Tetraphenylphosphonium is represented by a space-filling model in A and is pointed to in B (red arrow). Reprinted with permission from Elsevier.

The RND family members usually form complexes with membrane fusion proteins and outer membrane components to allow drug transport across both the inner and outer membranes of gram-negative bacteria. The RND-type efflux proteins are mostly drug/proton antiporters consisting of 12 transmembrane helices (Saier *et al.*, 1994). They recognize a variety of clinically relevant antimicrobials such as carbenicillin, thiolactomycin, tetracycline, and chloramphenicol (Nikaido, 1998). The *E. coli* AcrB transporter functions as a homotrimer, wherein each subunit contains 12 transmembrane helices and two large periplasmic domains as shown in Figure 2.4. The two large periplasmic domains are a pore domain and what is believed to be a TolC docking domain. AcrB is hypothesized to be in a multi-protein complex with the outer membrane channel protein TolC and the periplasmic linker protein AcrA (Figure 2.5) (Murakami *et al.*, 2002). This structure allows for the direct export of drugs from the cell interior to the external medium.

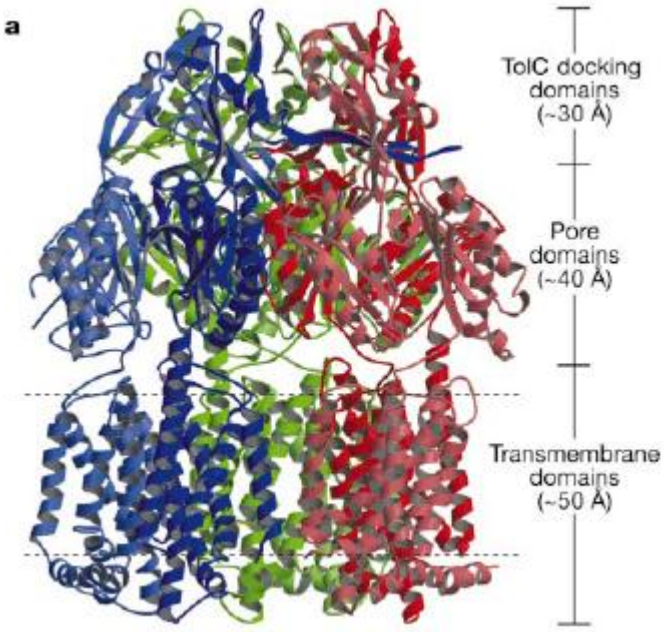


Figure 2.4. Side view of the three dimensional structure of AcrB. Red, green and blue represent the three individual protomers. Reprinted with permission from Macmillan Publishers Ltd: (Nature) (Murakami *et al.*, 2002) (PDB ID 1IWG).

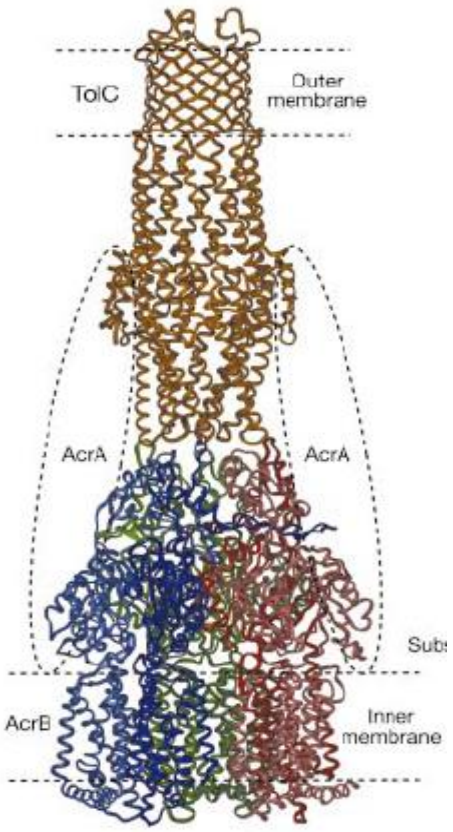


Figure 2.5. Side view of the proposed model for the AcrB-AcrA-TolC complex. Reprinted with permission from Macmillan Publishers Ltd: (Nature) (Murakami *et al.,* 2002) (PDB ID 1IWG).

The MFS transporters facilitate the transfer of sugars, anions, metabolites, and drug molecules across the cell membrane, driven by ion electrochemical gradients, typically proton-motive force (Marger and Saier, 1993; Pao *et al.,* 1998; Saier *et al.,* 1999). The Major Facilitator Superfamily consists of sugar uptake facilitators (Maiden, *et al.,* 1987; Henderson and Maiden, 1990), multidrug transporters such as EmrD of *E. coli* (Naroditskaya, 1993), single drug transporters such as Tet (L) of *Bacillus subtilis* (Krulwich *et al.,* 2001), Krebs cycle intermediate transport facilitators (Griffith *et al.,* 1992; Paulsen and Skurray, 1994), organophosphate:phosphate exchangers, and oligosaccharide:H^+ symporters (Marger and Saier, 1993). The MFS transporters can be further classified into two families based on transmembrane topology. While MFS transporters with 14 transmembrane helices are classified under the DHA14 family, MFS transporters with 12 transmembrane helices are classified under the DHA12 family (Paulsen *et al.,* 1996a; Pao *et al.,* 1998).

GlpT and LacY are two DHA12 MFS transporters from *E. coli* whose three dimensional structures have been solved by X-ray crystallography to 3.3 and 3.5 Ångstrom resolution respectively (Abramson *et al.*, 2003; Huang *et al.*, 2003). GlpT facilitates the transport of glycerol-3-phosphate in exchange for organic phosphate while LacY facilitates the transport of lactose using the proton gradient. A central cavity opening into the cytosol is seen both in GlpT and LacY when the substrate is not present (Figure 2.6 and 2.7). LacY has been studied in depth, which resulted in the determination of the amino acids involved in substrate binding (Glu^{126}, Arg^{144}, Trp^{151} and Glu^{269}), as well as in proton translocation (Tyr^{236}, Glu^{269}, Arg^{302}, His^{322} and Glu^{325}) (Abramson *et al.*, 2003). The substrate binding site for GlpT is believed to be deep in the internal cavity and to contain two positively charged residues, Arg^{45} and Arg^{269} (Huang *et al.*, 2003). These two MFS transporters share similar overall structure, having an RMSD of 3.7 Ångstrom (Vardy *et al.*, 2004).

In order to treat patients with multidrug resistant bacterial infections that express one or several multidrug transporters, it is critical that multidrug transporters are studied in detail. Proteins such as MdfA, which belongs to the MFS transporters, make good candidates for structural studies.

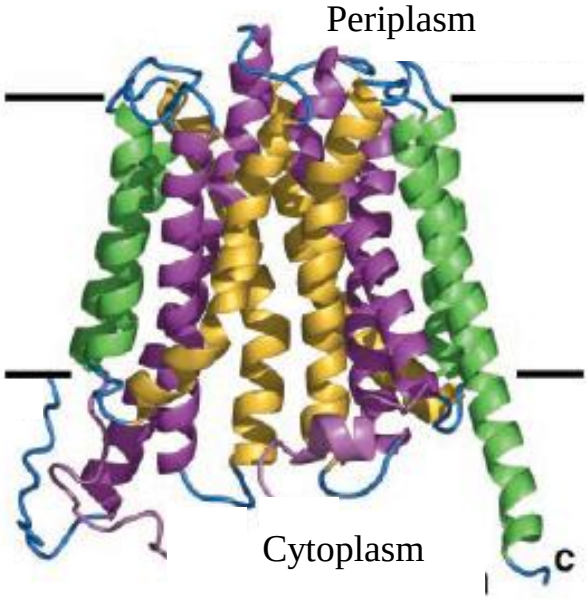


Figure 2.6. Side view of the ribbon representation of GlpT (Huang *et al.*, 2003) (PDB ID 1PW4). Reprinted with permission from AAAS.

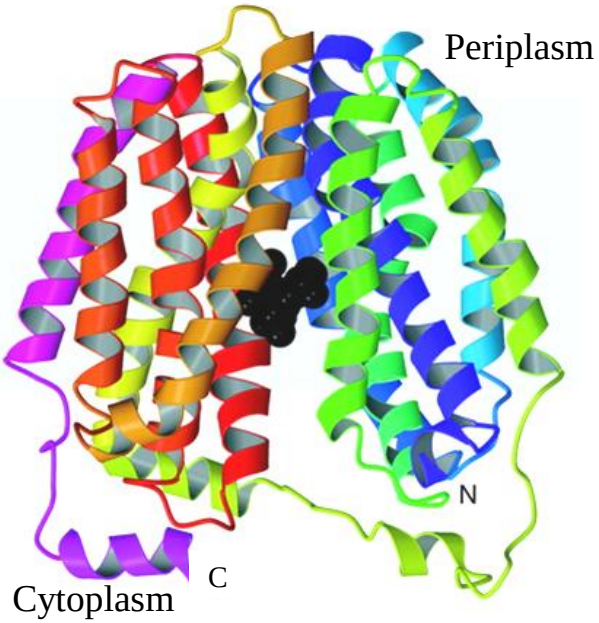


Figure 2.7. Side view of the ribbon representation of LacY (Abramson *et al.*, 2003) (PDB ID 1PV6). β-D-galactopyranosyl-1-thiol-β-D-galactopyranoside is represented by black spheres. Reprinted with permission from AAAS.

2.3. MdfA, an E. coli Multidrug Transporter with Extremely Broad Substrate Specificity

MdfA is an *E. coli* multidrug resistance transporter consisting of 410 amino acids (Edgar and Bibi, 1997). This membrane protein belongs to the Major Facilitator Superfamily of secondary transporters (Saier and Paulsen, 2001). Recent studies have found that MdfA operates as a monomer (Sigal *et al*., 2007). MdfA facilitates the movement of substrates out of the cell in exchange for protons (Edgar and Bibi, 1997; Mine *et al.*, 1998; Lewinson, *et al*., 2002). MdfA translocates an extremely large range of substrates, including various cationic or zwitterionic lipophilic compounds, such as ethidium bromide, tetraphenylphosphonium, rhodamine, daunomycin, benzalkonium, rifampin, tetracycline, and puromycin (Edgar and Bibi, 1997). Additionally, MdfA can also translocate chemically unrelated antibiotics such as chloramphenicol, erythromycin, and certain aminoglycoside and fluoroquinolone molecules (Edgar and Bibi, 1997). MdfA also acts as a K^+-Na^+/proton antiporter and confers resistance to extreme alkaline pH conditions up to pH 10 (Lewinson *et al*., 2004).

Limited structural information is available for MdfA through a secondary structure model constructed based on its hydropathy profile and the distribution of positively charged residues. MdfA is predicted to contain twelve alpha-helices spanning the plasma membrane (Figure 2.8). The predicted transmembrane topology of MdfA agrees with the X-ray crystal structure of the homologous protein, EmrD, which was also found to have 12 transmembrane helices (Figure 2.9) (Yin *et al.*, 2006). MdfA has 26% sequence identity and 39% similarity to EmrD and therefore should have a structure similar to EmrD. The structure for EmrD was solved to 3.5 Ångstrom resolution (Yin *et al.*, 2006).

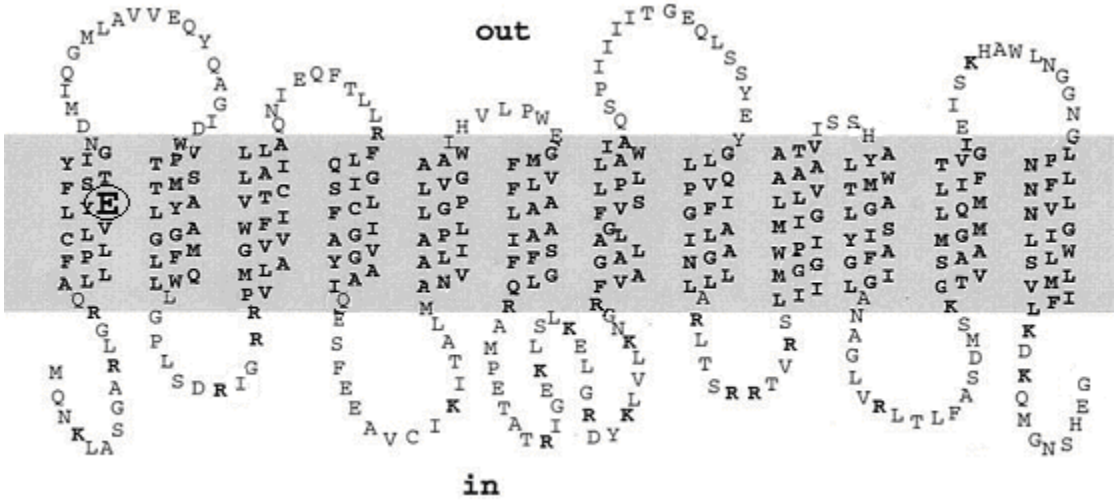


Figure 2.8. Secondary-structure model for MdfA based on hydropathy profile and distribution of positively charged amino acids (Bold) (Adler and Bibi, 2002). Reprinted with permission from the American society of Microbiology. Glu-26 (circled) has been implicated in the binding of cationic substrates.

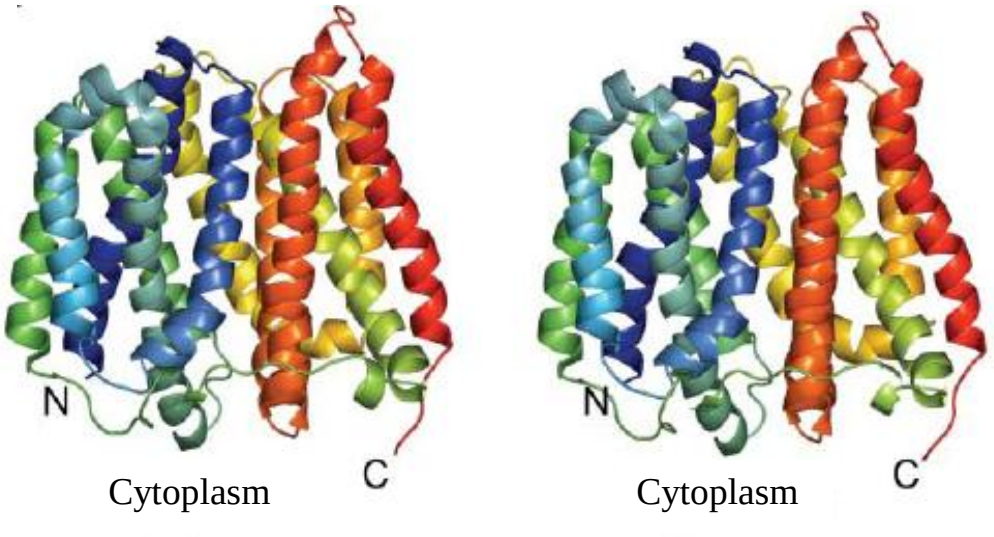


Figure 2.9. Stereo view of the three dimensional structure of EmrD (Yin *et al.*, 2006) (PDB ID 2GFP). Reprinted with permission from AAAS.

The drug extrusion by MdfA is believed to occur via the rocker switch mechanism, which is also proposed to function in EmrD, GlpT, LacY, and other MFS transporters (Figure 2.10) (Abramson *et al.*, 2003; Huang *et al.*, 2003; Yin *et al.*, 2006). In the rocker switch mechanism, the protein starts in a state where the central cavity is open to the cytoplasm. The substrate then enters the central cavity and binds to the protein. For EmrD, and other proteins that transport lipophilic molecules, the substrate can also enter from the inner membrane leaflet (Path 1, Figure 2.10) or through the cytoplasm (Path 2, Figure 2.10). The protein then changes conformation coupled with H^+ antiport, closing the internal cavity to the cytoplasm and opening it to the periplasm, which allows the release of the substrate. The structures for GlpT and LacY (Figure 2.6 and 2.7 respectively) show these proteins in the cytoplasm open state (A, Figure 2.10), whereas the structure of EmrD (Figure 2.9) shows it in the closed state (B, Figure 2.10).

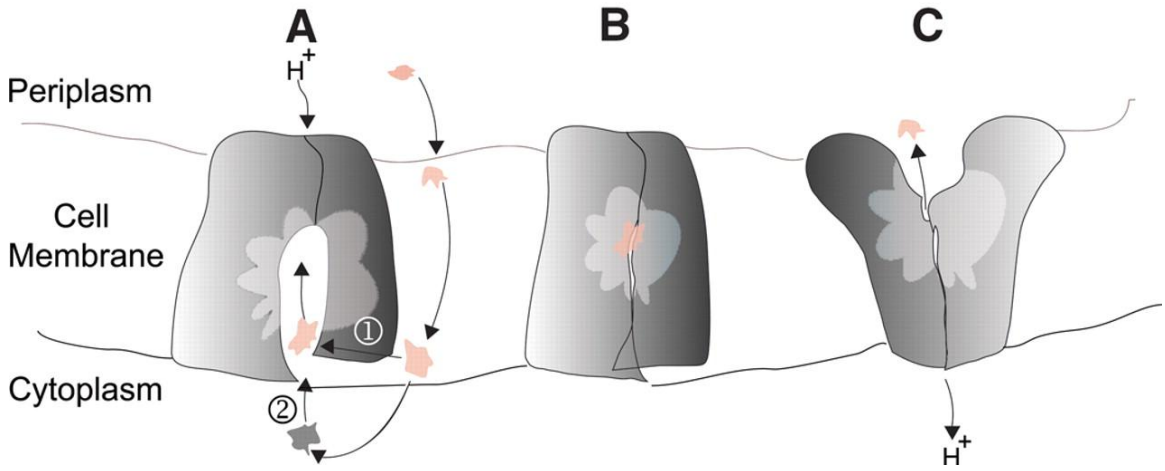


Figure 2.10. Rocker switch mechanism (Yin *et al.,* 2006). (**A**) The substrate enters the internal cavity and binds to the active site. (**B**) The protein closes. (**C**) The drug is released into the periplasm, coupled with proton antiport. Reprinted with permission from AAAS.

The ways in which MdfA recognizes a broad range of substrates can be hypothesized by contrasting the homologous protein EmrD with GlpT. EmrD has a broad substrate range, while GlpT transports only inorganic phosphate and organophosphates. GlpT has two positively charged amino acids (Arg^{45} and Arg^{269}) involved in substrate binding (Huang *et al.,* 2003), while EmrD is believed to have seven residues with hydrophobic side-chains involved in substrate binging (Ile^{28}, Ile^{217}, Ile^{253}, Tyr^{52}, Tyr^{56}, Trp^{300}, and Phe^{249}) (Yin *et al.,* 2006). GlpT's two positively charged Arg side chains electrostatically interact with negatively charged phosphate groups, which results in tight and highly specific binding of the substrate molecules. EmrD has several amino acids involved in substrate binding resulting in several weaker interactions, probably van der Waals interactions, with the outcome being lower binding affinity. The weak interactions result in lower specificity, and different combinations of the interactions can account for recognizing a variety of different substrates. For example, figure 2.11 depicts the binding of three different substrates to the same binding site of the transcriptional regulator QacR. Different combinations of amino acids are involved in the binding of the different substrates. The substrate binding site for MdfA has been implicated to contain the amino acids Cys^{21}, Glu^{26}, Gly^{39}, Val^{54}, Thr^{56}, Ala^{128}, Ala^{147}, Ala^{191}, and Val^{335} (Edgar and Bibi, 1999; Adler and Bibi, 2004; Adler, *et al.,* 2004), which would allow for the recognition of a large range of substrates. MdfA has only one charged amino acid located

within the lipid bilayer, Glu^{26} (Fig 2.8), which has been found through mutagenesis to be essential for the recognition and binding of cationic substrates, but not for proton translocation (Adler and Bibi, 2004). Most of the other implicated amino acids are either hydrophobic or slightly hydrophobic and are likely involved in the binding of zwitterionic lipophilic substrates.

MdfA is an excellent candidate for structural and biochemical studies because of its relatively small size and its substrate specificity. The protein has a large substrate range, yet it is specific enough not to transport normal constituents out of the cell. Fully understanding the substrate binding site of this protein would be extremely interesting and the knowledge acquired could be applied to other multidrug transporters, potentially aiding in the treatment of multidrug resistant cancers, epilepsy, AIDS and bacterial infections which have become multidrug resistant due to the over-expression of multidrug transporters.

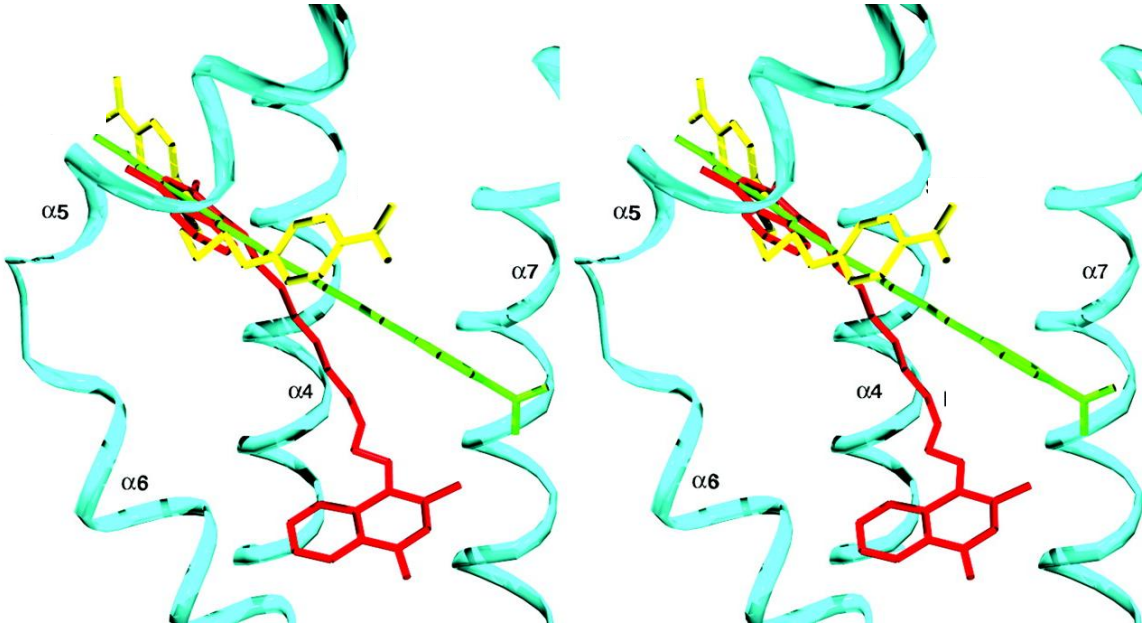


Figure 2.11. Stereo view of the superposition of three substrates in the QacR binding site (Murray *et al.*, 2004) (PDB ID 1RPW). The substrates are pentamidine (yellow), hexamidine (green) and dequalinium (red).

2.4 Protein Structural Studies

Protein structural studies are a very important part of biochemistry. The elucidation of protein structures can yield detailed information about how proteins and enzymes function. There are two main methods for solving three-dimensional protein structures: X-ray crystallography and NMR. X-ray crystallography is the most common method used for structural studies and has produced the most structures. On March 10, 2009 there were 56,217 structures deposited in the Protein Data Bank (PDB), and of those structures 48,162 were X-ray crystal structures, 7,729 were NMR structures, and 221 were cryo-electron microscopy structures.

2.4.1. X-ray Crystallography

Protein crystals of high quality are required for solving a high-resolution protein structure by X-ray crystallography. Protein crystals are made up of regularly repeating units of identical shape and size called unit cells. Protein crystals are formed by slowly supersaturating a purified protein sample, forcing the protein out of the liquid phase and into the solid phase as a highly-ordered crystal. The most common method for obtaining a supersaturated protein solution is by adding a precipitant which leads to a change in water content. A protein crystal is then placed into an X-ray beam, where the electrons from the atoms in the crystal lattice diffract the incident X-ray beam. An electron density map can be then generated from the observed X-ray diffraction pattern. Phase angles cannot be directly recorded during diffraction. Prior to generating an electron density map, the phase angles must be determined using one of three methods: multiple isomorphous replacement, multiwavelength anomalous dispersion, or molecular replacement. Once the electron density map is obtained, the protein model must be fit into the electron density map and refined.

There are several benefits to X-ray crystallography. A major advantage of X-ray crystallography is that there is no protein size limitation. Proteins of all sizes have been crystallized and their structures solved. The structure for the complete 70S ribosome was solved by X-ray crystallography to the resolution of 2.8 Ångstroms (Selmer *et al.*, 2006). The 70S ribosome is about 2.5 MDa in size and is made up of three large RNA molecules and over 50 small proteins. The largest advantage to X-ray crystallography is the ability to obtain extremely high resolution structures. At the resolution of 5.5 Ångstrom, the overall shape of

the molecule can be determined. At 3.5 Ångstrom resolution, the main chain of the protein can be observed. At 3 Ångstrom resolution, the side chains can be partially resolved. At 1.5 Ångstrom resolution, the amino acid side-chains can be completely resolved and the atoms in the structure will be located to about +/- 0.1 Ångstrom accuracy. The best resolution obtained using X-ray crystallography is 0.54 Ångstroms for the protein Crambin (Jelsch *et al.*, 2000). Membrane proteins generally have poorer resolution compared to soluble protein. However, X-ray structures for membrane proteins have been solved to an extremely high resolution. The structure of Aqy1 yeast aquaporin was solved to 1.15 Ångstrom resolution (Fisher *et al.*, 2009).

X-ray crystallography also has its drawbacks. The biggest drawback to X-ray crystallography is the need to obtain suitable protein crystals. The process of protein crystallization can be very difficult and time consuming. This is especially true for membrane proteins. Membrane proteins are relatively hydrophobic and need to be solubilized in order to be extracted from the lipid bilayer. Solubilization can be achieved using detergents. Detergents interact with hydrophobic surfaces on the protein, forming micelles around them which allows for their extraction from the lipid bilayer. The presence of detergents greatly hampers crystallization attempts. Detergent micelles conceal many potential crystal contacts, which are required for the formation and growth of crystals. The few remaining contacts can result in crystal formation, but the protein crystals are generally delicate and fragile. The micelle that surrounds the protein generally has a flexible and dynamic nature which also interferes with crystal formation. The magnitude of this problem is reflected in the statistics for crystallized proteins. In the PDB, there are over 48,000 protein crystal structures, and of those structures, only approximately 200 are of unique membrane proteins (http://blanco.biomol.uci.edu/ Membrane_Proteins_xtal.html, June 24/2009).

A second disadvantage of X-ray crystallography is encountered with substrate binding studies. The substrate either needs to be soaked into the crystal or co-crystallized with the protein, which can also result in crystallization difficulties. If the protein has multiple substrates, this opens up the possibility of multiple crystal forms and multiple crystallization conditions which makes crystallization trials time consuming.

Another drawback to X-ray crystallography is that the chemical composition of the crystallization liquor can potentially lead to conformational changes, resulting in a non-

physiological protein structure. Crystal formation itself may select one of several protein conformations present in the solution, or force the protein into a region of conformational space, sparsely populated under physiological conditions. X-ray structures give static models, whereas proteins are dynamic. Crystallization restricts the protein's movements and may result in a physiologically irrelevant structure. Mobile or disordered regions of the protein will not be seen in X-ray crystallography, and may hinder crystallization attempts. That being said, most X-ray structures provide physiologically relevant information. The structures of many proteins have been solved using both X-ray crystallography and NMR. Garbuzynskiy and co-authors compared the X-ray and NMR structures for some 60 proteins. In most cases, the X-ray and NMR structures of the same protein had only small differences, possibly due to packing restrictions in crystals (Garbuzynskiy *et al.*, 2005). The pairwise RMSDs of the backbone atoms were calculated comparing the X-ray structures to the NMR structures of the 60 different proteins (Garbuzynskiy *et al.*, 2005). The RMSD values for 42 of the 60 proteins were under 2.00, including 0.52 for ubiquitin, 0.57 for interleukin-1 beta, and 0.83 for human cyclophilin A. There were only five proteins with RMSD values higher than 3, the highest being 3.99. These low RMSD values indicate that there was no significant difference between the X-ray and NMR structures for most proteins. Despite the drawbacks to X-ray crystallography, it is still the most powerful method of protein structural analysis.

2.4.2. Nuclear Magnetic Resonance

Nuclear Magnetic Resonance spectroscopy can also be used to solve protein structures. Chemically distinct protons, carbon atoms and nitrogen atoms in the protein can be identified through the measurement of the chemical shifts of their respective nuclei. The interatomic distances can be then estimated by measuring Nuclear Overhauser effect, magnetization transfer through space, which makes protein structure elucidation by NMR possible. Measuring magnetization transfer through space allows the computation of distances between protons, and thus the distances between structural elements. The compilation of multiple distance constraints within the protein can then be used to produce a three-dimensional protein model.

Uniform isotope labeling allows for the collection of information on the entire protein. Isotope labeling is necessary in protein NMR because naturally-occurring isotopes of carbon

and nitrogen are not useful for high resolution NMR studies. ^{12}C has no magnetic moment and therefore is not NMR-active, while the high nuclear quadrupole moment of ^{14}N prevents obtaining high resolution spectra. Nuclei with spin number of ½, such as ^{13}C and ^{15}N, have the most favorable properties for high resolution NMR. Therefore, uniform isotope labeling of proteins using ^{13}C and ^{15}N isotopes allows for high resolution NMR studies of proteins. However, when investigating specific portions of a protein molecule, such as using chemical shift perturbation experiments to map a drug binding site, the use of uniform isotope labeling would result in a very complicated spectrum wherein the chemical shifts would be difficult to observe. Such experiments are made easier by using selective isotope labeling. By labeling only amino acids believed to be involved in the binding of substrates, the resulting NMR spectrum is greatly simplified. Through selective isotope labeling, it is possible to map active sites of protein without having to solve the entire structure. By labeling only the amino acids thought to be involved in substrate binding, the intermolecular distances can be calculated and a three-dimensional map of the active site can be generated.

The range of proteins that can be effectively studied by high-resolution NMR is limited by molecule size restrictions. The larger the protein, the faster the relaxation of the NMR signals. Fast relaxation results in broad NMR spectral lines, which in turn leads to loss of resolution and sensitivity. Additional difficulties arise in the NMR studies of membrane proteins, such as the choice of detergent, which can play a very important role in the NMR experiment. The detergent will form a micelle around the protein, adding to the overall size of the molecule. It is important to choose a detergent that will result in the smallest micellar size, yet still maintain the proper folding of the protein being studied.

Several methods have been developed to deal with large proteins in NMR studies. One method is deuteration of the protein. Deuteration improves NMR spectra for large proteins by reducing the dipolar interactions between ^{13}C or ^{15}N and the directly bound proton spins, which is the main source of relaxation in ^{13}C- and ^{15}N-labeled proteins (Browne *et al.*, 1973; Grzesiek *et al.*, 1993), significantly increasing relaxation time. The most significant method developed to address the problem of magnetization relaxation of large proteins is transverse relaxation optimized spectroscopy (TROSY) (Pervushin, *et al.*, 1998). The TROSY experiment results in the spectral peaks appearing as multiplets because decoupling

has not been applied. TROSY is designed to then select only the narrow spectral component of N-H multiplets, resulting in narrower line widths, and dramatically higher resolution.

An advantage of NMR is that proteins are studied in solution, which avoids the difficulties of protein crystallization. Unlike crystal structures, NMR protein structures are not static. Structures solved by NMR contain information on the conformational flexibility of the protein. Additional NMR techniques can be used to obtain detailed information on protein dynamics. Heteronuclear Nuclear Overhauser effects and relaxation rates provide direct information on the range and timescale of local molecular motions. The extent of conformational flexibility can be calculated as an order parameter using the Lipari-Szabo model free approach (Lipari and Szabo, 1982). This in combination with Nuclear Overhauser effects and residual dipolar coupling measurements (Bouvignies *et al.,* 2007) can be used to describe molecular motions fairly accurately.

Another advantage of NMR is the ability to determine what amino acids are involved in substrate binding by using chemical shift perturbation analysis. Chemical shifts are sensitive to changes in the local environment. Chemical shifts for those amino acids involved in substrate binding will experience perturbations upon the addition of substrate to the protein solution. These chemical shift perturbations can be observed by comparing NMR spectra from a protein sample with and without substrate added. Interpretation of such experiments may be complicated by the fact that chemical shift perturbations will also be observed for an amino acid residue located in the region, which experience significant structural changes because of substrate binding.

The structures of several membrane proteins have been solved using NMR experiments. Two examples are KcsA and diacylglycerol kinase. KcsA is a potassium channel which was extracted from the membrane using the detergent foscholine. NMR studies revealed that KcsA exhibits a tetrameric arrangement of helical KcsA monomers (Figure 2.12). The NMR structure obtained for KcsA is consistent with the previously determined X-ray structure (Yu *et al.,* 2005). Unlike KcsA, the structure determined for diacylglycerol kinase is unique, as there is no X-ray structure available. The NMR structure of diacylglycerol kinase solubilized in dodecylphosphocholine revealed a homotrimer with interlinking helices (Figure 2.13) (Van Horn *et al.,* 2009).

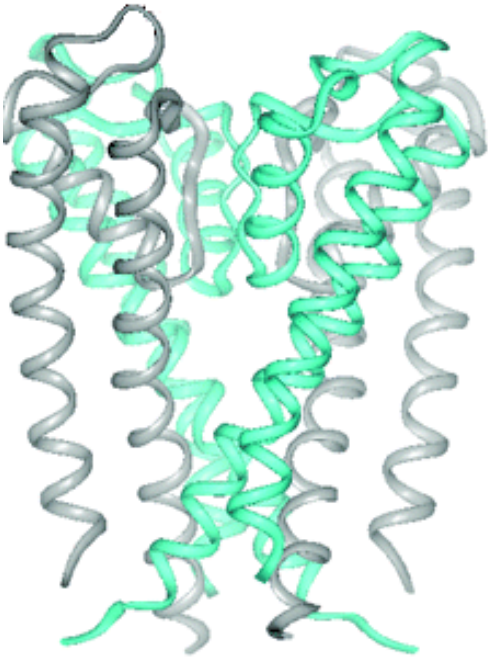


Figure 2.12. Side view of the ribbon representation of KcsA (Yu *et al.*, 2005) (PDB ID 2A9H). Two monomers are colored green and two monomers are colored grey. Reprinted with permission from the American Chemical Society.

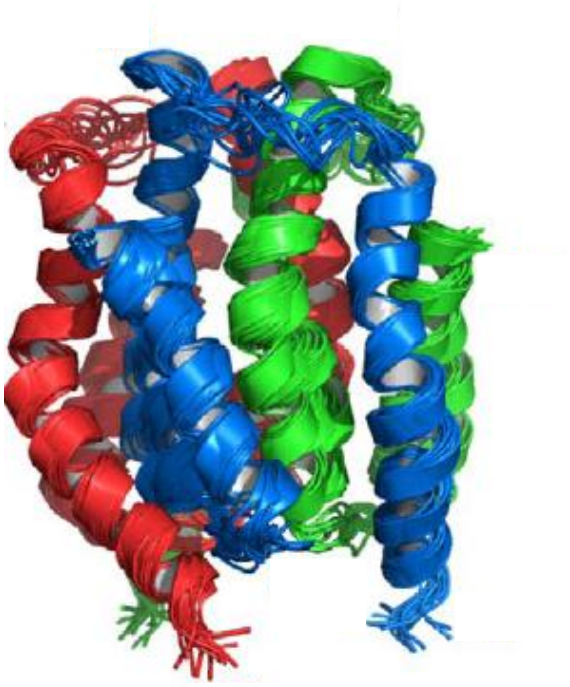


Figure 2.13. Diacylglycerol kinase structure ensemble comprised of the 16 lowest energy structures (Van Horn *et al.*, 2009) (PDB ID 2KDC). Reprinted with permission from AAAS.

2.5. Overview of Protein Preparation for Structural Studies.

In order to conduct protein structural studies using X-ray crystallography or NMR, it is important to obtain pure protein. In order to obtain pure protein, one must choose a suitable expression system. Generally, it is preferable to produce the protein in the correctly folded conformation and not within inclusion bodies. Once an appropriate expression plasmid is chosen with a compatible bacterial host, the correctly folded, over-expressed protein must then be purified from the bacterial cells. This is done by rupturing the bacteria and then using protein fractionation methods to remove contaminants and isolate the protein of interest. If the protein is not expressed in a soluble form, it may be necessary to use detergents to solubilize the protein prior to purification.

Initial screening of crystallization conditions in X-ray crystallographic studies typically requires 1 mL solution at a protein concentration of 10 mg/mL and at 95% purity or higher. Protein NMR spectroscopy typically requires a sample between 300-600 μL in volume with a protein concentration of 0.1 mM- 3.0 mM. It is also extremely important that the pure protein is stable. Many structural studies require proteins to be stable for weeks or months in a range of temperatures and experimental conditions. Protein stability can be affected by many parameters, including pH, salt concentration, presence of inhibitors and cofactors, residual protease activity and detergents employed in the study.

Another important characteristic of the protein in structural studies is its aggregation state. It is vital that the protein is present in solution in a monodisperse, though not necessarily monomeric state. Dimers, trimers and aggregates of higher order can be analyzed by X-ray crystallography and NMR, as long as the protein is present as a single species. A non-uniform aggregation state of a protein solution can inhibit the formation of high resolution crystals in X-ray crystallography. Protein aggregation in NMR will lead to very broad spectral lines resulting in peak overlap, signal disappearance and an overall poor NMR spectrum. The aggregation state of a protein in solution can be analyzed using dynamic light scattering experiments. Dynamic light scattering is a technique in which a beam of monochromatic light is directed through a purified protein solution and the fluctuations of intensity of scattered light is analyzed. This information can be used to determine the particle size distribution of the solution, and thus the aggregation state of the protein in solution. The aggregation state of the

protein can be affected by several different factors, including protein concentration, salt concentration and detergent concentration.

2.5.1. Cloning the Gene of Interest into an Expression Vector

Expression plasmids are commonly constructed to encode at least one antibiotic resistance gene for selection, an inducible promoter, and a purification tag that can be linked to the protein to be expressed. Some examples are the pET vectors (EMD Biosciences, San Diego, California), the pTYB vectors (New England BioLabs, Pickering, Ontario), and the pBAD vectors (Invitrogen, Burlington, Ontario). The pBAD vector features an ampicillin resistance gene for selection, a thioredoxin gene for possible purification of a fusion protein, a V5 epitope for antibody detection, and an (L)-arabinose inducible promoter. The isopropyl β-D-1-thiogalactopyranoside inducible promoter, found in the pET and pTYB vectors, is the most commonly used inducible promoter. The pBAD plasmid with the (L)-arabinose inducible promoter was chosen for this study because it offers tight control of expression, the inducing agent is relatively inexpensive, and MdfA has been successfully expressed previously using a similar system (Lewinson and Bibi, 2001). As well, some other membrane proteins such as GlpT and the MexA, B-OprM extrusion pump have also been successfully expressed using this system (Guan *et al.,* 1999; Huang *et al.,* 2003).

The arabinose promoter offers extremely tight control of gene expression. The pBAD plasmid also encodes the araC gene. AraC both positively and negatively regulates transcription based on arabinose and cyclic AMP (cAMP) concentrations (Ogden *et al.,* 1980; Schleif, 1992). AraC functions as a dimer, and in the absence of arabinose it binds the O_2 and I_1 sites in the plasmid (Figure 2.14), creating a DNA loop which inhibits transcription. In the presence of arabinose, the dimer will release the O_2 site and bind to the I_2 site, which releases the DNA loop and allows transcription to begin. The cAMP activator protein (CAP) binds to DNA in the presence of cAMP further promoting AraC binding to I_2. Upon the addition of glucose, cAMP levels are lowered, resulting in reduced binding of CAP, which leads to the decrease of transcriptional activation. Arabinose concentrations in the cytoplasm, and therefore gene expression level, can be easily adjusted by using a strain which is unable to metabolize (L)-arabinose. Being able to tightly regulate gene expression may be necessary if the protein is

toxic to the bacteria or if high concentrations of the protein results in the aggregation and precipitation of the protein.

Another option for protein expression is to use a cell-free expression system such as the one described by Torizawa and co-authors (Torizawa *et al.*, 2004). Cell-free synthesis exploits the cellular protein synthesis machinery to direct protein synthesis outside intact cells using exogenous messenger RNA or DNA as a template (Figure 2.15). This is achieved by combining a crude lysate from growing cells, containing necessary enzymes and machinery for protein synthesis, with an exogenous supply of amino acids, nucleotides, salts and energy-generating factors. Since the protein synthesis is done *in vitro*, and not in a bacterial cell, this method allows production of toxic proteins. The pEXP-DEST (Invitrogen, On, Canada) cell-

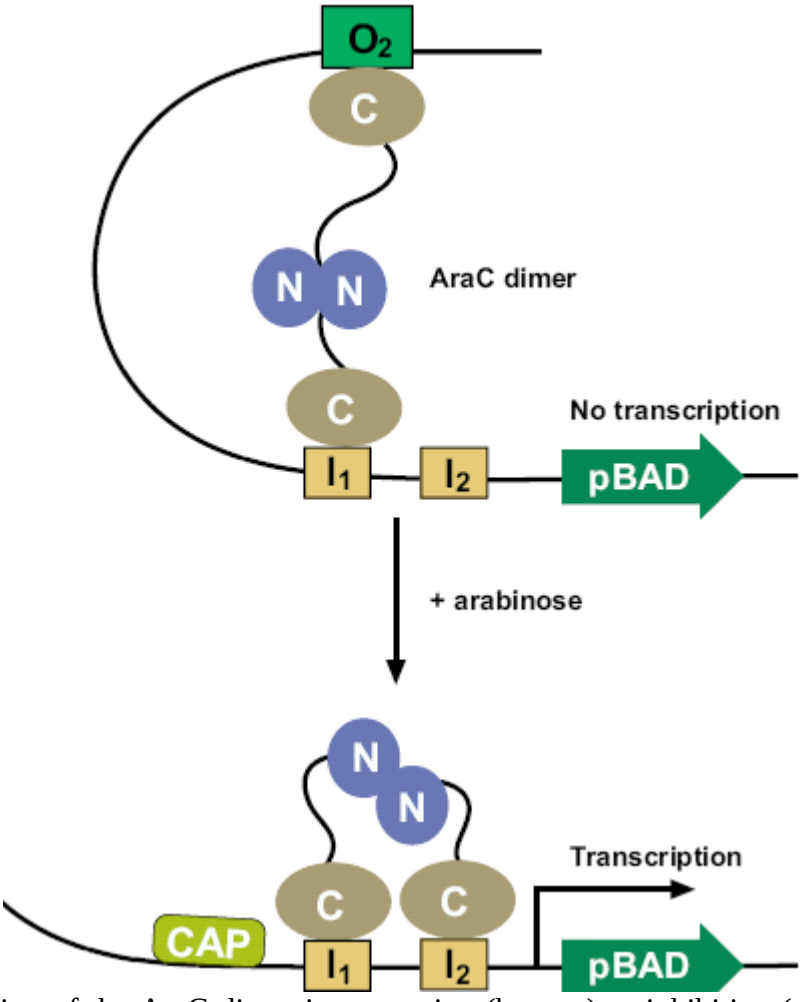


Figure 2.14. The action of the AraC dimer in promoting (bottom) or inhibiting (top) protein expression from the pBAD expression plasmid. (pBAD Directional TOPO Expression Kits instruction manual (Invitrogen, On, Canada)).

free synthesis plasmid includes a T7 RNA polymerase promoter to ensure selective gene expression from this plasmid using T7 RNA polymerase. Once an expression system is selected, the cDNA encoding the protein is sub-cloned into the plasmid. This can be accomplished by modifying the gene sequence encoding the protein to be flanked by restriction sites chosen for cloning into the specific vector. The restriction sites can be inserted at the 5' and 3' ends of the gene of interest through polymerase chain reaction with DNA primers encoding the restriction sites. Once the matching restriction sites are added, both the gene and the vector can be digested with the restriction enzymes, ligated together and the ensuing product is then transformed into the bacterial host. Constructs are screened via restriction analysis and then verified by DNA sequencing. An alternate option is to use homologous DNA recombination methods.

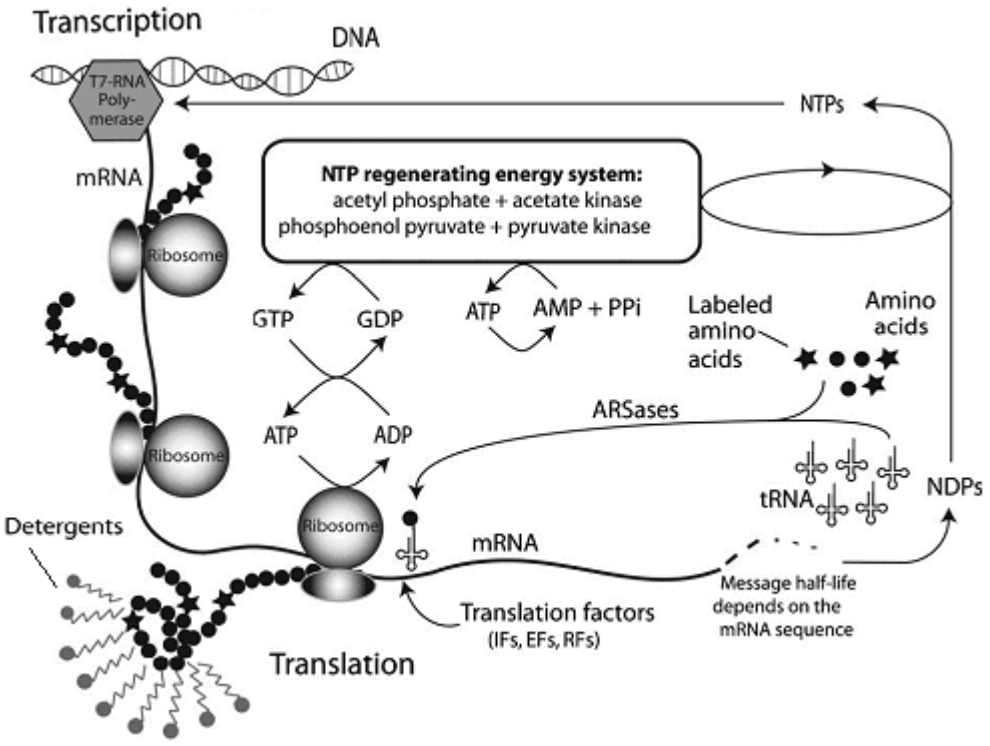


Figure 2.15. The principle of cell-free synthesis (Schwarz *et al.*, 2007). A crude lysate from growing cells containing the necessary enzymes and machinery for protein synthesis is combined with exogenous supply of amino acids, nucleotides, salts and energy-generating factors. ARSases: aminoacyl-tRNA synthetases. Detergents can be added to aid in protein folding and solubility. Reprinted with permission from Elsevier.

2.5.2. Optimization of Protein Production.

There are many ways to increase protein expression. One method is to test expression of the gene of interest in different bacterial strains. For example, BL21 (DE3) derivatives such as C43 (DE3) and C41 (DE3) were designed for the over-expression of toxic proteins as well as membrane proteins (Miroux and Walker, 1996). Varying the inducing agent concentration can also result in higher levels of protein expression. However, it is important that the expressed protein is properly folding. If too much protein is being expressed too quickly, the protein may not fold correctly, which can result in aggregation leading to formation of the inclusion bodies. Therefore, optimal inducing agent concentration may not be that which results in the highest level of expression. Increasing incubation time can also result in higher expression levels.

2.5.3. Producing Isotopically Labeled Protein samples.

NMR experiments may require uniformly or selectively isotopically labeled protein samples. There are several ways to accomplish such a labeling. For uniformly labeled protein samples, it is common to express the protein in bacteria growing on a minimal medium containing the source of the required isotope for incorporation into the amino acids during their biosynthesis in the bacterial cell. An example would be to use a growth medium where the only nitrogen source is ^{15}N ammonium chloride, which would result in the uniform labeling of the backbone amide groups of the protein with ^{15}N, as well as Asn, Gln, Arg, Trp, His, and Lys side chains. A similar approach can be used to obtain selectively labeled protein samples by using a medium supplemented with isotopically labeled amino acids. This, however, requires a large amount of labeled amino acids and can lead to improper labeling and isotope dilution as a result of metabolic scrambling. Metabolic scrambling occurs when one amino acid is converted into another through metabolic pathways in the bacterial cell. Metabolic scrambling can be overcome by using auxotrophic strains of bacteria for protein expression, or by using cell-free protein synthesis. Auxotrophic strains of bacteria contain mutations which render them unable to synthesize specific amino acids. These amino acids can be added externally in a labeled form and will be incorporated into the newly synthesized proteins largely avoiding metabolic scrambling. There is also no metabolic scrambling in cell-free protein synthesis because the amino acid metabolic activity in cell-free extracts is low (Kigawa *et al.,* 1995). The amount of

isotopically labeled amino acids needed for cell-free synthesis is much lower, because only the protein of interest is produced.

2.5.4. Purification of Cell Membranes

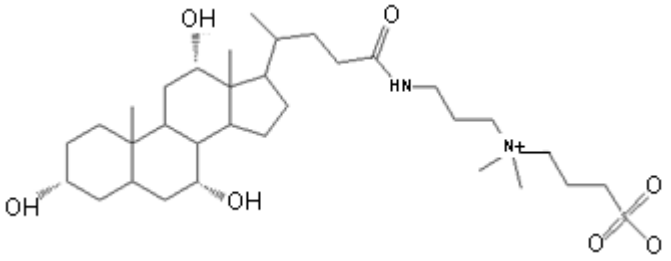
When purifying membrane proteins, it is necessary to separate the cell membrane from soluble cytosolic protein, as well as other cell debris. Once the cell culture has been pelleted using centrifugation, the cells can be disrupted using a method such as French press or sonication, then the cell membranes can be separated from cytosolic proteins through a series of centrifugation steps. The first centrifugation is usually carried out at speeds resulting in the force of around 6000 x g, separating the membranes and soluble protein from other cell debris, including inclusion bodies. Ultracentrifugation at around 100,000 x g is then used to separate the soluble cytosolic proteins from the membranes, as the membranes will pellet out of solution. The cell membranes are collected and used for protein purification.

2.5.5. Protein Extraction from the Cell Membrane

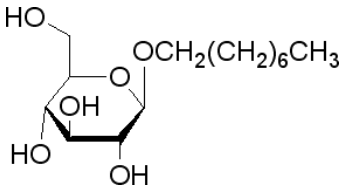
Membrane proteins need to be extracted from the cell membrane in order to be purified. This is accomplished by adding a detergent which will interact with the hydrophobic surfaces of the protein, thereby allowing the protein solubilization. Detergents are amphipathic molecules comprising a polar head group and a non-polar hydrocarbon tail. Detergents used for solubilization must maintain both the structural integrity and biological activity of the protein of interest. Membrane solubilization trials are performed using a wide range of detergents at concentrations higher than their critical micelle concentrations (CMC). The CMC is a minimum concentration where detergent monomers aggregate, forming micelles into which membrane proteins can be inserted. Membrane proteins have been purified in a wide range of detergents. Some detergents which have been used to solubilize membrane proteins include: 3-[3-(Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) (Banerjee *et al.,* 1995; Cladera *et al.,* 1997; Chattopadhyay, *et al.,* 2002), *n*-dodecyl-β-D-maltoside (DDM) (Ambramson *et al.,* 2003; Yin *et al.,* 2006), sodium cholate (Rivnay and Metzger, 1982) and Triton X-100 (Aller *et al.,* 2009). Several membrane transporters of MFS type, including LacY, EmrD, EmrE, GlpT and MdfA have been successfully solubilized using DDM (Auer *et al.,* 2001; Ambramson *et al.,* 2003; Adler, *et al.,* 2004; Yin *et al.,* 2006; Korkhov and Tate, 2009). When conducting NMR experiments on the solubilized protein, the detergent must not

interfere with the quality of the NMR spectra. Krueger-Koplin and co-authors (Krueger-Koplin *et al.*, 2004) evaluated the quality of NMR spectra of subunit *c* of the *E. coli* F_1F_o ATP synthase, subunit *c* from *B. pseudofirmus OF4*, Smr from *S. aureus*, and LH1 α and β subunits from *R. sphaeroides*, recorded in different detergents. Of these detergents, CHAPS, DDM and Triton resulted in poor quality NMR spectra, while detergents such as 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-1-(glycerol)] (LMPG) and 2-Diheptanoyl-*sn*-Glycero-3-Phosphocholine (DHPC) resulted in NMR spectra of high quality. The structures of the detergents discussed can be seen in figure 2.16 and are listed in table 2.1.

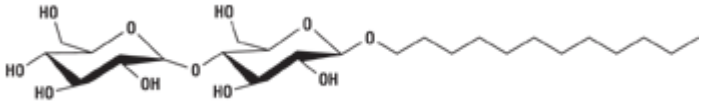
a) 3-[3-(Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)



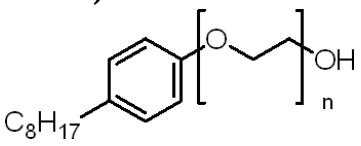
b) octylglucoside (OG)



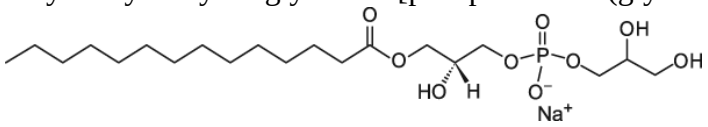
c) *n*-dodecyl-β-D-maltoside (DDM)



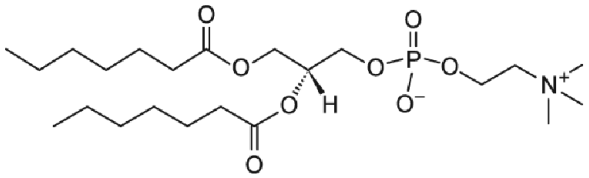
d) Triton X-100



e) 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-1-(glycerol)] (LMPG)



f) 2-Diheptanoyl-*sn*-Glycero-3-Phosphocholine (DHPC)



g) Sodium Cholate

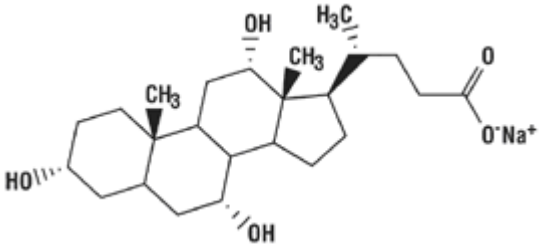


Figure 2.16. Chemical structures of detergents commonly used to solubilize membrane proteins.

Table 2.1. A list of detergents commonly used to solubilize membrane proteins.

Figure 2.16	Detergent Name	CMC	Common Abbreviation
a)	3-[3-(Cholamidopropyl)dimethylammonio]-1-propanesulfonate	6-10 mM	CHAPS
b)	octylglucoside	19 mM	OG
c)	*n*-dodecyl-β-D-maltoside	0.17 mM	DDM
d)	Triton X-100	0.23 mM	na
e)	1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-*rac*-1-(glycerol)]	0.05 mM	LMPG
f)	2-Diheptanoyl-*sn*-Glycero-3-Phosphocholine	1.4 mM	DHPC
g)	3 ,7 ,12 -Trihydroxy-5β-cholan-24-oic acid, monosodium salt	9-14 mM	Sodium Cholate

2.5.6. Protein Purification

There are several different protein purification methods, including size exclusion chromatography, ion-exchange chromatography, and affinity chromatography. Size exclusion chromatography separates proteins based on their molecular size and shape. A gel matrix of porous beads is used for the immobile phase in the column chromatography. Smaller proteins follow a longer trajectory through the pores, whereas larger proteins are unable to enter the smaller pores, which effectively leads to separation based on molecule size and shape. Ion-exchange chromatography separates proteins based on charge. The chromatography matrix can either be positively charged (anion-exchange) binding negatively charged proteins, or be negatively charged (cation-exchange) binding positively charged proteins. The bound proteins can then be serially eluted by gradually increasing the salt concentration in the solvent. Affinity chromatography is a chromatographic method for separating proteins based on highly specific interactions, such as those between antigen and antibody, or receptor and ligand. Affinity chromatography methods commonly used for purification of recombinant proteins generally involve modification of the protein sequence by adding a specific amino acid sequence, called a tag, which has an affinity for the immobile phase of the chromatography column. The elution step will vary based on what affinity chromatographic method is used.

At present, affinity chromatography is the most common method of purifying proteins. Some examples of tags used are Glutathione S-transferase (GST) which binds glutathione resin

(Smith and Johnson, 1988), poly-histidine patches which bind divalent metals (Hochuli *et al.*, 1987), maltose binding protein which binds amylase resin (Bedouelle *et al.,* 1987), and intein-chitin binding domain (CBD), which binds chitin (Chong *et al.*, 1997). In GST affinity chromatography, the vector has been modified so the fusion protein can be cleaved from GST by digestion with the site specific protease thrombin (Smith and Johnson, 1988). Poly-histidine-tagged proteins can be eluted from their interaction with divalent metals, such as nickel ions chelated by nitrilotriacetate groups, by washing the column with imidazole which out-competes histidine for the divalent metal. Maltose-binding protein-tagged proteins are eluted off amylase resin by washing it with maltose (Bedouelle, *et al.,* 1987). Intein-CBD-tagged proteins are eluted from immobilized chitin by incubating with a reducing agent such as 1, 4-dithiothreitol which reduces the disulfide bond connecting the protein to the tag (Chong *et al.,* 1997). These affinity purification methods often offer efficient one-step purification, resulting in pure protein.

3. Materials and Methods[1]

3.1. Plasmids and Strains

3.1.1. Generation of a Plasmid for Cell-Based Protein Expression

Plasmids were purified from the *E. coli* cell culture using a Miniprep kit (Qiagen, Mississauga, Ontario). The *mdfA* gene and its native ribosome-binding site were cloned into a pBAD102 vector (Invitrogen) under the control of an *araBAD* promoter (Dmitriev, unpublished data). The native ribosome-binding site is located immediately upstream of the *mdfA* gene. In the resulting plasmid pOD1016, the m*dfA* gene was fused to the V5 epitope sequence for immunodetection and a C-terminal hexa-histidine tag for purification (Figure 3.1 A). The plasmid also contained the thioredoxin gene, which was present in the pBAD102 vector for purification of fusion proteins, but was out of frame with *mdfA* and therefore was not used in our procedure. The pOD1016 plasmid was further modified to delete the sequences encoding for thioredoxin and the V5 epitope. The V5 epitope was deleted because it was not needed, since MdfA was successfully detected using anti-penta-histidine antibody. The gene encoding thioredoxin was deleted to determine if its absence would result in higher levels of *mdfA* expression. This new construct was named pCOG3 (Figure 3.1 B). To construct the pCOG3 plasmid, PCR was carried out using the pOD1016 vector as a template, with primers designed to create a *NcoI* cleavage site at the 5' end (MFM1NCO, Table 3.1), and to delete the V5 epitope sequence (MRM2061, Table 3.1) at the 3' end of the *mdfA* gene. The resulting 1.3 kb product (Figure 3.2, *lane 2*) was digested with the restriction endonuclease *Nco*I. The pOD1016 vector was digested with the restriction endonucleases *Nco*I and *Pme*I, resulting in 1.7 Kb and 4.0 kb fragments (Figure 3.2, *lane 4*). The digested PCR product was ligated to the *Nco*I-*Pme*I fragment of the pOD1016 vector, transformed into DH5α cells, and selected with ampicillin. Plasmid identity was confirmed by restriction digest with *NcoI*, which produced one fragment with the expected size of 5.3 kb (Figure 3.2., *lane 6*) and by DNA sequencing using the primer MRS146 (Table 3.1). Plasmids with the correct sequence were transformed into LMG194 cells (Invitrogen).

[1] Chemicals used in these experiments were purchased from Sigma-Aldrich (Oakville, Ontario) unless otherwise stated.

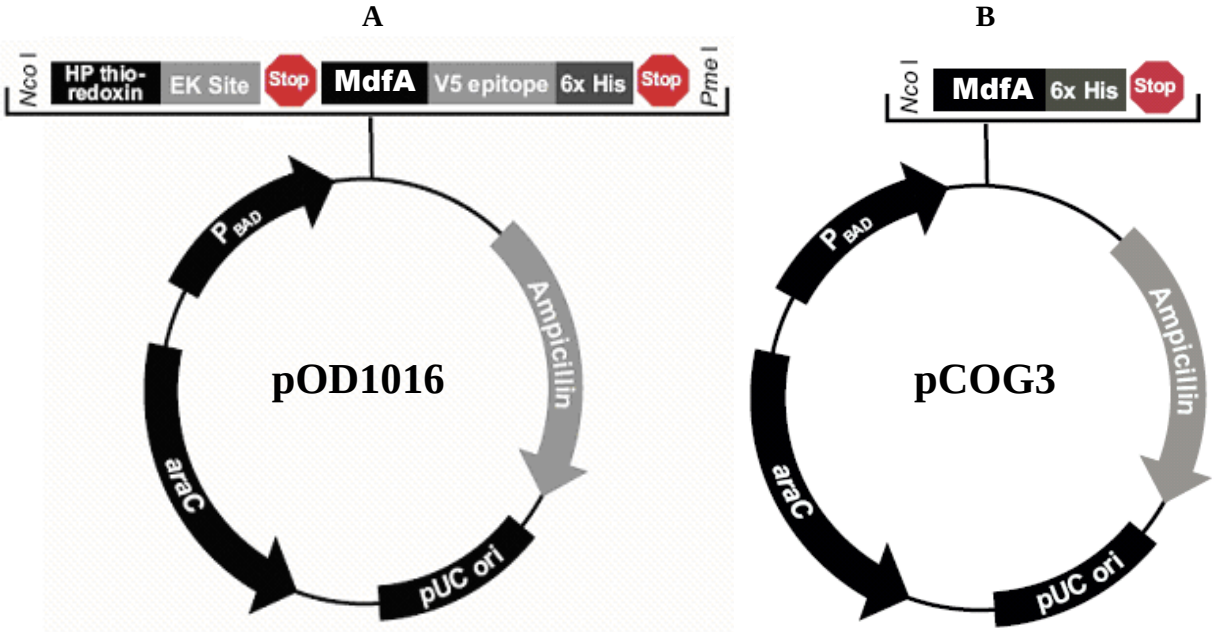


Figure 3.1. The pOD1016 vector (A) and pCOG3 vector (B). The pOD1016 vector has the *mdfA* gene fused to a V5 epitope and the sequence encoding a hexa-histidine tail. A thioredoxin tag and enterokinase (EK) recognition site are present upstream of *mdfA* for use in purification, but are out of frame with *mdfA*, and not used in purification. MdfA was cloned directly into the linearized pBAD vector using blunt-end PCR product and a proprietary process called "directional TOPO cloning" (pBAD Directional TOPO Expression Kits instruction manual (Invitrogen)). In the pCOG3 vector, the *mdfA* gene is fused to the sequence encoding a hexahistidine tag. Included in both vectors are the pBAD promoter, the *araC* gene encoding the regulatory protein for the pBAD promoter, and a pUC origin allowing high-copy replication and maintenance in *E. coli*. *Pme*I and/or *Nco*I cleavage sites are labeled.

Table 3.1. A list of primers used in generating pCOG2 and pCOG3.

Primer Name	Sequence (5'-3')
MFM1NCO	GAAATTCCATGGAAAATAAATTAGCTTCCGGTGCCAGG
MRM2061	TCAATGGTGATGGTGATGATGCCCTTCGTGAGAATTTC
MRS146	ATCAATGCCCGCCTGATATTG
MFM1NDE	TTGCATATGCAAAATAAATTAGCTTCCGGTGCCAGGCTTG
MRM1233	TGCTAGCTCAATGGTGATGGTGATGATGCCCTTCGTGAG
PDEST1F	GACTCACTATAGGGAGACCAC
PDEST1R	TGTTAGCAGCCGGATCAAGC

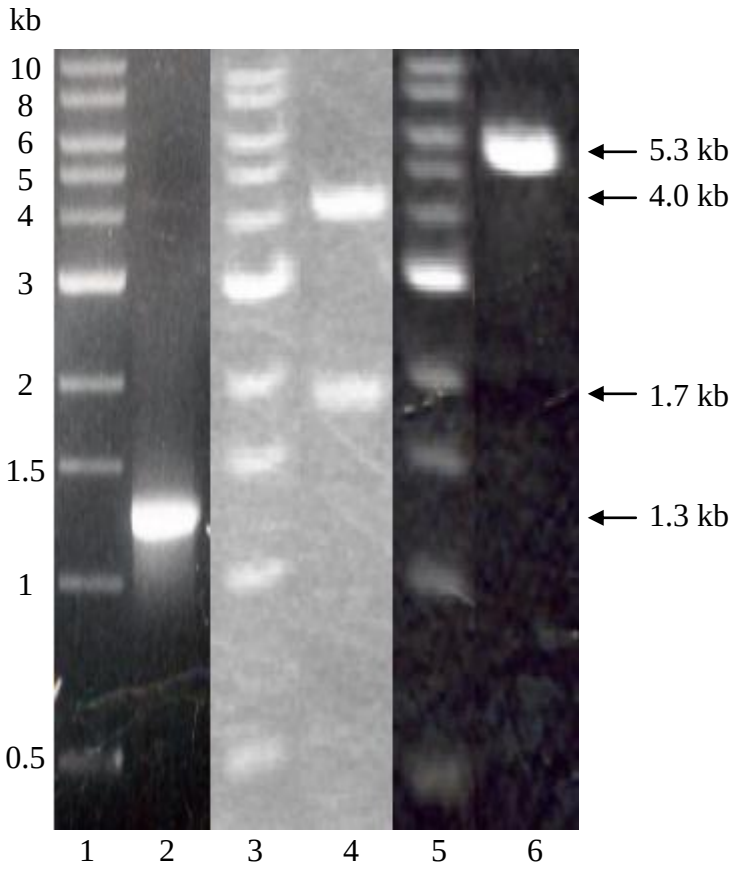


Figure 3.2. Generation of pCOG3. DNA fragments were separated on a 1% agarose gel and stained with ethidium bromide Lanes 1, 3, and 5: DNA standards; lane 2: PCR product containing the *mdfA* gene; lane 4: pOD1016 plasmid digested with the restriction enzymes *Pme*I and *Nco*I; lane 6: pCOG3 digested with the restriction enzyme *Nco*I.

3.1.2. Generation of a Plasmid for Cell-Free Synthesis

The *mdfA* gene was also cloned into the plasmid pEXP1-DEST (Invitrogen) for cell-free synthesis. This new construct was named pCOG2 (Figure 3.3). To construct the pCOG2 plasmid, PCR was carried out using the pOD1016 vector as a template, with primers designed to create both an *Nde*I cleavage site at the 5' end (MFM1NDE, Table 3.1) and an *Nhe*I cleavage site at the 3' end (MRM1233, Table 3.1). The resulting 1.3 kb product (Figure 3.4, *lane 2*) was digested with the restriction endonucleases *Nhe*I and *Nde*I. The pEXP1-DEST vector was also digested with the endonucleases *Nhe*I and *Nde*I, resulting in a 4.6 kb fragment (Figure 3.4, *lane 4*). The digested PCR product was ligated to the *Nhe*I-*Nde*I fragment of the pEXP1-DEST vector, transformed into DH5α cells, and selected with ampicillin. Plasmid identity was confirmed by a restriction digest with *Nhe*I and *Nde*I, which produced two fragments with the expected sizes of 1.3 kb and 4.6 kb (Figure 3.4, *lane 6*), and also by DNA sequencing using the

primers PDEST1F and PDEST1R (Table 3.1). Plasmids with the correct identity were transformed into One-shot ccdB survival T1 phage-resistant cells (Invitrogen).

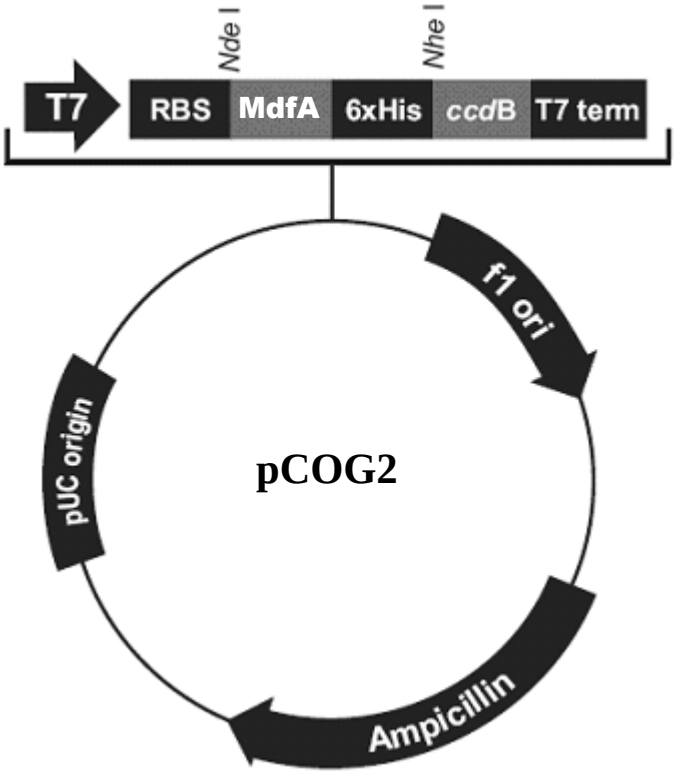


Figure 3.3. The pCOG2 plasmid. This plasmid contains a T7 promoter for specific expression by T7 RNA polymerase, a ribosome binding site, the *ccd*B gene for negative selection of the plasmid, a T7 transcription termination sequence, f1 origin to allow for the rescue of single-stranded DNA, and the pUC origin to permit high-copy replication and maintenance in *E. coli*.

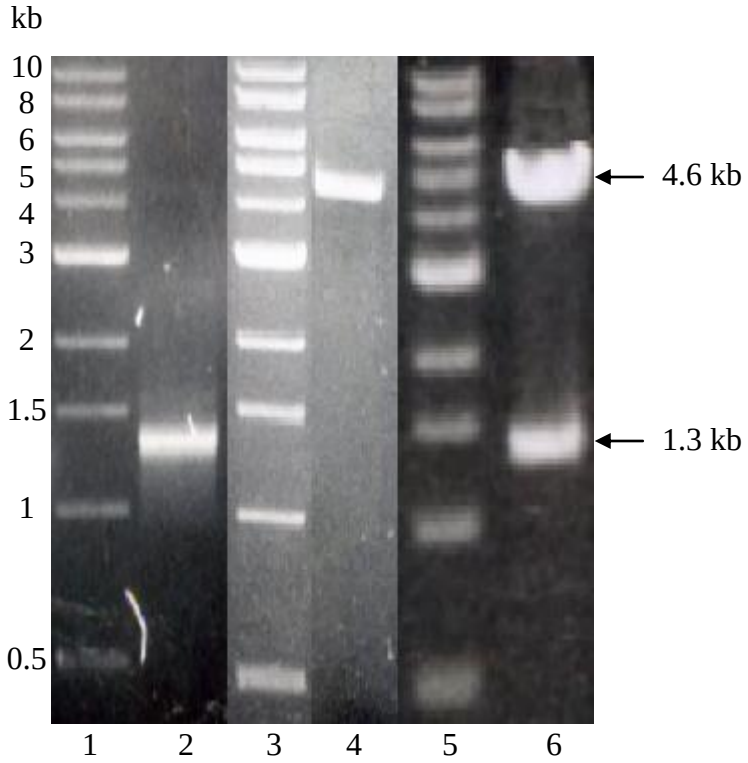


Figure 3.4. Generation of pCOG2. DNA fragments were separated on a 1% agarose gel and stained with ethidium bromide. Lanes 1, 3 and 5: DNA standards; lane 2: PCR product containing the *mdfA* gene; lane 4: pEXP1-DEST plasmid digested with the restriction enzymes *Nhe*I and *Nde*I; lane 6: pCOG2 digested with the restriction enzyme *Nhe*I and *Nde*I.

3.1.3. Bacterial Strains used in the Cloning and Expression of *mdfA*

E. coli strains used in this research include C43 (DE3), DH5α, LMG194, and One-shot ccdB survival T1 phage-resistant cells. The C43 (DE3) cells metabolize arabinose, and have at least one uncharacterized mutation, which prevents the cell death associated with the expression of toxic recombinant proteins (Miroux and Walker, 1996). The C43 (DE3) cells have the genotype F^-*ompT gal hsd*S_B ($r_B^-m_B^-$) *dcm lon* λDE3. The DH5α cells are able to take up large plasmids (nupG) and have the genotype F^- *end*A1 *gln*V44 *thi*-1 *rec*A1 *rel*A1 *gyr*A96 *deo*R *nup*G ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *hsk*R17(r_K^- m_K^+), λ-. LMG194 cells cannot metabolize arabinose (Δ*ara*714) and have the genotype F^- Δ*lac*X74 *gal*E *thi rps*L Δ*pho*A (*Pvu* II) Δ*ara*714 *leu*::Tn10. The One-shot ccdB survival cells are resistant to the ccdB gene product and have the genotype F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G *fhu*A::152.

3.2. Optimization of *mdfA* Expression

MdfA expression in *E. coli* cells was optimized by varying the concentration of the inducer agent (L)-arabinose present in the growth medium. *MdfA* expression from the plasmids pOD1016 and pCOG3 in the strains LMG194 and C43 (DE3) was investigated. Expression was induced in Luria Broth (LB) containing 170 mM NaCl, 10 g/L tryptone, and 5 g/L yeast extract; casamino acid medium (RM) containing 2% casamino acids, 0.5% glycerol, 1 mM $MgCl_2$, 42 mM Na_2HPO_4, 22 mM KH_2PO_4, 9 mM NaCl, and 20 mM NH_4Cl; or M63 minimal medium containing 15 mM NH_4Cl, 15 mM Na_2SO_4, 1 mM $MgSO_4$, 0.5% glycerol, 15 μM thiamine, 2 mM leucine, 100 mM potassium phosphate buffer, pH 7.0. Cell cultures were grown at 37°C on a shaker platform at 200 r.p.m. to an optical density (OD_{600}) of 0.4, and then (L)-arabinose was added to a final concentration of 0.0002% - 2.0%. The cultures were incubated for an additional 4 hours under the same conditions. Culture samples were then taken and diluted to an equal optical density (OD_{600}). The 1 mL aliquots of each sample were then pelleted and resuspended in a lysis buffer containing 50 mM potassium phosphate, pH 7.8, 400 mM NaCl, 100 mM KCl, 10% glycerol, 0.5% Triton X-100, and 10 mM imidazole, and then disrupted in a sonication bath for 5 minutes. Samples were then centrifuged at 18,000 x g for 1 minute. The resulting pellets and supernatants were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 10% polyacrylamide gels (Schagger and von Jagow, 1987), transferred to a nitrocellulose membrane by blotting at 0.4 Amp for 90 minutes, and probed by Western blot analysis (Towbin *et al.*, 1979) using a monoclonal antibody against penta-histidine epitope (Qiagen).

MdfA was also expressed in a cell-free system from the pCOG2 plasmid using the method previously described by Torizawa and co-authors (Torizawa *et al.*, 2004). The cell-free protein synthesis reactions were carried out in 100 μL samples, with or without the presence of 100 μg liposomes, containing from 2 to 8 μg of the pCOG2 plasmid. Expression was performed at 37°C for 120 minutes. *MdfA* expression levels were analyzed by SDS-PAGE followed by Western blot analysis.

3.3. Gel Electrophoresis of Protein

All SDS-PAGE was carried out using 10% polyacrylamide gels with or without the addition of 6 M urea (Schagger and von Jagow, 1987). Western blot analysis was carried out as

explained by Towbin and co-authors (Towbin *et al.*, 1979). Protein samples were transferred from 10% polyacrylamide gels to nitrocellulose membrane by blotting at 0.4 Amp for 90 minutes. The nitrocellulose membrane was blocked using blocking reagent (Qiagen), and detection was accomplished using an anti-penta-histidine-HRP-conjugate antibody (1:10000 dilution). Chemo luminescence was achieved through incubation with 100 mM Tris, pH 8.5 containing 1.25 mM luminol, 0.225 mM d-cumaric acid, and 0.01% hydrogen peroxide.

3.4. Preparation of Cell Membranes for MdfA Extraction.

LMG194 cells containing the pCOG3 plasmid were grown in LB at 37^{o}C on a shaker platform at 200 r.p.m. to an optical density (OD_{600}) of 0.4. (L)-arabinose was added to a final concentration of 0.002% and the cultures were incubated for an additional 4 hours. The cells were then pelleted by means of centrifugation (6,000 x g for 15 minutes at 4^{o}C). The cells were lysed via French press at 18,000 p.s.i. in a buffer containing 50 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$ 1 mM dithiothreitol, 10% glycerol (v/v) and 1 mM phenylmethylsulfonyl fluoride (TMDG buffer). Cell debris was removed by centrifugation (7,700 x g for 15 minutes) and then the membranes were purified by centrifugation of the low-speed supernatant at 100,000 x g for 75 minutes. The pellet was then resuspended in TMDG buffer and centrifuged under the same conditions. The membrane pellet was resuspended once more in TMDG buffer and frozen in liquid nitrogen.

3.5. MdfA Detergent Extraction from Cell Membranes

MdfA extraction from the membrane was attempted using the following five detergents: Triton X100 (BioRad, Mississauga, Ontario), octylglucoside (Fisher BioReagents, Ottawa, Ontario), DHPC (Avanti, Alabaster, Alabama), DDM and LMPG (Avanti) at concentrations ranging from 0.5% to 2.0% (w/v). The membrane samples were diluted to 10 mg protein /mL with TMDG buffer, mixed with the detergent at the appropriate final concentration and sonicated in a bath sonicator for 5 minutes. Samples were then ultracentrifuged (220,000 x g for 20 minutes) to separate soluble and insoluble protein. The soluble and insoluble fractions were analyzed by SDS-PAGE, followed by Western blot analysis.

3.6. MdfA Purification

MdfA was extracted from the cell membrane using DHPC at a final concentration of 2.0%. Membrane samples were diluted to a protein concentration of 10 mg/mL and incubated with detergent, followed by gentle agitation for 30 minutes at 4^{o}C. Samples were ultracentrifuged (100,000 x g for 60 minutes) to pellet the insoluble fraction. The soluble fraction was loaded onto a column containing nickel nitrilotriacetic acid agarose (Ni-NTA agarose) (Qiagen). The column was washed with TMDG buffer containing 0.4% DHPC and 20 mM imidazole to remove contaminating proteins. MdfA was eluted in a TMDG buffer containing 0.4% DHPC and 250 mM imidazole. Purification through extraction with DDM was performed using the same procedure, except that DDM was substituted for DHPC.

For NMR sample preparation, MdfA was transferred into LMPG micelles. Briefly, MdfA was extracted from the membrane and loaded onto a Ni-NTA column as described above. The column was washed with TMDG containing 0.4% DHPC and 20 mM imidazole and then with TMDG containing 0.2% LMPG and 20 mM imidazole, effectively replacing DHPC micelles with LMPG micelles. MdfA was eluted with TMDG containing 0.2% LMPG and 250 mM imidazole. Due to the high cost of LMPG, it was not used in the extraction procedure.

3.7. Isotopic Labeling of MdfA for NMR Studies

In order to isotopically label MdfA, LMG194 cells containing the pCOG3 plasmid were grown on LB at 37^{o}C on a shaker platform at 200 r.p.m. to an optical density (OD_{600}) of 0.4 then pelleted by centrifugation (6,000 x g for 15 minutes at 4^{o}C). Cells were resuspended in M63, RM, or *E. coli* OD2 medium (Silantes, München Germany), supplemented with isotopically labeled compounds as described in detail below. (L)-arabinose was added to a final concentration of 0.002% and cultures were incubated for an additional 4 hours. A ^{15}N- labeled MdfA sample was generated by inducing *mdfA* expression in M63 medium containing 15 mM ^{15}N ammonium chloride (Cambridge Isotope Labs, Andover, Massachusetts) as the sole source of nitrogen. A ^{13}C, ^{15}N, ^{2}H- labeled MdfA sample was generated inducing *mdfA* expression in M63 medium formulated with 100% deuterium oxide (Cambridge Isotope Labs), containing 15 mM ^{15}N ammonium chloride as the sole source of nitrogen and 0.5% (v/v) ^{13}C glycerol (Cambridge Isotope Labs) as the sole source of carbon. A ^{13}C-Glu, ^{15}N-Phe-labeled MdfA sample was generated by inducing *mdfA* expression in M63 medium supplemented with 0.25

mM ^{13}C Glu and 0.25 mM ^{15}N Phe. Fluoro-tryptophan labeled MdfA was generated by expressing m*dfA* in RM medium supplemented with 1 mM 5-fluoro-tryptophan. A ^{13}C, ^{15}N-double-labeled MdfA sample was made by expressing *mdfA* in *E.coli* OD2 ^{13}C, ^{15}N-labeled medium. MdfA was extracted with DHPC and purified as described above into LMPG micelles.

3.8. Dynamic Light Scattering

Dynamic light scattering experiments were conducted at the Saskatchewan Structural Sciences Center (SSSC) using the DynaPro/Micro Sampler with Dynamics Software package (Dynamics ™ version 5.26.60). Dynamic light scattering was performed on MdfA purified using DHPC, DDM, and LMPG at protein concentrations ranging from 2 mg/mL to 7 mg/mL.

3.9. MdfA Stability Optimization

The stability of purified MdfA was assessed by measuring the percentage of total protein that remained soluble after incubation at 27^{o}C for periods of up to 72 hours. Stability of MdfA was tested in a 25 mM sodium phosphate buffer containing 0.4% DHPC or 0.2% LMPG, at pH 5, 6, 7 or 8, with or without 100 mM NaCl. Buffer exchange was achieved by diluting 1:10 purified MdfA in an appropriate buffer and concentrating it to the starting volume, using an Amicon Ultra protein concentrator (Millipore, Billerica, Massachusetts) with a 10,000 Da molecular weight cutoff. This procedure was repeated three times. After buffer exchange, samples were incubated at 27^{o}C for 0, 48, or 72 hours. At those times, aliquots were taken, pelleted in a desktop centrifuge at 18,000 x g for 3 minutes, and the protein concentration of the supernatant was determined using the Lowry protein assay (Lowry *et al.,* 1951).

3.10. Nuclear Magnetic Resonance Experiments

NMR spectroscopy on ^{15}N-, ^{15}N, ^{13}C, ^{2}H-, ^{15}N, ^{13}C- and ^{13}C-Glu, ^{15}N-Phe- labeled MdfA were performed by Dr. Oleg Dmitriev either at the SSSC using a 600 MHz Bruker spectrometer equipped with a Cryoprobe, or at the National Magnetic Resonance Facility in Madison, Wisconsin (NMRFAM) on a 750 MHz Bruker spectrometer with a Cryoprobe. ^{19}F NMR experiments were conducted at the SSSC on a 500 MHz Bruker spectrometer equipped with a ^{1}H, ^{19}F-probe at 315K. The samples contained between 0.133-0.266 mM protein, 25 mM

sodium phosphate buffer, pH 6.0, 5% (v/v) D_2O, 0.2% LMPG, 50 mM NaCl, and 0.5 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid for chemical shift referencing.

3.11. Reconstitution of MdfA into Proteoliposomes

Liposomes were prepared from an *E. coli* total lipid extract in chloroform (Avanti). The organic solvent was evaporated to dryness under a gentle stream of argon, and the lipid was resuspended in a solution containing 100 mM potassium phosphate, pH 7.0, 2 mM β-mercaptoethanol, and 1.5% octylglucoside to a final lipid concentration of 25 mg/mL. Liposomes were dialyzed against 100 mM potassium phosphate, pH 7.0 and 2 mM β-mercaptoethanol using a 12,000-14,000 Da molecular weight cutoff SpectraPor dialysis tubing (Spectrum Laboratories Inc., Rancho Domingeuz, California).

Preformed liposomes were passed through a mini-extruder (Avanti Polar Lipids) using a Whatman polycarbonate membrane (Ge Healthcare, Piscataway, New Jersey) with a 100 nm pore size to ensure uniform size distribution. Liposomes were resuspended in a 300 μL solution containing 100 mM potassium phosphate at pH 7.0, 2 mM β-mercaptoethanol, and 2% Triton-X 100 at 12.5 mg/mL lipid concentration, mixed with 60 μg of purified MdfA and incubated at 30^{o}C for 15 minutes. Detergents were removed by incubating liposomes at 30^{o}C with several changes of Bio-beads (BioRad) for a total time of 150 minutes, using a total of 120 mg Bio-beads per 300 μL liposome suspension. Bio-beads were pre-treated as follows: 10 g of Bio-beads were added to 70 mL methanol and gently stirred for 15 minutes. The beads were washed with 700 mL ddH_2O and then transferred to a vacuum flask and degassed for 6 hours, changing the ddH_2O every 2 hours.

3.12. MdfA Activity Assay by ACMA Fluorescence

MdfA is an H^+/Drug antiporter, and therefore drug transport can be indirectly monitored by measuring the H^+ transport coupled to drug translocation. A transport assay using the fluorophore 9-amino-6-chloro-2-methoxyacridine (ACMA) as a reporter of H^+ transport has been developed. The assay principle is illustrated in Figure 3.5. Liposomes were loaded with K^+, while the assay buffer contained Na^+. Valinomycin was added to the assay solution, allowing K^+ to diffuse down its concentration gradient out of the liposome, creating an electric potential difference across the membrane ($\Delta\psi$), positive outside. Addition of carbonyl cyanide

4-(trifluoromethoxy)-phenylhydrazone (FCCP) allowed flow of protons down the electrochemical gradient into the liposome, lowering its internal pH. ACMA can pass freely through the lipid bilayer in the electroneutral form, but not in the positively charged protonated form. ACMA added to the solution became protonated upon entering the liposome and accumulated inside, causing concentration-dependent fluorescence quenching. Addition of an MdfA substrate, such as ethidium bromide, to the solution resulted in the pumping of the substrate into the liposome in exchange for H^+, increasing internal pH, which in turn increased concentration of the deprotonated ACMA inside the liposome. The uncharged form of ACMA diffused out of the liposome, diluting in the external medium, which resulted in the increase of the ACMA fluorescence.

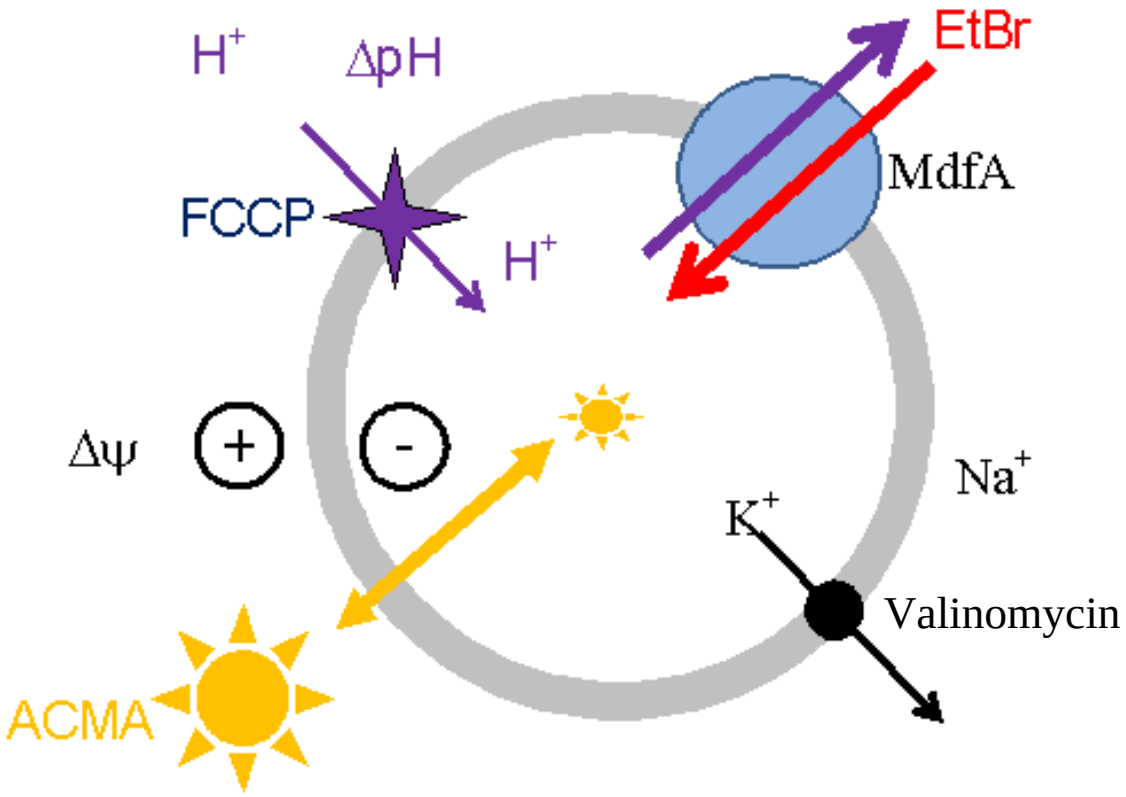


Figure 3.5. The principle of the fluorescence assay of MdfA activity. The addition of valinomycin and FCCP allow for the diffusion out of K^+ and the influx of H^+ down the electrochemical gradient ($\Delta\mu H^+$). The lowering of liposomal pH results in the sequestering of ACMA, causing concentration-dependent fluorescence quenching. Ethidium bromide (EtBr) is added and is pumped into the liposome in exchange for H^+, allowing ACMA to leave the liposome, and increasing fluorescence intensity.

The MdfA activity assay detected by ACMA fluorescence was carried out in a 10 mM Tricine buffer at pH 7.0, containing 5 mM magnesium sulfate and 200 mM sodium sulfate. Fluorescence excitation was set to 410 nm and emission at 490 nm. ACMA was added to a final concentration of 1 μM. A total of 250 μg of proteoliposomes by lipid content was added to the 2 mL assay solution followed by addition of valinomycin to a final concentration of 0.0125 μM. To achieve ACMA quenching, FCCP was added to a final concentration of 0.025 μM. Either chloramphenicol, or ethidium bromide was added to final concentrations ranging from 0.25 μM to 25 μM and ACMA fluorescence recovery was monitored.

4. Results

4.1. *MdfA* Expression Optimization

4.1.1. Comparison of *mdfA* Expression Levels from pOD1016 and pCOG3 Plasmids at different Arabinose Concentration.

In order to generate a large enough quantity of MdfA protein suitable for structural studies, the level of *mdfA* expression had to be optimized. We tested *mdfA* expression in the pBAD102 vector, which offers tight control of expression using (L)-arabinose to induce expression and glucose to inhibit background expression prior to induction. The gene encoding MdfA was previously cloned into the pBAD102 vector yielding pOD1016 plasmid (Dmitriev, unpublished data). The pCOG3 plasmid was generated from the pOD1016 plasmid by deleting the genes encoding thioredoxin and the V5 epitope. The thioredoxin gene was deleted to determine if the absence of its expression would affect *mdfA* expression. The V5 epitope was deleted because a commercially-available anti-penta-histidine antibody effectively detected the hexa-histidine tag fused to MdfA for purification purposes. We compared *mdfA* expression from the plasmids pOD1016 and pCOG3 in LMG194 strain under identical conditions (Figure 4.1). There was no significant difference in *mdfA* over-expression between the pCOG3 and pOD1016 plasmids (Figure 4.1). Deletion of the thioredoxin gene did not result in higher *mdfA* expression. Additional bands seen in Figure 4.1 are likely MdfA aggregates as discussed in detail below. Maximum *mdfA* expression levels were achieved upon the addition of (L)-arabinose to the final concentration of 0.0002 %.

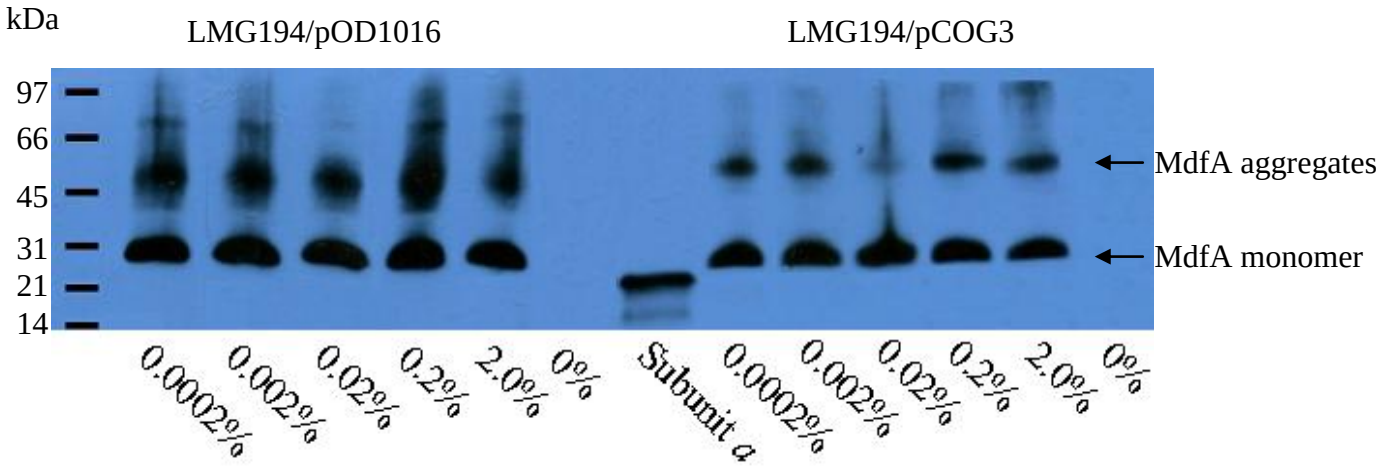


Figure 4.1. *MdfA* expressed from LMG194/pOD1016 and LMG194/pCOG3 cells in a range of (L)-arabinose concentrations. SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". All cell samples were diluted to OD_{600} of 0.7. The 1 mL fractions were then sonicated and centrifuged as described under "Materials and Methods". The pelleted material was resuspended in 1 mL sample buffer, and 20 μL of the pellet samples were loaded into the corresponding lanes. Subunit *a* of ATP synthase containing a hexa-histidine tag was used as a control. The concentrations of (L)-arabinose used for induction are listed below their corresponding lanes.

4.1.2. Comparison of *mdfA* Expression Levels in the LMG 194 and C43 (DE3) Cells

The pCOG3 plasmid was also transformed into C43 (DE3) cells to determine if expression levels of *mdfA* would increase when compared to LMG194. C43 (DE) cells are effective in expressing toxic and membrane proteins (Miroux and Walker, 1996). C43 (DE3) cells metabolize (L)-arabinose, and as such, were expected to require more (L)-arabinose to induce expression compared to a non-(L)-arabinose metabolizing strain such as LMG194. Indeed, maximum *mdfA* expression in C43 (DE3) strain was achieved at (L)-arabinose concentrations of 0.02% and above (Left at panel, Figure 4.2). There was no significant difference in maximum *mdfA* expression levels between the strains C43 (DE3) and LMG194 (Figure 4.2). LMG194/pCOG3 cells were subsequently used for *mdfA* expression for purification purposes because of the smaller requirement of (L)-arabinose. Due to inconsistent expression at 0.0002% (L)-arabinose, further *mdfA* expression for purification purposes was performed using 0.002% (L)-arabinose.

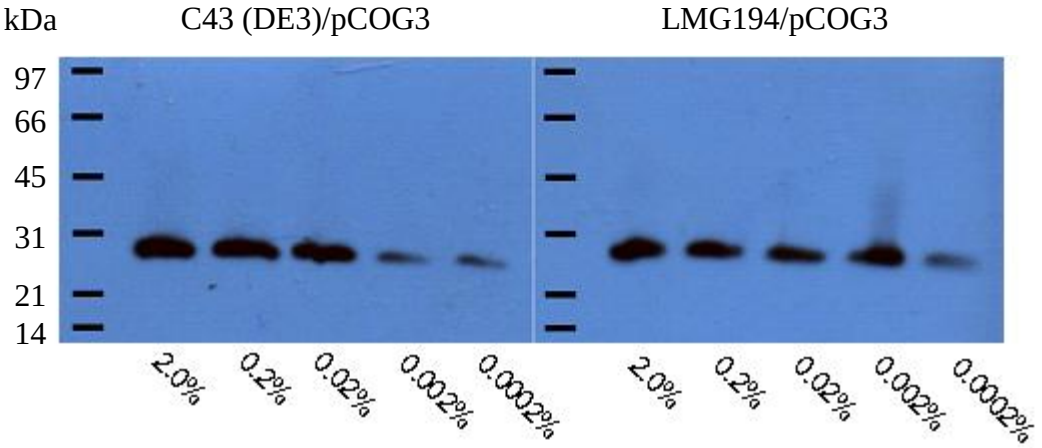


Figure 4.2. *MdfA* expressed from C43 (DE3)/pCOG3 and LMG194/pCOG3 cells in a range of (L)-arabinose concentrations. SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". All cell samples were diluted to OD_{600} of 0.7. The 1 mL fractions were then sonicated and centrifuged as described under "Materials and Methods". The pelleted material was resuspended in 1 mL sample buffer, and 20 μL of the pellet samples were loaded into the corresponding lanes. Subunit *a* of ATP synthase containing a hexa-histidine tag was used as a control. The concentrations of (L)-arabinose used for induction are listed below their corresponding lanes. Lanes 1 and 6: molecular weight standards.

4.1.3. Determination of Optimal *mdfA* Expression Levels on Different Media

Prior to commencing preliminary NMR experiments, it was necessary to determine if *mdfA* could successfully be expressed in a minimal medium. Gene expression in minimal medium allows for simple uniform isotope labeling as well as selective amino acid isotope labeling of the protein.

MdfA expression levels were tested in several different growth media. First, LMG194/pCOG3 cells were grown to mid-logarithmic phase in the LB medium. The cells were pelleted, resuspended and induced in either LB, M63 minimal medium, or, RM. After four hours of induction, the cells were pelleted and the level of *mdfA* expression was then examined by Western blot analysis (Figure 4.3). Visual inspection of Figure 4.3 indicates *mdfA* was expressed slightly higher in RM or M63 minimal medium than in LB. Similar or higher levels of *mdfA* expression in RM and M63 minimal medium made it possible for the generation of both selective and uniform isotope-labeled MdfA samples using these media.

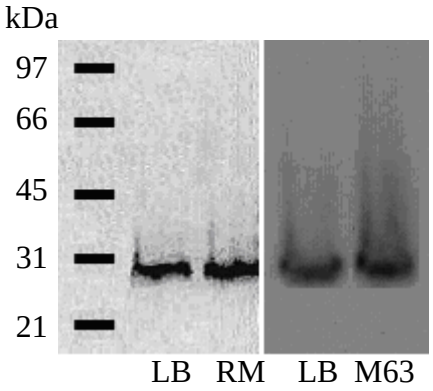


Figure 4.3. *MdfA* expressed by LMG194/pCOG3 in LB, RM, and M63 minimal medium. SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". All cell samples were diluted to OD_{600} of 0.7. The 1 mL fractions were then sonicated and centrifuged as described under "Materials and Methods". The pelleted material was resuspended in 1 mL sample buffer, and 20 μL of the pellet samples were loaded into the corresponding lanes.

4.1.4. MdfA Synthesis in a Cell-free System

Amino acid selective isotopic labeling can significantly facilitate NMR of large proteins. We have tested the feasibility of using cell-free synthesis for amino acid selective isotopic labeling of MdfA for NMR studies.

For cell-free protein expression of *mdfA*, the gene encoding MdfA was cloned into the pEXP-DEST vector, generating the plasmid pCOG2. This plasmid features a T7 promoter, which is essential to prevent transcription of the endogenous DNA present in the cell extract used for cell-free synthesis. Only genes under the control of T7 promoter will be expressed using T7 RNA polymerase added to the reaction mixture. MdfA was synthesized using pCOG2 in a cell-free synthesis system and analyzed by Western blot as shown in Figure 4.4. The amount of expressed *mdfA* (Figure 4.4, *lane 3*), was similar to the amount of subunit *a* of *E. coli* ATP synthase expressed under identical conditions. The yield of subunit *a* in the cell-free synthesis system was about 0.1 mg/mL reaction mixture (Dmitriev and Uhlemann, 2007), which is sufficient to make sample preparation for NMR practical. *MdfA* was also expressed in the presence of liposomes to promote proper folding (Figure 4.4, *lane 5*). Liposomes were prepared as described under section 2.10. The prepared liposomes were passed through a mini-extruder (Avanti Polar Lipids) using a Whatman polycarbonate membrane (Ge Healthcare) with a 100 nm pore size to ensure uniform size distribution prior to use. However, experiments to

determine if MdfA was incorporated into the liposomes were inconclusive and further investigation is needed (data not shown).

4.2. Membrane Incorporation of MdfA

Overproduction of a membrane protein can lead to three possible outcomes. The protein can be degraded upon production, it can be inserted into the membrane, or it can form inclusion bodies in the cytoplasm. It was important that MdfA be inserted into the membrane, because proteins which form inclusion bodies are commonly improperly folded. To determine distribution of over-produced MdfA in the subcellular fractions, a cell culture induced with 0.002% (L)-arabinose was pelleted, passed through French Press to disrupt the cells and centrifuged at a low speed. Low speed centrifugation pelleted much of the cell debris, including inclusion bodies, whereas cellular membranes remained in the supernatant. Most of the MdfA was indeed found in the membrane-containing fraction of the induced cells (Figure 4.5, *lane 4*), and only minute quantities of MdfA were detected in the low speed pellet (Figure 4.5, *lane 3*). Membranes were purified through ultracentrifugation and subsequently used for MdfA extraction.

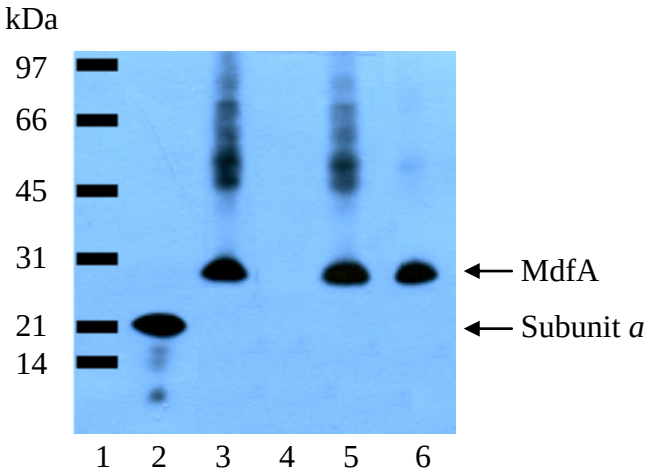


Figure 4.4. MdfA synthesized in the cell-free synthesis system. SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". 5 μL of the cell-free synthesis solution was loaded in each lane. Lane 1: molecular weight standard; lane 2: cell-free synthesized subunit *a* from ATP synthase; lane 3: cell-free synthesized MdfA; Lane 4: cell-free synthesis solution lacking expression plasmid; lane 5: cell-free synthesized MdfA in the presence of 100 μg liposomes; lane 6: 10 μL membrane preparation containing MdfA.

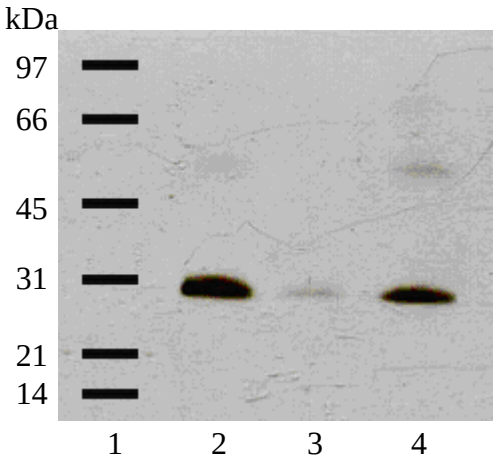


Figure 4.5. Distribution of MdfA between the subcellular fractions. SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". The pellet sample was resuspended in a volume equal to that of the supernatant and then 10 μL of each sample were loaded into each lane. Lane 1: molecular weight standard; lane 2: crude cell extract; lane 3: pellet after centrifugation of the crude cell extract at 7700 x g; lane 4: supernatant after centrifugation of the crude cell extract at 7700 x g.

4.3. Detergent Extraction of MdfA from Cell Membranes

Several detergents, including Triton-X 100, octylglucoside, DHPC, DDM, and LMPG, were tested for MdfA extraction from the lipid bilayer. The detergents were added to membrane samples containing MdfA. Soluble protein was separated from insoluble material by ultracentrifugation. The effectiveness of MdfA extraction by each detergent was evaluated by comparing the quantity of MdfA present in the soluble fractions with the amount in the insoluble fractions. While Triton-X 100 did not solubilize MdfA to any significant extent (Figure 4.6, Triton-X 100), octylglucoside achieved maximum MdfA solublization at a concentration of 2.0% (Figure 4.6, Octylglucoside). DHPC was also found to achieve maximum MdfA solublization at a concentration of 2.0% (Figure 4.6, DHPC) while LMPG reached maximum MdfA solublization at 0.5% (Figure 4.6, LMPG). Lastly, DDM achieved maximum solublization at a concentration of 1.0% (Figure 4.6, DDM). Both LMPG and DDM appear to be the most effective at extracting MdfA from the cell membrane (Figure 4.6).

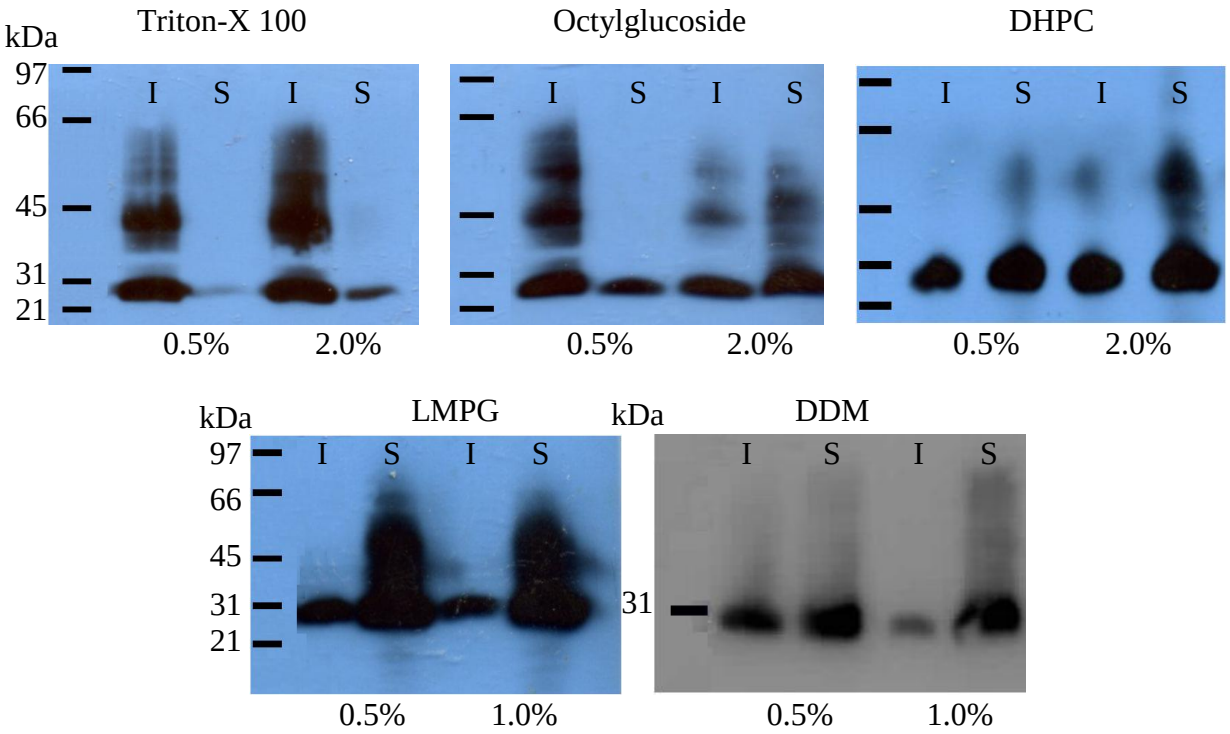


Figure 4.6. The extraction of MdfA from membrane samples using the detergents Triton-X 100, octylglucoside, DHPC, LMPG, and DDM. SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". Each sample, containing 10 mg/mL protein in a volume of 250 μL was centrifuged to separate the soluble and insoluble fractions. The insoluble fraction was resuspended in 250 μL TMDG and 10 μL of either the soluble (S) or insoluble (I) fractions were loaded into each lane. The first lane in each picture contains molecular weight standards. The final concentrations of the detergents used are listed below their corresponding lanes.

4.4. Optimization of MdfA Purification.

In order to determine if MdfA can be purified from the LMPG extract of the cell membrane by means of its C-terminal hexa-histidine tag, we tested binding of the protein to Ni-NTA agarose followed by elution with imidazole (Figure 4.7). MdfA bound to Ni-NTA agarose and was eluted with 100 mM imidazole. Comparison of samples loaded on the column (Figure 4.7, *lane 10*) with fractions eluted with imidazole (Figure 4.7, *lane 7*) shows the majority of MdfA loaded onto the column was recovered after elution. No bound MdfA was detected on the Ni-NTA beads after elution (Figure 4.7, *lane 8*), indicating that MdfA could potentially be purified via its hexa-histidine tag.

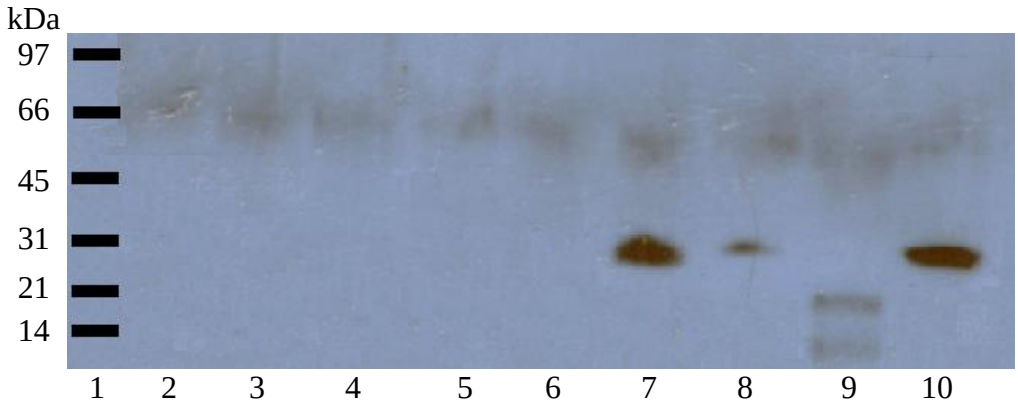


Figure 4.7. Ni-NTA chromatography of MdfA in the presence of 1% LMPG SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". 10 μL from each fraction was loaded on the gel. Lane 1: molecular weight standard; lanes 2-6: column wash samples; lane 7: the eluted fraction; lane 8: bead sample after elution; lane 9: subunit *a* from ATP synthase; lane 10: soluble fraction of the membrane extract loaded onto the column.

LMPG is an ideal detergent for NMR experiments (Krueger-Koplin *et al.,* 2004), but it is expensive. We attempted to purify MdfA using DHPC as an alternative to LMPG. When DHPC was used in the purification of MdfA, results similar to the LMPG purification were obtained (Figure 4.8).

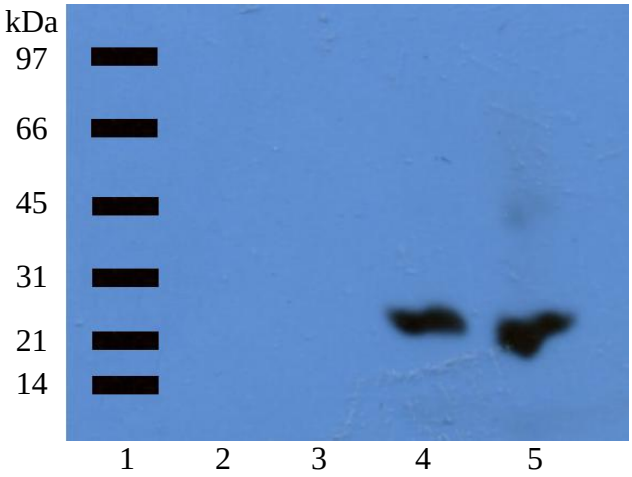


Figure 4.8. Ni-NTA chromatography of MdfA in the presence of 2% DHPC. SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". 10 μL of each fraction was loaded into its corresponding lanes. Lane 1: molecular weight standard; lanes 2-3: column wash; lane 4: bead sample prior to elution; lane 5: the eluted fraction.

Following pilot experiments, a large scale purification of MdfA was performed. MdfA was first purified using solely DHPC detergent, but the purification procedure was subsequently modified to include a detergent exchange stage, replacing DHPC with LMPG. MdfA was extracted from the membrane fraction with DHPC as described under Materials and Methods. Detergent extract was loaded onto a Ni-NTA column and washed with buffer containing 0.4% DHPC and 20 mM imidazole to remove non-specifically bound protein (Figure 4.9, *lane 5*). MdfA binding to Ni-NTA beads was visualized by comparing the crude membrane extract loaded on the column with the column flow-through (Figure 4.9, *lanes 3 and 4*). The MdfA band was present in the former, but was absent in the latter (circled region in Figure 4.9, *lanes 3 and lanes 4*). Thus MdfA remains bound to the column. To improve elution of non-specifically bound protein and to exchange detergent, a second wash was performed with a buffer containing 0.2% LMPG and 20 mM imidazole (Figure 4.9, *lane 6*). MdfA was eluted using a buffer containing 0.2% LMPG and 250 mM imidazole (Figure 4.9, *lanes 7-10*). Large scale purification in this manner yielded an average of 1.0 - 1.4 mg of MdfA per 1 L culture as determined using Lowry protein assay (Lowry *et al.*, 1951). MdfA was also successfully purified using the detergent DDM (Figure 4.10).

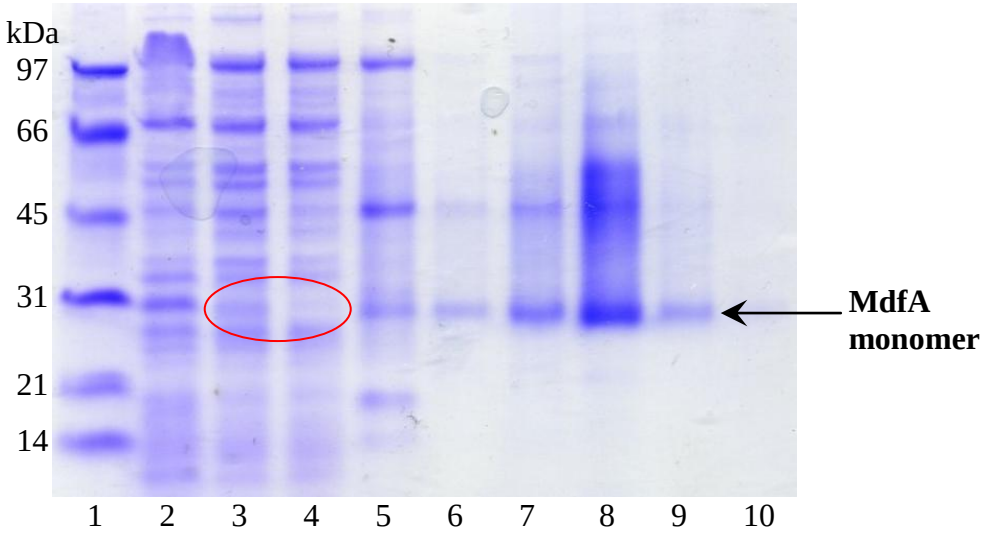


Figure 4.9. Coomassie stained 10% polyacrylamide gel showing the purification of DHPC-extracted MdfA by means of Ni-NTA chromatography. 10 μL of each sample was loaded on the gel. Lane 1: molecular weight standard; lane 2: cell membranes (10 mg/mL protein); lane 3: DHPC extract loaded onto the column; lane 4: flow-through during column loading; lane 5: 20 mM imidazole wash containing 0.4% DHPC; lane 6: 20 mM imidazole wash containing 0.2% LMPG; lanes 7-10: purified MdfA samples eluted by 250 mM imidazole. Bands corresponding to MdfA monomer are shown by the arrow. The area containing monomeric MdfA band in the DHPC extract (lane 3) and column flow-through (lane 4) is circled.

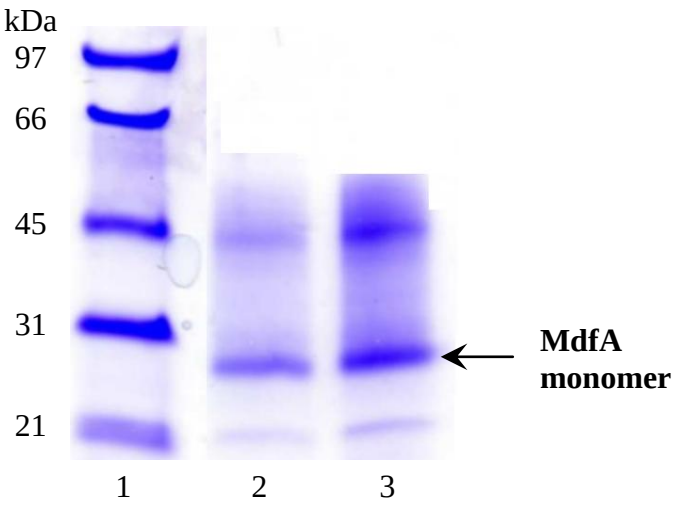


Figure 4.10. Coomassie stained 10% polyacryalmide gel of MdfA purified in DDM detergent micelles. Lane 1: molecular weight standards; lane 2: 4 μg purified MdfA; lane 3: 10 μg purified MdfA.

4.5. Aggregation Propensity of MdfA.

In addition to monomeric MdfA, MdfA purified by Ni-NTA chromatography contained several diffuse bands at higher molecular weights as revealed by SDS-PAGE, followed by either Coomassie staining or silver staining (Figure 4.11). These bands were detectable via Western blot analysis with anti-penta-histidine antibody used to detect MdfA (Figure 4.11). The molecular weights of the corresponding proteins were determined based on band migration distances using molecular weight standards for calibration. Molecular weights corresponding to the four bands detected by Western blot (Figure 4.11 B) were determined to be 28 kDa, 57 kDa, 87 kDa, and 111 kDa (Figure 4.12). These values are consistent with the molecular weights of MdfA monomer, dimer, trimer, and tetramer respectively. MdfA is 47 kDa in size, but it was found to migrate at around 30 kDa during SDS-PAGE. It is not unusual for the molecular weight of a membrane protein to be underestimated from SDS-PAGE migration distance (Miyake *et al*., 1978; Rath *et al.*, 2009). Membrane proteins can bind more SDS per unit weight due to their extreme hydrophobicity and therefore have a larger negative charge and migrate farther through the SDS-PAGE gel than hydrophilic proteins of the same size. Previous studies have also found that MdfA migrates at an apparent molecular weight of around 30 kDa via SDS-PAGE (Lewinson and Bibi, 2001).

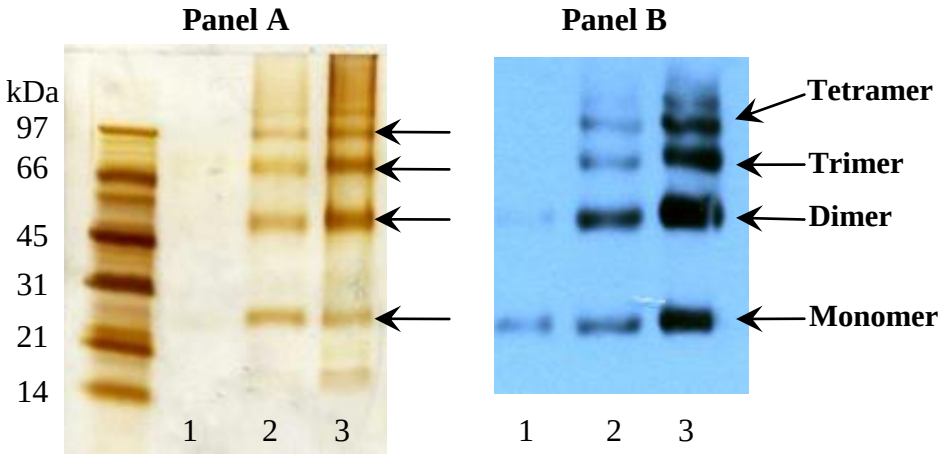
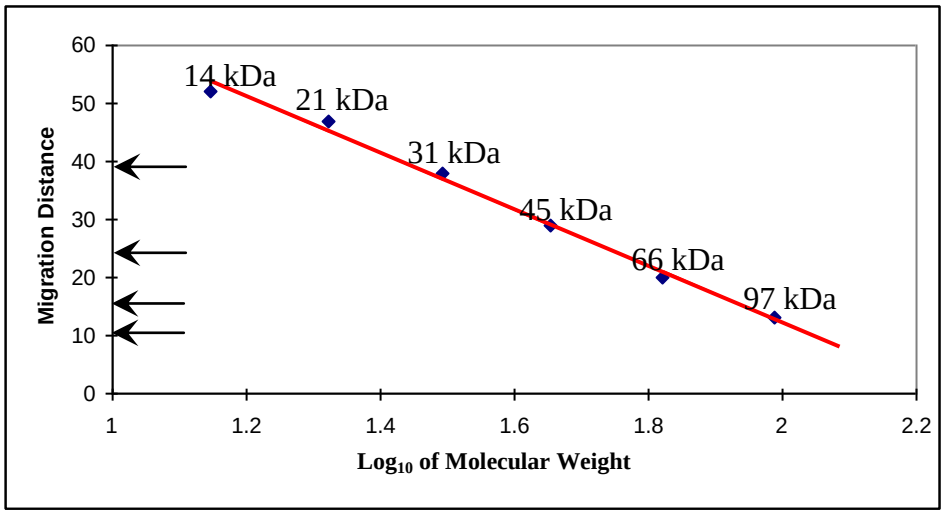


Figure 4.11. Silver-stained gel (A) and Western blot analysis (B) of purified MdfA. SDS-PAGE and subsequent Western blot analysis was performed as described under "Materials and Methods". Lane 1: 0.2 µg MdfA; Lane 2: 2 µg MdfA; Lane 3: 5 µg MdfA. The arrows point to the bands corresponding to MdfA and its oligomers.



Band	Distance, mm	logMW	MW (kD)	*Aggregation number*
1	39	1.449881	**28.18**	Monomer
2	24	1.758432	**57.34**	Dimer
3	15	1.943562	**87.81**	Trimer
4	10	2.046413	**111.28**	Tetramer

Figure 4.12. The molecular weight determination of MdfA oligomers. Migration distance of the protein standards is plotted as a function of the logarithm of molecular weight (◆). Migration distance of the MdfA aggregates is shown by the arrows.

Prior to the realization that the additional bands observed in SDS-PAGE gels were likely MdfA aggregates, all SDS samples were treated at the standard 100°C for 3 minutes. Since membrane protein aggregation in SDS is often temperature-dependent, I investigated the effects of temperature treatment of MdfA SDS samples on band intensity in SDS-PAGE. Strong temperature dependence of the band intensity was observed and provided further evidence that the additional bands correspond to MdfA oligomers (Figure 4.13). Pure MdfA was divided into four aliquots, which were either incubated at 37°C, 75°C, 100°C, or autoclaved prior to analysis via SDS-PAGE. An intense band corresponding to MdfA monomer was observed in the sample treated at 37°C (Figure 4.13, panel A and B, row 4). The relative intensity of additional bands corresponding to MdfA oligomers increased dramatically when the sample was treated at higher temperatures. If the additional bands corresponded to proteins other than MdfA, varying the temperature should not have significantly affected the resolution pattern seen in the SDS-PAGE gels. Since the additional bands are MdfA oligomers, the purified MdfA samples are in fact pure, and contain very little contaminating protein.

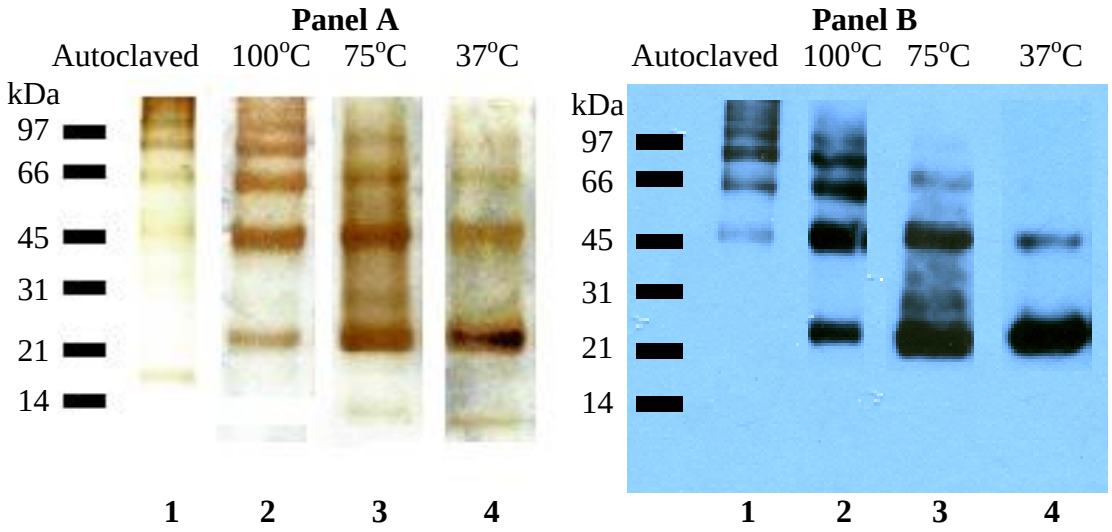


Figure 4.13. Silver-stained gel (A) and Western blot analysis (B) of purified MdfA samples incubated at different temperatures prior to SDS-PAGE. Samples were loaded onto a 10% SDS-PAGE gel containing 6 M urea. Western blot analysis was performed as described under "Materials and Methods". Lanes 1-4: 2 μg MdfA. Temperature treatments are shown above the lane.

4.6. Analysis of MdfA Aggregation State by Dynamic Light Scattering

The yield of 1.0 - 1.4 mg pure MdfA per 1 L culture was sufficient to initiate crystallization trials. Monodisperse protein samples are generally required for successful crystallization. Dynamic light scattering experiments showed that the purified MdfA was monodisperse in DHPC (Figure 4.14), DDM (Figure 4.15), and LMPG detergent micelles (Figure 4.16) up to 4 mg/mL protein concentration and at a range of temperatures up to 30°C. Nearly all the particles present in the solution were found to be around 3-4 nm in size, which is around the size expected for a globular protein of 50 kDa (Dynamics Software package, version 5.26.60). MdfA however, is not a globular protein, but a 3-4 nm diameter is a reasonable estimate for MdfA in a detergent micelle.

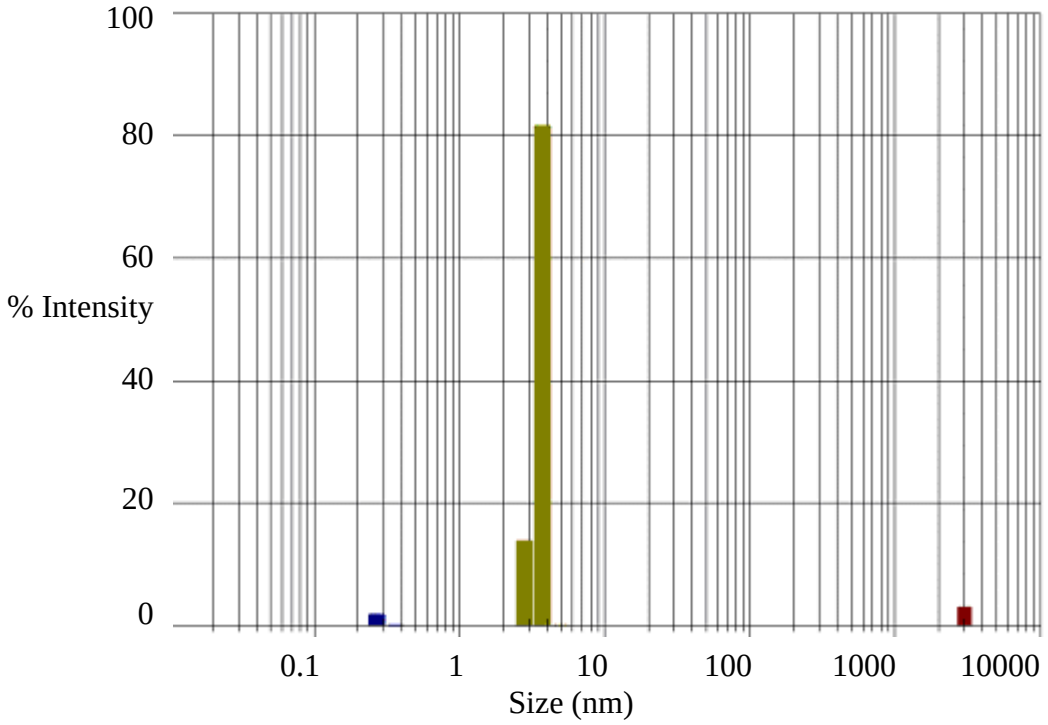


Figure 4.14. Dynamic Light Scattering analysis on MdfA. A purified MdfA sample at a protein concentration of 4 mg/mL in 0.4% DHPC was tested at 25°C using the DynaPro/Micro sampler at the SSSC.

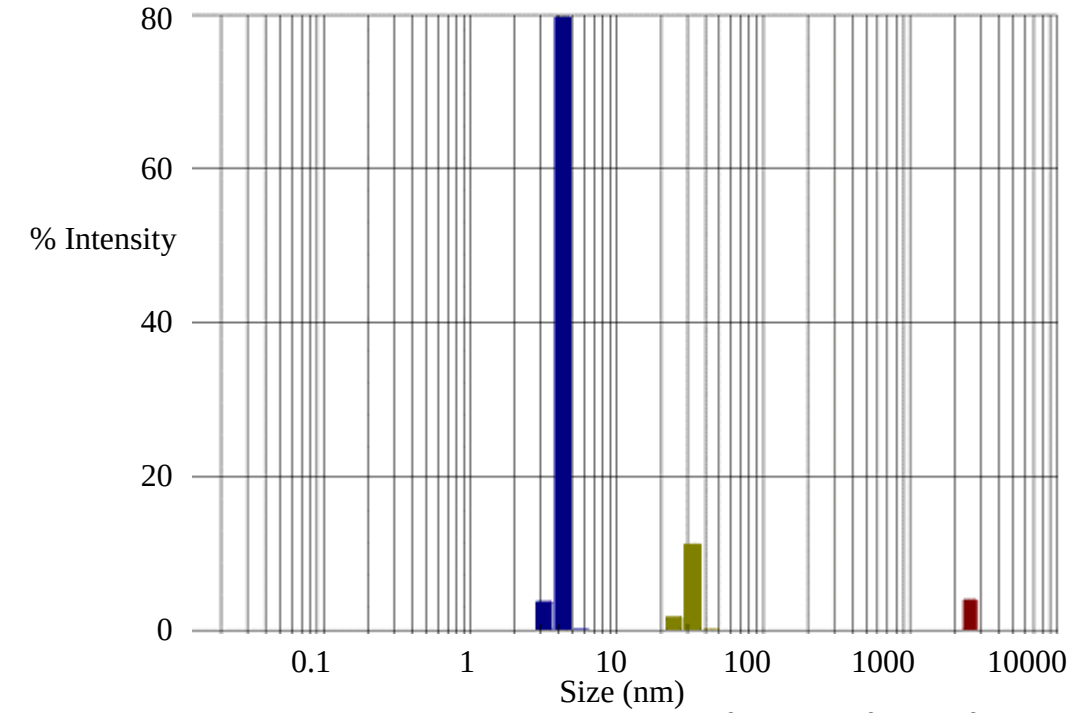


Figure 4.15. Dynamic Light Scattering analysis on MdfA. A purified MdfA sample at a protein concentration of 4 mg/mL in 0.4% DDM was tested at 25°C using the DynaPro/Micro sampler at the SSSC.

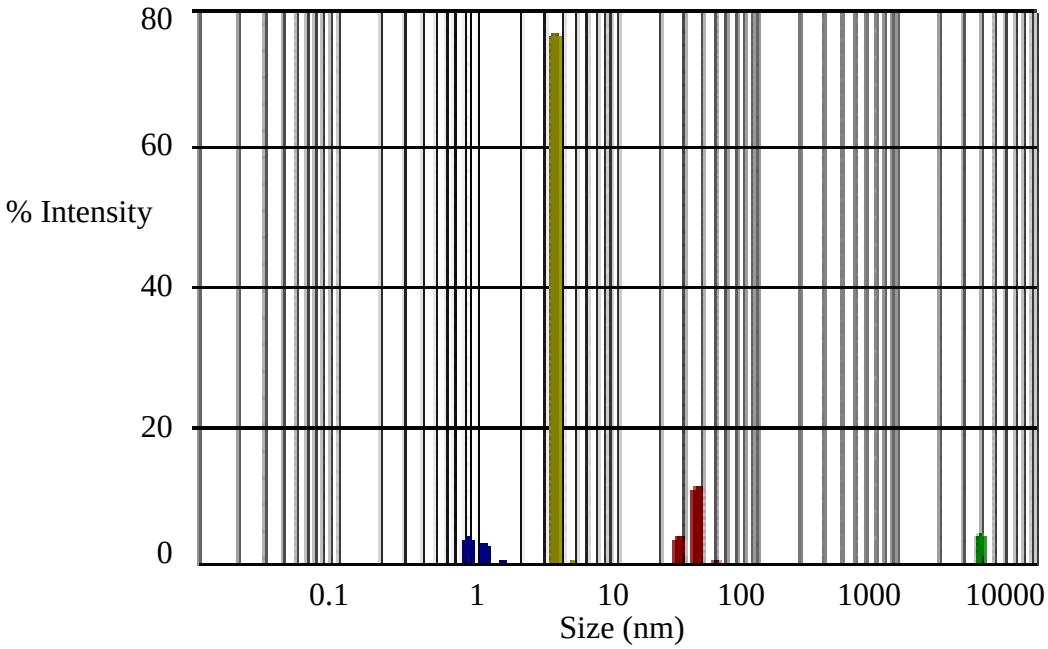


Figure 4.16. Dynamic Light Scattering analysis on MdfA. A purified MdfA sample at a protein concentration of 4 mg/mL in 0.2% LMPG was tested at 25°C using the DynaPro/Micro sampler at the SSSC.

4.7. MdfA Stability Optimization

In order to conduct crystallization trials or NMR experiments on MdfA, it was necessary to ensure that MdfA was stable for long periods of time as crystallization trials and NMR experiments can last for several weeks, or even several months. Initially, MdfA in solution was only stable for two days at 27^{o}C. In an attempt to increase the stability of MdfA in solution, different buffer conditions were investigated (see Materials and Methods for details). MdfA stability was measured in DHPC and LMPG detergent micelles, at pH 5, 6, 7, and 8, and in the presence or absence of 100 mM NaCl (Figure 4.17). MdfA was most stable in LMPG detergent micelles at pH 5-6 in the presence of 100 mM NaCl. Under these conditions, nearly 100% of the protein remained soluble after 72 hours at 27^{o}C. Further visual observation revealed that it remained soluble for several more days at 27^{o}C and for two months at room temperature.

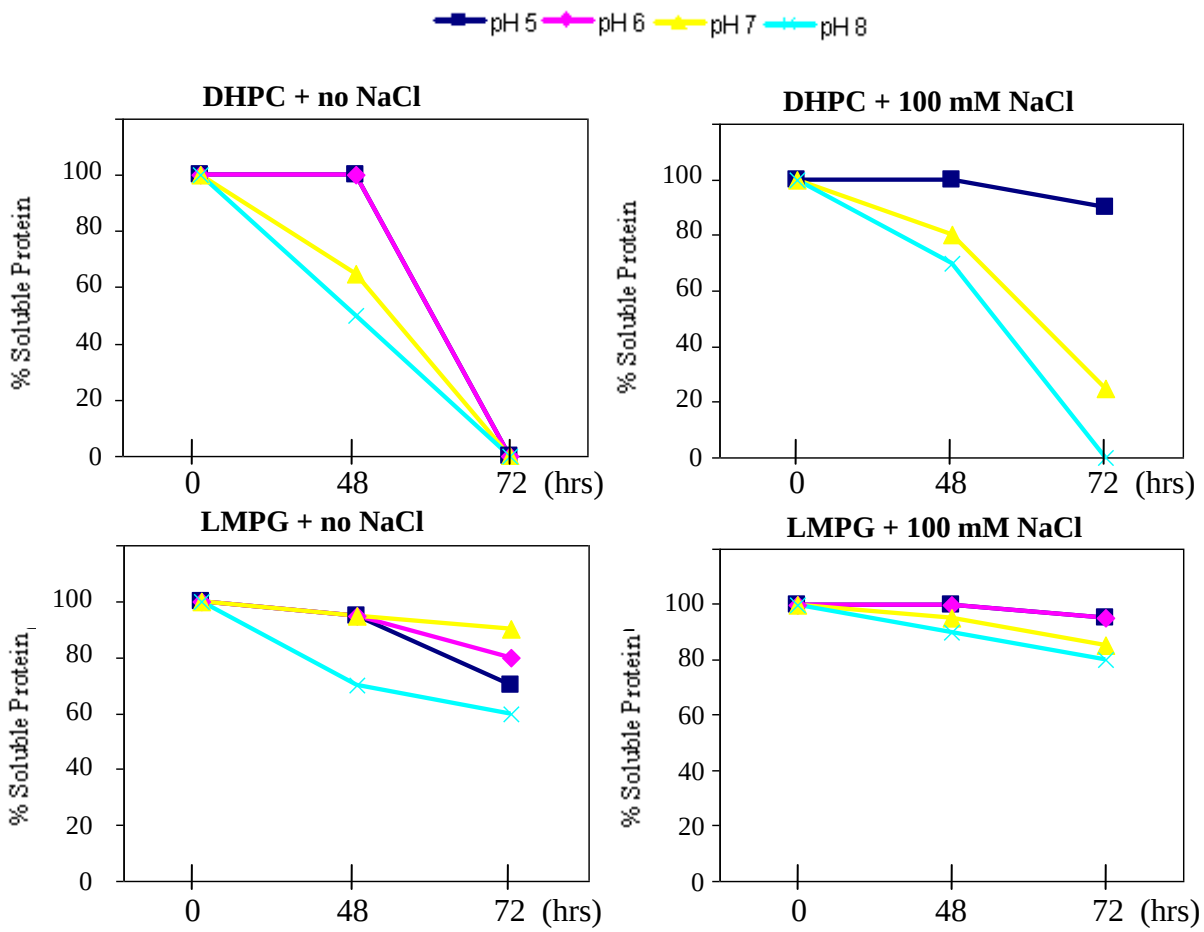


Figure 4.17. Stability of the purified MdfA protein in a 25 mM sodium phosphate buffer containing either 0.2% LMPG or 0.4% DHPC, at pH 5-8, with or without 100 mM sodium chloride.

4.8. Preliminary MdfA Crystallization Trials

An MdfA sample purified in LMPG detergent micelles was sent to the High-Throughput Crystallization Lab at the Hauptman-Woodward Institute for Crystal Screening. The screens resulted in several conditions (Table 4.1) that yielded possible protein crystals (Figure 4.18). Unfortunately, upon reproduction of the conditions, the crystals that were found were not made up of protein. Further screening was conducted using the MPD Suite (Qiagen), which resulted in two sets of conditions that formed crystals (Figure 4.19). Initial testing by Ponceau S staining indicates that these crystals may be protein and experiments to grow larger crystals are underway.

Table 4.1. List of crystallization conditions from the high-throughput crystallization screens which were chosen for reproduction and further investigation.

Crystal Screen Composition
0.05 M $CsCl_2$, 0.1 M MES monohydrate pH 6.5, 30% (v/v) Jeffamine M-600
3.82 M $MnCl_2$, 0.1 M sodium citrate, pH 4.2
2.13 M $(NH_4)_2HPO_4$, 0.1 M HEPES, pH 7.5
3.73 M LiBr, 0.1 M Bis-Tris Propane, pH 7
1.59 M $MgSO_4*7H_2O$, 0.1 M TAPS, pH 9
0.39 $Zn(C_2H_3O_2)_2$, 0.1 M sodium acetate, pH 5
0.1 M NH_4NO_3 0.1 M TAPS, pH 9, 20% (w/v) PEG 20000
0.1 M $Zn(C_2H_3O_2)_2$, 0.1 M sodium acetate, pH 5, 20% (w/v) PEG 20000
0.1 M LiBr, 0.1 M CAPS, pH 10, 40% (w/v) PEG 20000
0.1 M $Mg(C_2H_3O_2)_2*4H_2O$, 0.1 M TAPS, pH 9, 40% (w/v) PEG 20000
0.1 M K_2HPO_4, 0.1 M Bis-Tris Propane, pH 7, 20% (w/v) PEG 8000
0.1 M K_2HPO_4, 0.1 M Tris, pH 8, 20% (w/v) PEG 8000
0.1 M $Mg(NO_3)_2*6H_2O$, 0.1 M TAPS, pH 9, 20% (w/v) PEG 8000
0.1 M LiCl, 0.1 M Tris pH 8, 20% (w/v) PEG 4000
0.1 M $MgSO_4*7H_2O$, 0.1 M TAPS, pH 9, 20% (w/v) PEG 4000
0.1 M $MnCl_2$, 0.1 M Bis-Tris Propane pH 7, 20% (w/v) PEG 4000
0.1 M $MgCl_2*6H_2O$, 0.1 M TAPS, pH 9, 20% (w/v) PEG 1000
0.1 M $MnCl_2$, 0.1 M sodium acetate, pH 5, 20% (w/v) PEG 1000
0.1 M $Mg(C_2H_3O_2)_2*4H_2O$, 0.1 M sodium acetate, pH 5, 40% (w/v) PEG 400
0.1 M $MgSO_4*7H_20$, 0.1 M MES, pH 6, 40% (w/v) PEG 400
0.1 M $Zn(C_2H_3O_2)_2$, 0.1 M sodium acetate, pH 5, 40% (w/v) PEG 400
0.1 M $Ca(C_2H_3O_2)_2$, 0.1 M sodium acetate, pH 5, 20% (w/v) PEG 400
0.8 M Lithium sulfate monohydrate, 0.1 M Tris, pH 8.5

Panel A **Panel B**

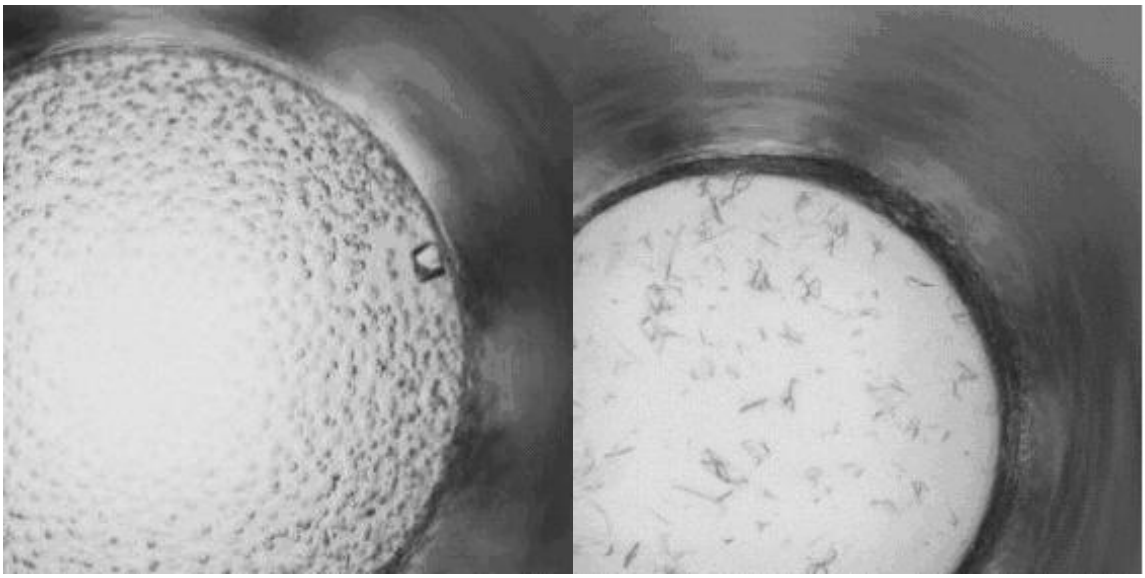


Figure 4.18. Examples of crystallization conditions chosen from the high-throughput crystallization trials for further investigation. (A) A sample containing 0.05 M $CsCl_2$, 0.1 M MES monohydrate pH 6.5 and 30% (v/v) Jeffamine M-600; (B) sample containing 3.82 M $MnCl_2$ and 0.1 M sodium citrate, pH 4.2.

Panel A **Panel B**

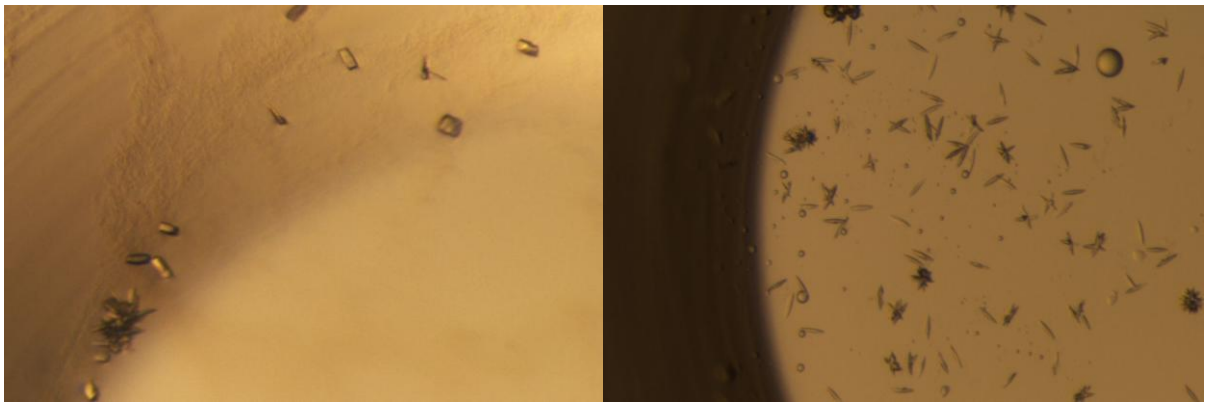


Figure 4.19. Photographs of Ponceau S stainable crystals found in MdfA crystal screens. (A) A sample containing 40% 2-methyl-2,4-pentanediol, and 0.2 M $CaCl_2$; (B) sample containing 40% 2-methyl-2,4-pentanediol and 0.2 M ammonium nitrate. Photographs courtesy of Carla Protsko.

4.9. Preliminary MdfA Nuclear Magnetic Resonance Experiments

After stability optimization, several MdfA samples with different isotope labeling patterns were generated for NMR studies. The samples included uniform ^{15}N-labeled, uniform ^{13}C, ^{15}N, ^{2}H-labeled, ^{13}C-Glu, ^{15}N-Phe-labeled, and 5-fluoro-tryptophan-labeled proteins. ^{15}N- and ^{13}C, ^{15}N, ^{2}H-labeling can be used to solve the three-dimensional protein structures. ^{13}C-Glu, ^{15}N-Phe-labeling and 5-fluoro-tryptophan-labeling can be used in substrate binding studies. It may also be possible to use 5-fluoro-tryptophan-labeling for NMR sample optimization.

4.9.1. Substrate Binding to MdfA Measured by ^{19}F NMR

^{19}F NMR chemical shift perturbation experiments were carried out to determine if they could be used to detect whether purified MdfA binds its substrates or not. 5-fluoro-tryptophan labeling was chosen as reporter groups because the substrate binding site of MdfA is believed to be hydrophobic and the 10 hydrophobic tryptophan residue are evenly spaced throughout the transmembrane region of the MdfA sequence, where it is likely that some of the tryptophan residues will be located in the substrate binding site and will register substrate binding in the chemical shift perturbation assays. Since there are 10 tryptophan residues in MdfA, the ^{19}F NMR spectrum of a 5-fluoro-tryptophan-labeled MdfA sample should therefore consist of ten peaks. MdfA labeled with 5-fluoro-tryptophan (Figure 4.20) resulted in a spectrum exhibiting several overlapping wide peaks (Fig 4.21). The overlapping peaks and wide spectral lines are probably due to fast relaxation of magnetization of the protein during the NMR experiment. The peaks in the ^{19}F NMR spectra that have a signal-to-noise ratio greater than two were very reproducible, being observed in all successful ^{19}F NMR experiments. Upon addition of ethidium bromide, a substrate of MdfA, minor spectral changes were observed, suggesting that MdfA bound ethidium bromide (Figure 4.22).

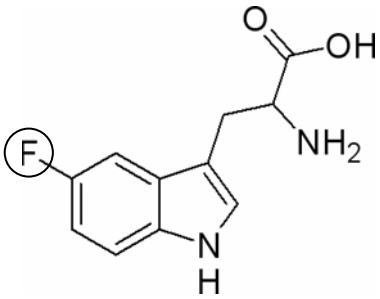


Figure 4.20. The structure of 5-fluoro-tryptophan. The fluorine is encircled.

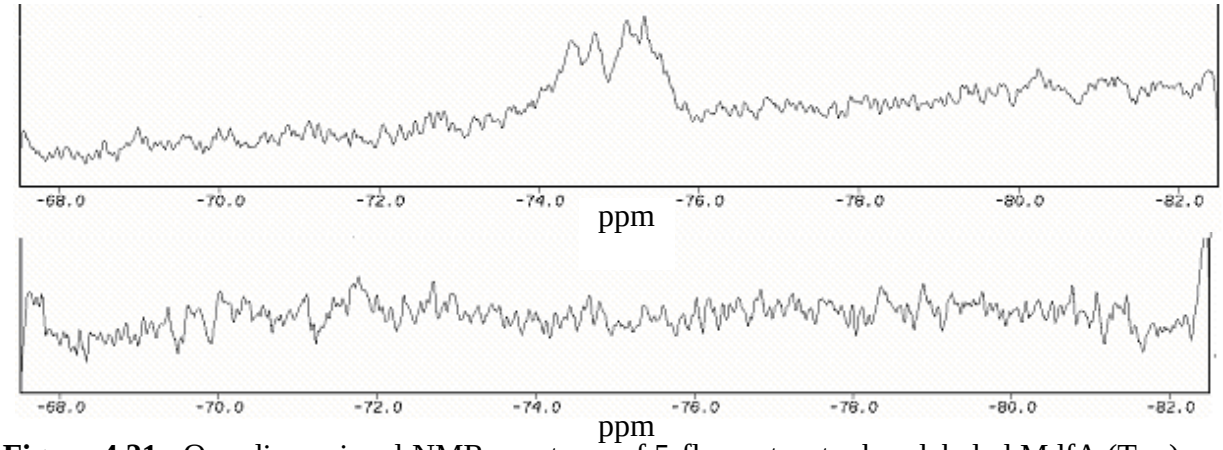


Figure 4.21. One-dimensional NMR spectrum of 5-fluoro-tryptophan-labeled MdfA (Top) and the one-dimensional NMR spectrum of the buffer solution (Bottom).

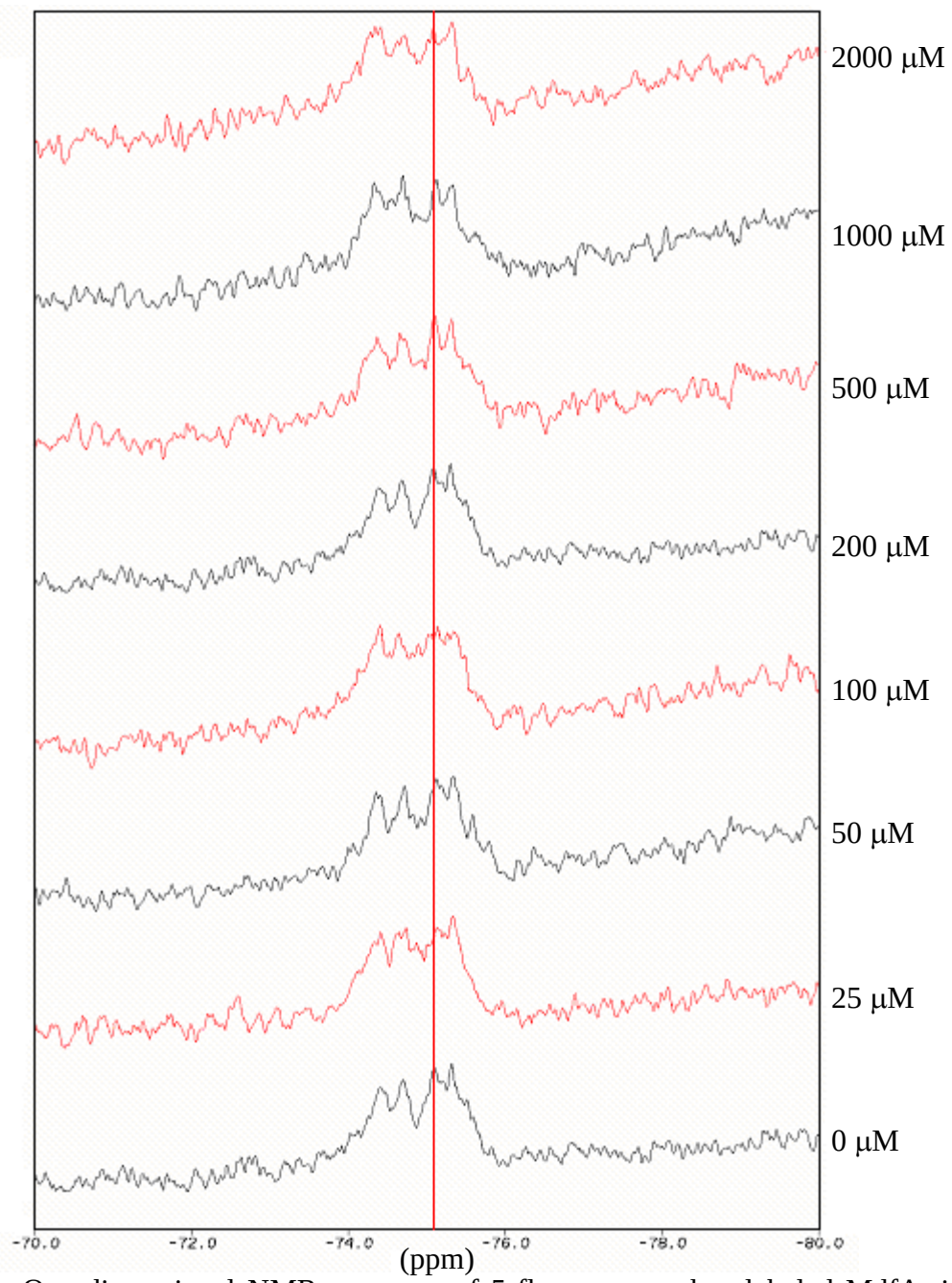


Figure 4.22. One dimensional NMR spectrum of 5-fluoro-tryptophan-labeled MdfA titrated with EtBr. The EtBr concentrations in the NMR samples are labeled beside the corresponding spectra. The vertical red line is added for reference.

The ^{19}F NMR experiment was repeated using another MdfA substrate, chloramphenicol. Upon addition of chloramphenicol, a minor but detectable chemical shift change was observed for one of the peaks (Figure 4.23), which indicated protein-substrate interaction. The observation of chemical shift perturbations upon the addition of either ethidium bromide or chloramphenicol is indicative of properly folded MdfA able to bind its substrates. Upon optimization of NMR conditions, it may be possible to calculate binding constants for these MdfA substrates using ^{19}F NMR.

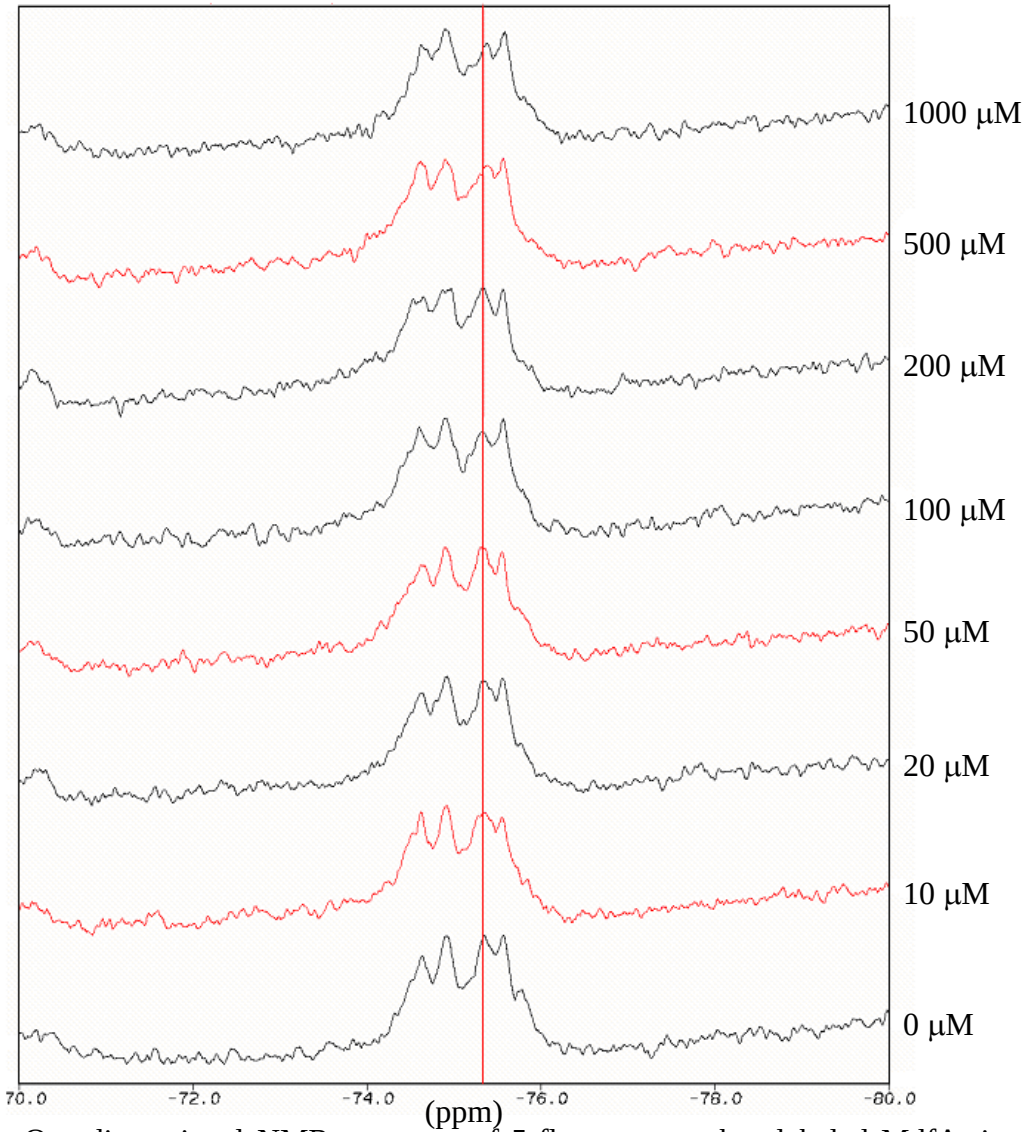


Figure 4.23. One-dimensional NMR spectrum of 5-fluoro-tryptophan-labeled MdfA titrated with chloramphenicol. The chloramphenicol concentrations in the NMR samples are labeled beside the corresponding spectra. The vertical red line marks the position of the peak, which shifts at higher concentrations of chloramphenicol.

4.9.2. Preliminary NMR Studies of MdfA using ^{15}N- and ^{15}N, ^{13}C, ^{2}H- isotope labeling.

Preliminary ^{1}H, ^{15}N-heteronuclear single quantum coherence (HSQC) NMR resulted in poor NMR spectra (Figure 4.24). The chemical shifts for backbone amides and side-chains of Asn, Gln, Arg and Trp are observed in the ^{1}H, ^{15}N-HSQC NMR 2D spectrum (Figure 4.24). Fast relaxation of magnetization lead to weak signal and overlap, resulting in poor resolution of backbone amide signals. Poor dispersion of amide signals indicates that the majority of MdfA exists in an α helical conformation, consistent with secondary structure prediction (Alder and Bibi, 2002). Comparison of the 1D slices taken through the backbone amide region (Figure 4.25 (A)) and through the Asn and Gln side-chain region (Figure 4.25 (B)) respectively, shows a large difference in the signal-to-noise ratio. The Asn and Gln side-chain region has a much higher signal-to-noise ratio than the backbone amides because the Asn and Gln side-chains are likely to have higher mobility than backbone amides in an α-helix. Figure 4.26 is a ^{1}H, ^{15}N-HSQC spectra of *E. coli* subunit *c* where the NMR signals are well dispersed, and experience relatively little overlap. *E. coli* subunit *c* is a very small protein, only 8.3 kDa in size, and therefore experiences much slower magnetization relaxation than MdfA in NMR experiments, resulting in strong, dispersed NMR signals. In the study of MdfA, NMR experiments were carried out at higher temperatures in order to increase molecular tumbling rate, which in turn reduced the overall correlation time. The shorter correlation time resulted in slower magnetization relaxation, which then led to a stronger signal, less overlap and better resolution. Optimal temperature for recording ^{1}H, ^{15}N-HSQC spectra of MdfA was determined to be 37^{o}C. The spectral overlap seen in the ^{15}N-MdfA HSQC spectrum for backbone amides can be resolved using 3D HNCO NMR. The HNCO experiment correlates amide ^{1}H and ^{15}N chemical shifts with the ^{13}C chemical shift of the preceding residue. This addition of the carbonyl dimension should resolve the amide peaks in three-dimensions, which would dramatically reduce overlap. In preliminary HNCO experiments for MdfA, around 403 chemical shifts from the backbone, and around 30 from Gln and Asn side-chains were expected. However, only a few peaks were observed (data not shown). This is likely a result of fast relaxation as well. The HNCO experiment involves more magnetization transfer steps than HSQC, and therefore sensitivity losses due to relaxation are larger than in HSQC.

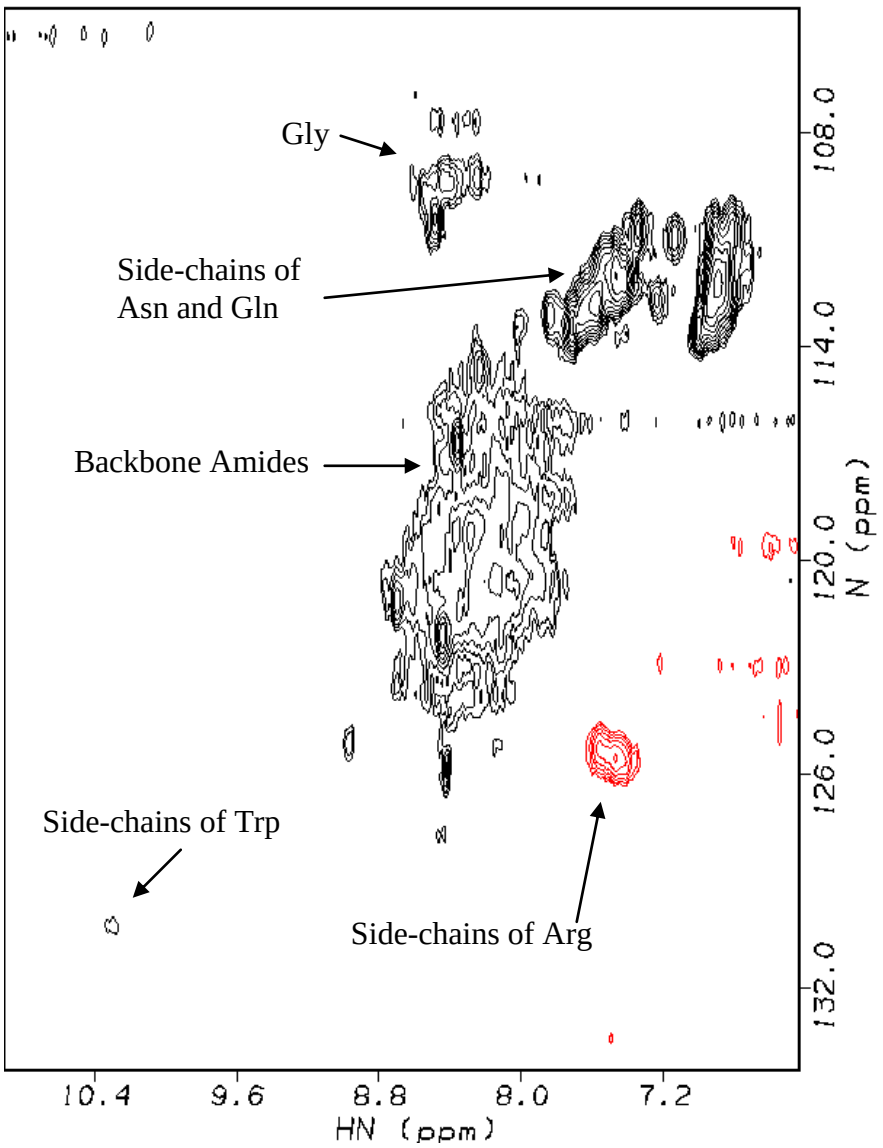


Figure 4.24. ^{1}H, ^{15}N-HSQC NMR spectra of ^{15}N-labeled MdfA in LMPG detergent micelles. NMR data was acquired at the National Magnetic Resonance Facility in Madison, Wisconsin (NMRFAM) using a 750 MHz Bruker spectrometer with a Cryoprobe. 64 scans were collected for every t1 increment, with a total of 64 t1 increments.

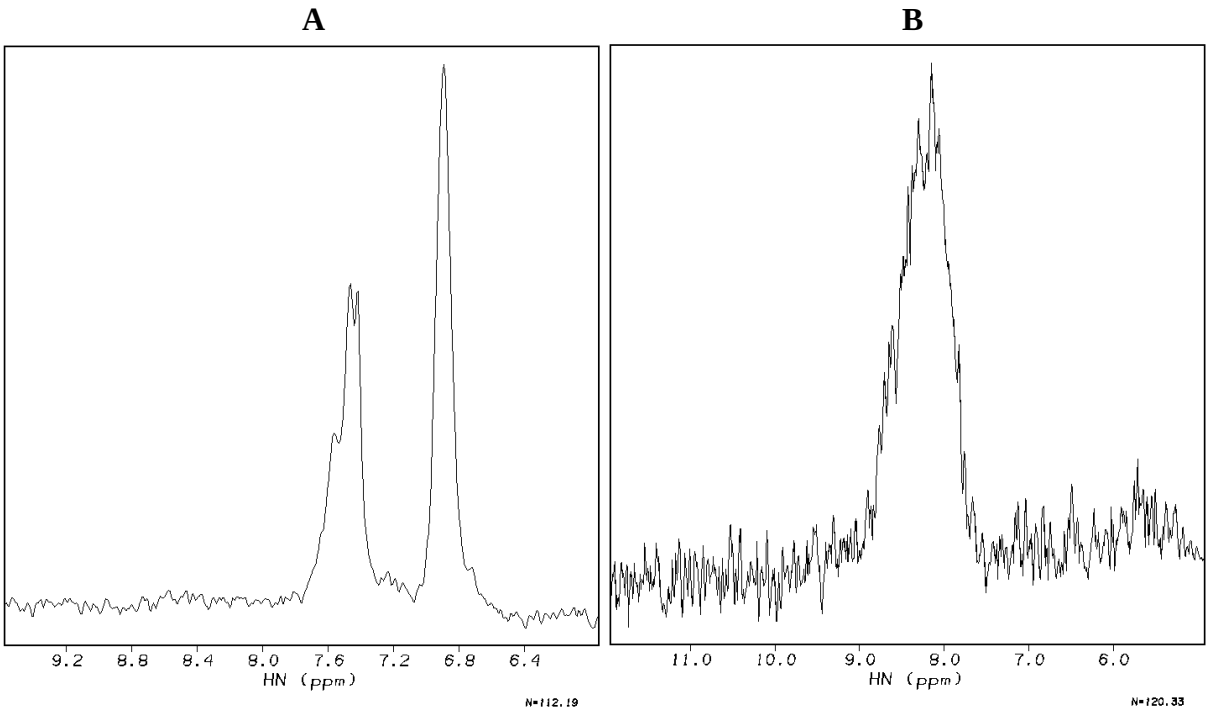


Figure 4.25. (A) One-dimensional slice taken through the Asn and Gln side-chain region from ^{1}H, ^{15}N-HSQC NMR spectra of ^{15}N-labeled MdfA. (B) One-dimensional slice taken through the amide backbone region from ^{1}H, ^{15}N-HSQC NMR spectra of ^{15}N-labeled MdfA. Note significantly higher signal-to-noise ratio for the side chain signals.

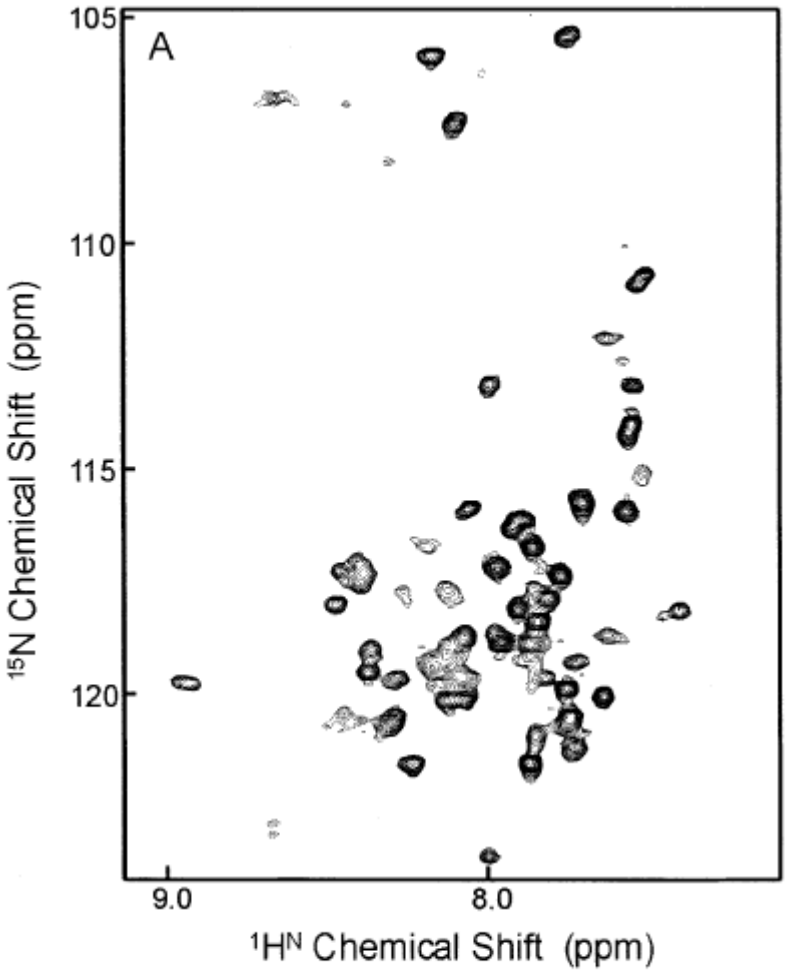


Figure 4.26. ^{1}H, ^{15}N-HSQC NMR spectra of ^{15}N-labeled *E. coli* subunit *c* in 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phosphor-*rac*-(1-glycerol)] detergent micelles (Krueger-Koplin *et al.*, 2004). NMR data was acquired at 600 MHz. Reprinted with permission from Springer Science+Business Media.

The ^{13}C-Glu-^{15}N-Phe-labeled sample was generated to help solve peak overlap problems observed in HNCO data and to conduct substrate binding studies. Since Phe^{27} follows Glu^{26} in MdfA, only the Phe^{27} chemical shift would appear in an HNCO NMR spectrum of the ^{13}C-Glu-^{15}N-Phe-labeled MdfA protein. This experiment would have no peak overlap and since Glu^{26} has been implicated in the binding of cationic substrates, the ^{13}C-Glu-^{15}N-Phe-labeled sample would be ideal for chemical shift perturbation experiments. If a substrate were to interact with Glu^{26}, a perturbation in the chemical shift for Phe^{27} would be

observed due to its close proximity to Glu[26]. Unfortunately, the expected signal of Phe[27] was not detected, possibly due to incomplete labeling.

4.10. MdfA Activity Detected through Proton Translocation Coupled to ACMA Fluorescence

In order to determine if MdfA purified in this research is biologically functional, an MdfA activity assay was developed. Purified MdfA was re-incorporated into liposomes for this activity assay. The activity assay uses ACMA fluorescence coupled to proton translocation as an indicator of transport activity and is described in detail under the Materials and Methods (page 40). Figure 4.27 shows recorded fluorescence through the course of MdfA activity assay. ACMA fluorescence was observed upon the addition of ACMA to the assay solution. A slight drop in fluorescence was observed after the addition of the liposomes and no change in fluorescence was observed upon the addition of valinomycin. An almost complete fluorescence quenching was observed after the addition of FCCP, which allows the influx of protons leading to the protonation and subsequent sequestering of ACMA resulting in concentration-dependent fluorescence quenching. Increase of ACMA fluorescence was observed upon addition of ethidium bromide (Figure 4.28), which indicated proton translocation, and thus MdfA activity because proton translocation is coupled to substrate uptake. Figure 4.28 indicates that the increase of ACMA fluorescence observed upon the addition of ethidium bromide was concentration dependent. This activity assay was repeated using chloramphenicol, where a chloramphenicol concentration dependent increase in ACMA fluorescence was also observed, which indicated proton translocation by MdfA (Figure 4.29). At the present, the time dependence of fluorescence for this assay does not allow for the accurate measurement of the initial rates of substrate transport which are necessary to calculate kinetics such as K_t and V_{max}. However, K_t values for MdfA regarding ethidium bromide and chloramphenicol can be roughly estimated from concentration dependence of fluorescence change shown in Figures 4.28 and 4.29. The K_t values were estimated to be 1 μM - 10 μM for ethidium bromide and 1 μM - 3 μM for chloramphenicol. Edgar and Bibi previously determined that *E.coli* HB101 cells not expressing MdfA could withstand the effects of ethidium bromide up to a concentration of 190 μM and chloramphenicol up to a concentration of 6.19 μM (Edgar and Bibi, 1997). The estimated K_t of 1 μM - 10 μM for ethidium bromide and 1 μM - 3 μM for chloramphenicol

measured for the purified MdfA are consistent with the protective effect of MdfA-catalyzed transport of these drugs *in vivo*.

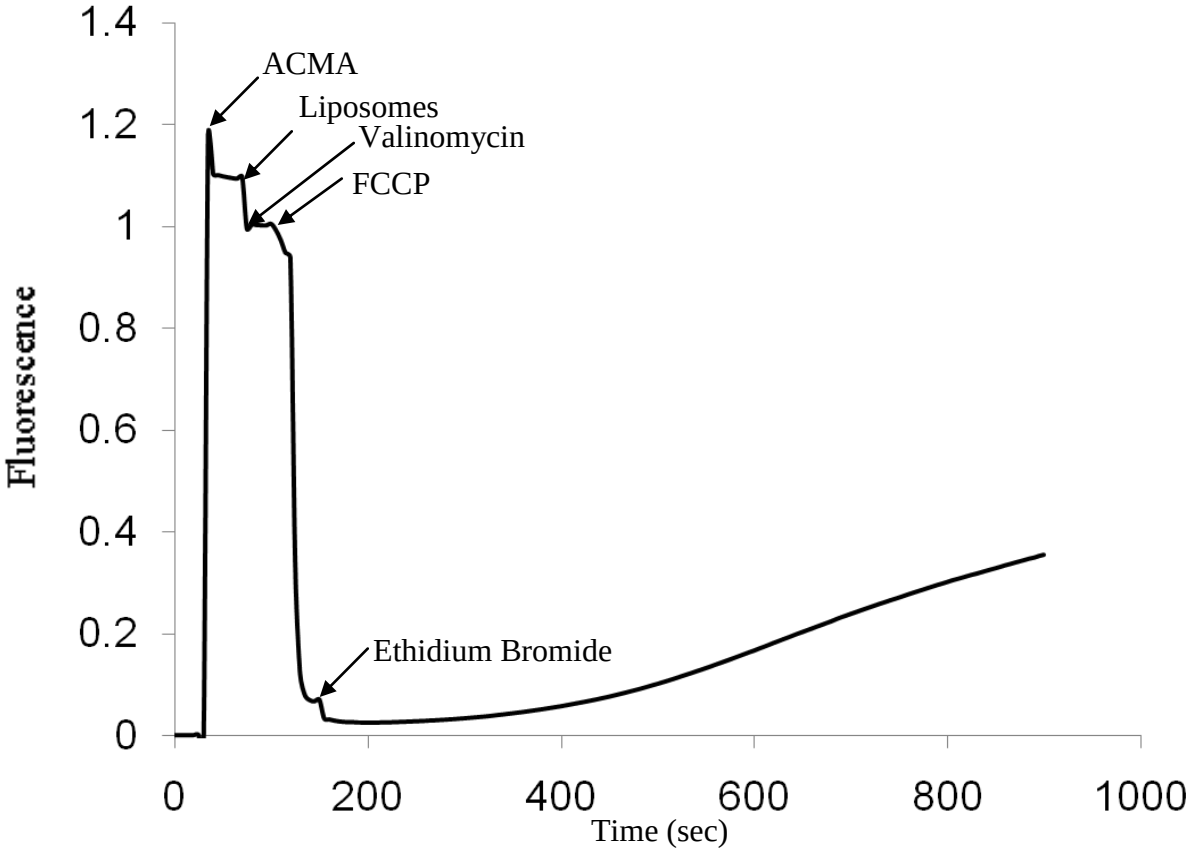


Figure 4.27. ACMA fluorescence graph for the assay developed for MdfA.

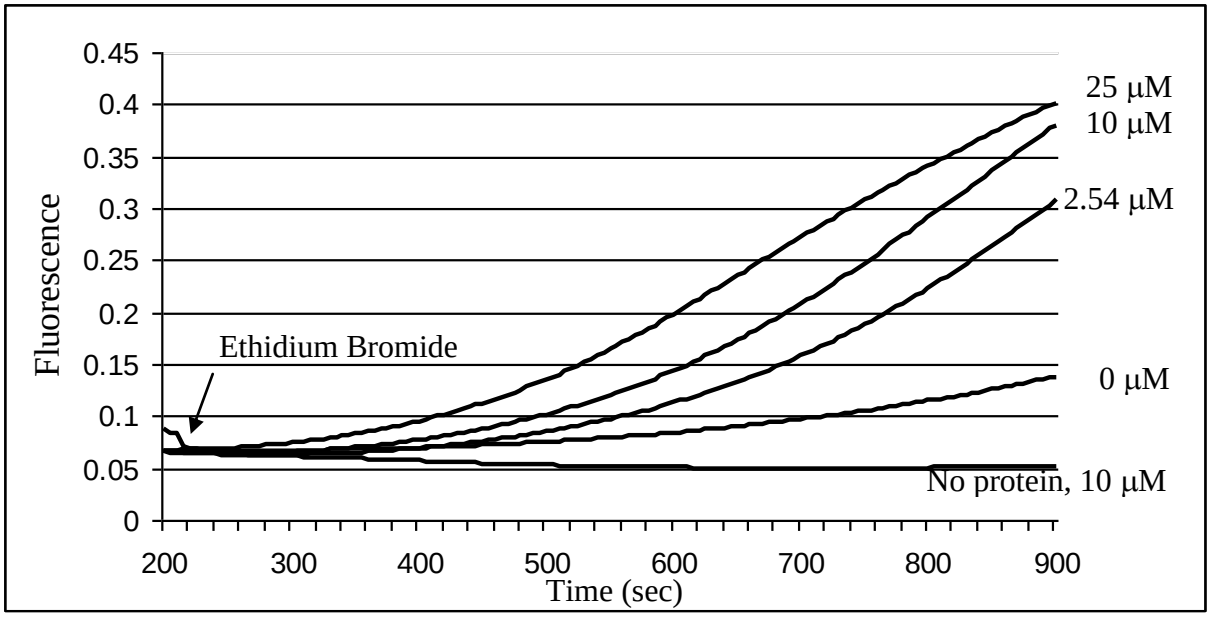


Figure 4.28. Time dependence of ethidium bromide uptake into the MdfA-containing proteoliposomes monitored by ACMA fluorescence. ACMA fluorescence indicates proton translocation, which is coupled to ethidium bromide uptake by MdfA. EtBr concentrations are listed to the right.

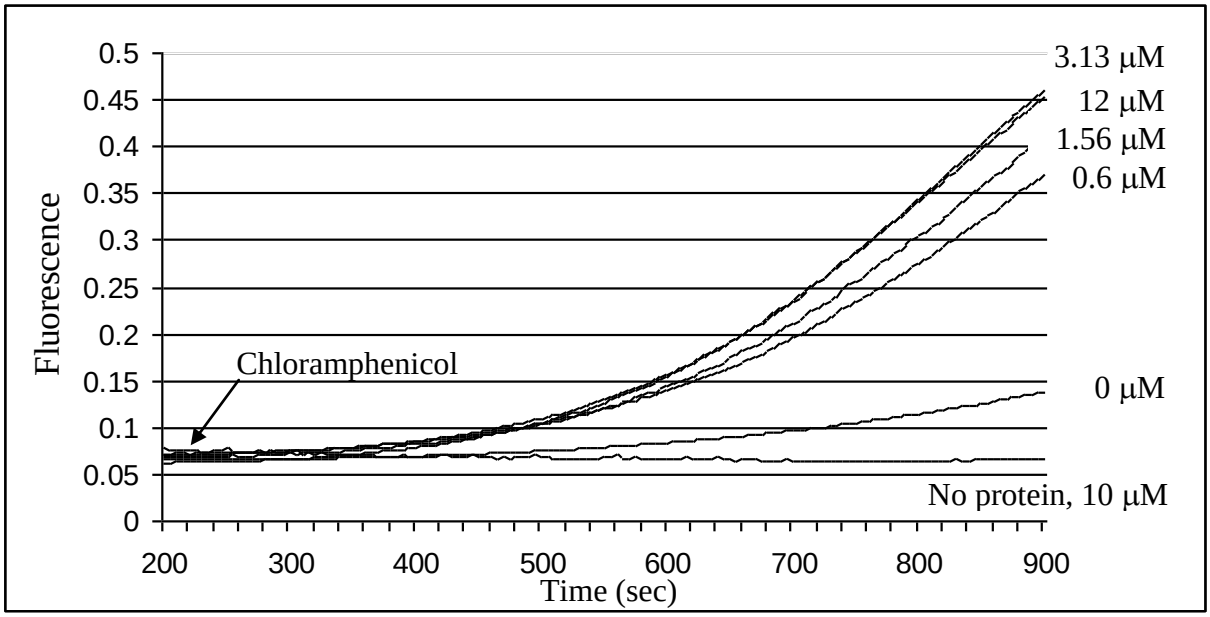


Figure 4.29. Time dependence of chloramphenicol uptake into the MdfA-containing proteoliposomes monitored by ACMA fluorescence. ACMA fluorescence indicates proton translocation, which is coupled to chloramphenicol uptake by MdfA. Chloramphenicol concentrations are listed to the right.

5. Discussion

The goals of this research project were to develop a practical method for purification of MdfA, to evaluate the feasibility of structure determination by NMR and X-ray crystallography, and to develop an activity assay for purified MdfA. MdfA was successfully expressed both in the *E. coli* cell culture and in a cell-free synthesis system. MdfA was successfully extracted from the cell membranes prepared from an overproducing *E. coli* strain and purified via Ni-NTA chromatography. Preliminary X-ray crystallography trials have resulted in several potential protein crystals which need to be further investigated. Initial NMR structural studies using ^{15}N-, ^{15}N, ^{13}C, ^{2}H- and ^{15}N, ^{13}C- labeled MdfA samples have found that using NMR to investigate the structure of MdfA is feasible, but needs more optimization. ^{19}F NMR spectroscopy was found to be a potentially useful tool for investigation of substrate binding to MdfA and for optimization of sample conditions for high-resolution NMR. An activity assay for the purified MdfA has been developed. Purified MdfA reconstituted into the liposomes was found to actively pump substrates in exchange for protons.

5.1. *MdfA* Expression and Purification

The gene encoding MdfA was originally cloned into the pBAD vector and the resulting vector was named pOD1016 (Dmitriev, unpublished data). pOD1016 in LMG194 cells was found to successfully express *mdfA*. In an attempt to increase *mdfA* expression, pOD1016 was modified and expression was tested in different bacterial strains. The genes encoding both thioredoxin and the V5 epitope were deleted from the pOD1016 vector, creating the vector pCOG3. The thioredoxin gene was deleted to determine if the absence of its expression would increase *mdfA* expression and the V5 epitope was deleted because we found that MdfA was detectible via its hexa-histidine tail using a pent-histidine antibody. We found that deleting the thioredoxin gene did not result in any higher expression from the pCOG3 vector compared to the pOD1016 vector. However, deleting the V5 epitope resulted in an MdfA product that was slightly smaller and better-suited for NMR studies. Therefore, the pCOG3 vector was used for protein expression in this research. Two strains, LMG194 and C43 (DE3), were tested to evaluate *mdfA* expression. LMG194 was chosen because it did not metabolize (L)-arabinose and because it was capable of growing on minimal media. C43 (DE3) was chosen because it is able to achieve high levels of expression for both toxic and membrane proteins. *MdfA*

expression levels did not significantly differ between LMG194 and C43 (DE3). We found that a lower (L)-arabinose concentration was required to induce expression in the LMG194 cells compared to the C43 (DE3) cells. Therefore, MdfA was purified from LMG194 cells containing the pCOG3 vector.

LMG194 cells transformed with pCOG3 plasmid were found to successfully express *mdfA* when induced in M63 minimal medium, as well as casamino acid medium, allowing for easy uniform labeling in NMR experiments. The ^{15}N-, ^{15}N, ^{13}C-, ^{15}N, ^{13}C, ^{2}H-, ^{13}C-Glu-, ^{15}N-Phe- and 5-fluoro-tryptophan-labeled MdfA samples were subsequently generated by expressing *mdfA* on minimal media supplemented with isotopically-labeled chemicals.

We attempted to express *mdfA* in a cell-free synthesis system because it allows for the simple generation of selectively-labeled MdfA samples for NMR studies. The pCOG2 plasmid was generated for cell-free synthesis purposes and expressed *mdfA* at levels near 0.1 mg protein per one milliliter reaction mixture, enough to generate samples for NMR studies. A problem that could potentially arise in the cell-free synthesis of membrane proteins is improper folding. Addition of liposomes and/or chaperone proteins can aid in proper folding and direct newly synthesized proteins into the lipid bilayer. MdfA has been expressed using cell-free synthesis in the presence of liposomes. However, further investigation is needed to determine if functional MdfA was incorporated into the liposomes.

In order to conduct NMR or X-ray crystallography structural studies on membrane proteins, the protein must first be extracted from the cell membrane. This is commonly achieved by employing detergents. Detergents can displace the lipids around a protein, thereby forming a detergent micelle. Detergents are generally chosen based on cost and suitability. A more affordable detergent can be used for initial purification then replaced with one more suitable for NMR experiments. MdfA was successfully extracted from cell membranes using several different detergents. LMPG was the most efficient at extracting MdfA from cell membranes, with over 60% of the MdfA present extracted at a final detergent concentration of 1.0%. However, LMPG was not used for extraction after preliminary tests due to its high cost. DHPC was chosen as an alternative to LMPG, and provided adequate extraction efficiency, extracting around half the MdfA present at a final concentration of 2.0%. MdfA was also successfully extracted using DDM. DDM is a mild detergent that is often able to preserve protein activity and is commonly used in membrane protein crystallization (Banerjee *et al.*,

1995; Auer *et al.*, 2001; Boutler and Wang, 2001; Engel *et al.*, 2002). Several structures have been solved for membrane proteins purified in DDM detergent micelles including LacY, EmrD, EmrE, and GlpT (Auer *et al.*, 2001; Ambramson *et al.*, 2003; Yin *et al.*, 2006; Korkhov and Tate, 2009). MdfA in DDM detergent micelles has not been used for crystallization trials in this work. However, due to the successful use of DDM in the crystallization of membrane proteins, it should not be overlooked in future studies.

In order to conduct structural studies via NMR or X-ray crystallography, it is essential to obtain protein of 95% purity or higher. MdfA was successfully purified by Ni-NTA affinity chromatography using a hexa-histidine-tag. The purification method developed in this research resulted in 1.0 - 1.4 mg of pure MdfA protein per 1 L bacterial culture, which is significantly more than the yield of 0.3 mg per 1 L culture obtained by the purification method developed by Sigal (Sigal *et al.*, 2006). This increase in yield greatly decreases the amount of culture that is needed to obtain MdfA in sufficient amount for NMR or X-ray crystallography, and in turn greatly reduces the quantities of potentially expensive detergents needed in the purification process. The purified MdfA contained very little contaminating protein according to visual inspection of Coomassie-stained and silver-stained SDS-PAGE gels. The purity of the preparation was similar to that obtained previously (Lewinson and Bibi, 2001; Sigal *et al.*, 2006). The purification procedure developed by this research is practical, easy, very reproducible, and cost effective.

5.2. Suitability for Structural Studies

5.2.1. MdfA Stability Optimization

For structural studies by NMR or X-ray crystallography, high protein stability over time is essential. Crystallization trials can last weeks or months and a single multidimensional NMR experiment can last several days. Initial purification trials resulted in unstable MdfA samples that denatured and precipitated within 7 days at 4^{o}C (visual observation). Initially, MdfA was purified in TMDG buffer (pH 7.5) containing 0.4% DHPC. By testing protein stability in a variety of solutions with different pH and containing different detergents, conditions were found which greatly increased MdfA stability. LMPG was found to have the largest impact on stability. We compared MdfA stability at 27^{o}C in DHPC and LMPG. Nearly all of the purified MdfA precipitated out of the DHPC solution after 72 hours. In comparison, nearly 100% of the

protein remained soluble in LMPG after the same time period. Investigation of the pH effect on stability showed that MdfA was most stable at pH 5 - 6, which is consistent with the estimated isoelectric point of 8.85 (Protein Calculator v3.3, http://www.scripps.edu/~cdputnam/protcalc.html). The lower overall net charge in TMDG buffer at pH 7.5 could decrease protein solubility as observed in our experiments. Investigation of the effect of sodium chloride showed that MdfA stability was slightly improved in the presence of 100 mM NaCl. Optimal stability was achieved using a 25 mM sodium phosphate buffer (pH 6.0), containing 0.2% LMPG and 100 mM NaCl. After optimization, the MdfA was found to be soluble for at least two months at room temperature (visual inspection). The stability of MdfA in solution is suitable for both NMR and X-ray crystallographic studies.

5.2.2. Investigation of the MdfA Aggregation State

Monodispersity of a protein solution is extremely important in NMR and X-ray crystallography structural studies. In addition to the MdfA monomer, several prominent extra bands were observed via SDS-PAGE of purified MdfA. Upon further examination, the extra bands were determined to be oligomers of MdfA. Three observations were used to draw that conclusion. First, the extra bands were detected by the anti-penta-histidine antibody used to visualize MdfA monomers. Second, molecular weights corresponding to the extra bands were consistent with the dimer, trimer, and tetramer of MdfA, respectively. Third, relative band intensity strongly depended on sample treatment temperature. If the extra bands were contaminating proteins, they would likely have similar relative intensities at all tested temperatures. Based on these experiments, pure MdfA has a strong tendency to aggregate at elevated temperatures. These aggregates were stable even in the presence of 6 M urea in the loading buffer and SDS-polyacrylamide gel. Formation of stable aggregates that do not dissociate even in the presence of SDS is not uncommon for membrane proteins (Sagne, *et al.*, 1996). When MdfA was not exposed to high temperatures prior to loading onto SDS-polyacrylamide gels, it predominantly migrated as a monomer.

To investigate the aggregation state of purified MdfA, dynamic light scattering experiments were performed demonstrating that MdfA behaved in a monodisperse manner up to the protein concentration of 4 mg/mL in DDM, DHPC, and LMPG micelles at temperatures up to 30^{o}C. The particle size distribution in the dynamic light scattering experiments was

between 3 - 4 nm, which is a reasonable estimate for the 47 kDa MdfA in the detergent micelle. Since MdfA behaves in a monodisperse manner, it is suitable for NMR experiments and X-ray crystallization trials.

5.3. Functional Studies

In order to determine if the extracted and purified MdfA remained correctly folded and functional, substrate binding and transport were investigated. ^{19}F NMR experiments were used to investigate substrate binding and an MdfA activity assay was developed to investigate substrate transport.

5.3.1. Substrate Binding by ^{19}F NMR

There are several advantages to using ^{19}F NMR. One advantage is that this variant of selective labeling can greatly simplify the NMR spectra of large proteins. Another advantage is that this 100% naturally occurring isotope (^{19}F) has a spin of one-half and a high magnetogyric ratio, making it very sensitive to NMR detection, second only to ^{1}H. As well, the nonexistence of naturally occurring fluorine atoms in protein results in ^{19}F protein NMR studies with no background signal interference. Another advantage is the low cost of generating fluorine labeled protein samples. When conducting NMR experiments on differently-labeled MdfA protein samples, we found that the ^{19}F NMR experiment resulted in a spectrum with the least overlap and a few nearly distinguishable peaks. Chemical shift perturbation experiments were conducted to determine if solubilized MdfA was able to bind its substrates. Upon addition of MdfA substrates ethidium bromide and chloramphenicol, minor changes in the chemical shifts were observed, suggesting that MdfA interacted with these substrates. However, due to peak overlap and corresponding broad peaks in the NMR spectra, it was not possible to calculate accurate binding constants for these experiments. ^{19}F NMR remains an attractive approach for substrate binding studies and NMR sample optimization experiments. Potentially very high sensitivity of ^{19}F-NMR is compromised by the outdated design of the available ^{19}F NMR probes that are no match for modern probes with ^{1}H-detection, developed specifically for protein NMR.

5.3.2. MdfA Activity Detected by ACMA Fluorescence

The transport of substrates by multidrug transporters has been mostly studied *in vivo*. This usually involves over-expression of the protein of interest to determine which drugs become less effective as a result of the protein expression. A simple *in vitro* assay was developed, which can potentially be used to study substrate specificity for all multidrug transporters driven by transmembrane proton gradients. This assay detects transport of substrates into liposomes by tracking the movement of protons via monitoring of the internal liposomal pH using ACMA fluorescence. Using this assay we determined that purified MdfA re-incorporated into liposomes bound its substrates ethidium bromide and chloramphenicol, and transported them into the liposomes in exchange for protons. At the present, the long time dependence of fluorescence for this assay does not allow for the accurate measurement of the initial rates of substrate transport which are necessary to calculate kinetics such as K_t and V_{max}. The long time dependence may be improved by using lower buffer concentration inside the liposomes to increase the magnitude of pH-change resulting from drug uptake. The highest tolerated drug concentrations for *E. coli* without artificial over-expression of MdfA were determined to be about 200 μM for ethidium bromide and about 6 μM for chloramphenicol (Edgar and Bibi, 1997). Thus, the estimated Kt values of 1 - 10 μM for ethidium bromide and 1 - 3 μM for chloramphenicol *in vitro* are consistent with the protective effect of MdfA *in vivo*.

When the activity assay was carried out in proteoliposomes lacking substrates, a small rise in fluorescence was observed, indicating a proton leak from the liposomes. There are several possible explanations for this phenomenon. One possibility is that the proteoliposomes are destabilized by the addition of MdfA, leading to nonspecific increase in ion permeability. Another possibility is the slow rate of proton transport observed without the substrate can be accounted for by the intrinsic Na^+/H^+-activity of MdfA. MdfA could be pumping H^+ out of the liposome in exchange for Na^+, which would result in the small rise of fluorescence. MdfA is also believed to have intrinsic K^+/H^+ -activity. However, we did not see any indications of K^+/H^+ antiport by MdfA in our transport assay.

The time dependence of fluorescence for this assay at present does not allow for the accurate measurement of the initial rates of substrate transport which are necessary to calculate kinetics such as K_t and V_{max}, but it holds a lot of promise for future studies. Once this activity assay is optimized, MdfA substrate binding and pumping can be studied in greater detail. This

assay could also be used for other multidrug transporters which use proton gradients to facilitate transport.

5.4. Initial Structural Studies

5.4.1. Initial Crystallization Screening

High-throughput crystallization screening of MdfA purified in LMPG was conducted at the Hauptman-Woodward Medical Research Institute in Buffalo, NY. This screening process tests over 1500 conditions in a one month period. Initial screens yielded several sets of conditions that produced crystals. These conditions were reproduced in our laboratory, but unfortunately, none of the crystals were found to consist of protein. Further microbatch screening using the MPD Suite (Qiagen) resulted in two sets of conditions yielding crystals. MPD has been used in the crystallization of several membrane proteins (Raman *et al.*, 2006). The crystals took up the dye Ponceau S, indicating that the crystals likely contained protein. These conditions are being optimized further to obtain larger crystals, which will be tested by X-ray diffraction to determine if they contain protein. If the crystals are not made up of protein, there are several options available for further crystallization trials. One option in membrane protein crystallization is to test different detergent micelles. A commonly used detergent in membrane crystallography is DDM, and it has been successfully used in the crystallization of several MFS transporters, including EmrD, GlpT and LacY (Abramson *et al.*, 2003; Huang *et al.*, 2003; Yin *et al.*, 2006). MdfA has been successfully extracted and purified using the detergent DDM which could be the next step in crystal screening. Another option in membrane protein crystal screening is to use the *in-meso* screening method (Cherezov *et al.*, 2002). The *in-meso* method makes use of a lipidic mesophase as the hosting medium and protein reservoir from which protein crystals grow (Caffery and Cherezov, 2009). This method is similar to typical crystal screening where similar screening conditions are used. The use of this method will eliminate the presence of detergents in the crystallization process since the protein is present in a lipidic mesophase instead of a detergent micelle, and the crystals themselves will grow in the lipidic mesophase.

5.4.2. Initial Structural Studies by NMR Spectroscopy

To date, there are only approximately 30 unique membrane protein structures that have been determined using NMR spectroscopy (http://www.drorlist.com/nmr/MPNMR.html, September 30, 2009). Many of the alpha-helical NMR structures are of small polypeptides of 4 kDa to 10 kDa in size. However, there are several structures of alpha-helical proteins solved which range in size from 30 kDa to 71 kDa including the structure for the 43 kDa homotrimer Diacylglycerol Kinase and the 71 kDa homotetramer KcsA (Introduction, page 21). Thus far, NMR spectroscopy was carried out on ^{15}N-, ^{15}N, ^{13}C- and ^{15}N, ^{13}C, ^{2}H-labeled MdfA samples to determine the feasibility of high-resolution studies by NMR. From our results, we determined that high-resolution NMR studies on MdfA are feasible. Not only have we developed an expression system capable of generating sufficient amounts of stable uniform isotopically labeled- MdfA, we have also obtained NMR spectra which are of the initial quality that can be expected for a mostly alpha-helical protein of this size in detergent micelles. The NMR spectra have poor resolution resulting in signal overlap which is likely a result of the fast relaxation of magnetization caused by the large size of MdfA and its detergent micelle. However, it is possible to reduce relaxation of magnetization by lowering the overall size of MdfA and its micelle. This can be achieved by purifying MdfA using organic solvent-water mixtures as alternatives to detergents. Several biologically-relevant structures have been solved for membrane proteins purified in organic solvent-water mixtures using NMR spectroscopy including subunit *c* from ATP synthase (Girvin *et al.*, 1998) and a partial structure of subunit *a* from ATP synthase (Dmitriev *et al.*, 2008). Signal overlap can also be resolved by three-dimensional experiments like HNCO and can be further resolved by using three-dimensional experiments in conjunction with amino acid selective labeling and/or TROSY experiments.

The ^{19}F NMR experiments which were carried out during this research project can be used to optimize sample preparation for NMR experiments. ^{19}F NMR is a good choice for optimization because of the low cost of labeling. Several overlapping peaks were observed in the ^{19}F NMR experiments for 5-fluoro-tryptophan-labeled MdfA where 10 distinct peaks were expected. The optimization of NMR conditions can be evaluated by monitoring the linewidth of tryptophan signals. When the linewidths are at their narrowest, the optimal conditions will be met. These conditions can then be applied to the multidimensional NMR experiments required for structural studies.

6. Conclusions and Future Work

6.1. Conclusions

MdfA was successfully over-expressed, extracted from the bacterial membrane using non-denaturing detergents, and subsequently purified by Ni-NTA chromatography. MdfA was also successfully expressed in a cell-free synthesis system, paving the way for the selective isotope labeling which is needed for NMR studies. The folding properties and activity of MdfA prepared in a cell-free system remains to be determined. The sample conditions for MdfA have been optimized for maximum stability and the protein preparation was found to behave in a monodisperse manner in solution, making MdfA suitable for initial structural studies. Preliminary ^{19}F NMR studies indicated that the purified MdfA protein interacted with the substrates ethidium bromide and chloramphenicol. Proper folding and function of purified MdfA was revealed through the ACMA fluorescence rebounding assay. Initial NMR studies of MdfA indicate that high-resolution NMR studies are feasible, but further sample optimization is needed. Crystallization trials for MdfA have yielded two conditions which have generated potential protein crystals. X-ray crystallography is likely better suited than NMR for structural determination of MdfA. However, NMR is suitable for substrate binding studies, protein dynamic analysis, and possibly global fold determination.

6.2. Future Work

Cell-free synthesis of MdfA should be optimized and examined in the presence of either liposomes or detergents to ensure proper folding. Proper folding can be determined by detecting MdfA activity using the coupled proton transport assay developed for this study. Once proper folding has been established, selective amino acid labeling of MdfA should be investigated and NMR experiments should follow to determine the feasibility of this approach for mapping the drug binding site.

Optimization of NMR experiments for MdfA needs to be accomplished. The initial optimization of sample conditions for NMR experiments can be done using the 5-fluoro-tryptophan-labeled MdfA to record ^{19}F spectra and assess linewidth of tryptophan signals. The multidimensional NMR spectra need to be improved in order to obtain information about the structure of MdfA. The alteration of experimental conditions, testing different buffers and

detergents, or using organic solvents in place of detergents, selective isotope labeling, and possibly carrying out NMR experiments at the higher magnetic field of 900 MHz could improve NMR spectra. It would be extremely valuable to conduct TROSY experiments on the new 1000 MHz NMR spectrometer developed by Bruker, because the optimal field strength for TROSY is between 950-1050 MHz (Wider and Wuthrich, 1999). The first 1000 MHz NMR spectrometer was installed at the Centre de RMN à Très Hauts Champs in Lyon, France in August 2009.

The activity assay requires further optimization to improve the speed of the assay. Altering buffer component concentrations, liposome concentration, liposome size, and MdfA concentration may all improve the speed of this assay. The optimization will allow for accurate kinetic measurements for any proton/drug antiporter.

The X-ray crystallographic studies on MdfA are promising and should continue. If the recently discovered crystals do not turn out to consist of protein, MdfA should be purified in several different detergents and analyzed by large scale crystallization trials in attempts to obtain protein crystals. Membrane proteins have been successfully crystallized using DDM and this would therefore be a good starting point (Yin *et al*., 2006). Another option is to use *in-meso* screening method which is done in liposomes, therefore eliminating detergents from the crystal screens.

7. References

Abraham, E.P., & Chain, E. (1940). An enzyme from bacteria able to destroy penicillin. Nature *146,* 837.

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., and Iwata, S. (2003). Structure and mechanism of the lactose permease of Escherichia coli. Science *301,* 610-615.

Adler, J., and Bibi, E. (2004). Determinants of substrate recognition by the Escherichia coli multidrug transporter MdfA identified on both sides of the membrane. J. Biol. Chem. *279,* 8957-8965.

Adler, J., and Bibi, E. (2002). Membrane topology of the multidrug transporter MdfA: complementary gene fusion studies reveal a nonessential C-terminal domain. J. Bacteriol. *184,* 3313-3320.

Adler, J., Lewinson, O., and Bibi, E. (2004). Role of a conserved membrane-embedded acidic residue in the multidrug transporter MdfA. Biochemistry *43,* 518-525.

Aller, S.G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P.M., Trinh, Y.T., Zhang, Q., Urbatsch, I.L., and Chang, G. (2009). Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science *323,* 1718-1722.

Auer, M., Kim, M.J., Lemieux, M.J., Villa, A., Song, J., Li, X.D., and Wang, D.N. (2001). High-yield expression and functional analysis of Escherichia coli glycerol-3-phosphate transporter. Biochemistry *40,* 6628-6635.

Banerjee, P., Joo, J.B., Buse, J.T., and Dawson, G. (1995). Differential solubilization of lipids along with membrane proteins by different classes of detergents. Chem. Phys. Lipids *77,* 65-78.

Bedouelle, H., Duplay, P., and Hofnung, M. (1987). Expression, export and one-step purification of proteins by fusion to the MalE protein of E. coli. C. R. Acad. Sci. III. *305,* 623-626.

Boulter, J.M., and Wang, D.N. (2001). Purification and characterization of human erythrocyte glucose transporter in decylmaltoside detergent solution. Protein Expr. Purif. *22,* 337-348.

Bouvignies, G., Markwick, P.R., and Blackledge, M. (2007). Simultaneous definition of high resolution protein structure and backbone conformational dynamics using NMR residual dipolar couplings. Chemphyschem *8,* 1901-1909.

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

Browne, D.T., Kenyon, G.L., Packer, E.L., Sternlicht, H., and Wilson, D.M. (1973). Studies of macromolecular structure by 13 C nuclear magnetic resonance. II. A specific labeling approach to the study of histidine residues in proteins. J. Am. Chem. Soc. *95,* 1316-1323.

Caffrey, M., and Cherezov, V. (2009). Crystallizing membrane proteins using lipidic mesophases. Nat. Protoc. *4,* 706-731.

Chattopadhyay, A., Harikumar, K.G., and Kalipatnapu, S. (2002). Solubilization of high affinity G-protein-coupled serotonin1A receptors from bovine hippocampus using pre-micellar CHAPS at low concentration. Mol. Membr. Biol. *19,* 211-220.

Chen, C.J., Chin, J.E., Ueda, K., Clark, D.P., Pastan, I., Gottesman, M.M., and Roninson, I.B. (1986). Internal duplication and homology with bacterial transport proteins in the mdr1 (P-glycoprotein) gene from multidrug-resistant human cells. Cell *47,* 381-389.

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

Cherezov, V., Clogston, J., Misquitta, Y., Abdel-Gawad, W., and Caffrey, M. (2002). Membrane protein crystallization in meso: lipid type-tailoring of the cubic phase. Biophys. J. *83,* 3393-3407.

Chong, S., Mersha, F.B., Comb, D.G., Scott, M.E., Landry, D., Vence, L.M., Perler, F.B., Benner, J., Kucera, R.B., Hirvonen, C.A. *et al*. (1997). Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene *192,* 271-281.

Choudhuri, B.S., Bhakta, S., Barik, R., Basu, J., Kundu, M., and Chakrabarti, P. (2002). Overexpression and functional characterization of an ABC (ATP-binding cassette) transporter encoded by the genes drrA and drrB of Mycobacterium tuberculosis. Biochem. J. *367,* 279-285.

Chung, Y.J., and Saier, M.H.,Jr. (2001). SMR-type multidrug resistance pumps. Curr. Opin. Drug Discov. Devel. *4,* 237-245.

Cladera, J., Rigaud, J.L., Villaverde, J., and Dunach, M. (1997). Liposome solubilization and membrane protein reconstitution using Chaps and Chapso. Eur. J. Biochem. *243,* 798-804.

Dawson, R.J., and Locher, K.P. (2006). Structure of a bacterial multidrug ABC transporter. Nature *443,* 180-185.

Dmitriev, O.Y., Freedman, K.H., Hermolin, J., and Fillingame, R.H. (2008). Interaction of transmembrane helices in ATP synthase subunit a in solution as revealed by spin label difference NMR. Biochim. Biophys. Acta *1777,* 227-237.

Dmitriev, O.Y., and Uhleman, E. (2007). Cell-free Synthesis of Membrane Subunits of E. coli ATP synthase. Book of abstracts.

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

Edgar, R., and Bibi, E. (1997). MdfA, an Escherichia coli multidrug resistance protein with an extraordinarily broad spectrum of drug recognition. J. Bacteriol. *179,* 2274-2280.

Endicott, J.A., and Ling, V. (1989). The biochemistry of P-glycoprotein-mediated multidrug resistance. Annu. Rev. Biochem. *58,* 137-171.

Engel, C.K., Chen, L., and Prive, G.G. (2002). Stability of the lactose permease in detergent solutions. Biochim. Biophys. Acta *1564,* 47-56.

Fath, M.J., and Kolter, R. (1993). ABC transporters: bacterial exporters. Microbiol. Rev. *57,* 995-1017.

Fischer, G., Kosinska-Eriksson, U., Aponte-Santamaria, C., Palmgren, M., Geijer, C., Hedfalk, K., Hohmann, S., de Groot, B.L., Neutze, R., and Lindkvist-Petersson, K. (2009). Crystal structure of a yeast aquaporin at 1.15 angstrom reveals a novel gating mechanism. PLoS Biol. *7,* e1000130.

Fralick, J.A. (1996). Evidence that TolC is required for functioning of the Mar/AcrAB efflux pump of Escherichia coli. J. Bacteriol. *178,* 5803-5805.

Garbuzynskiy, S.O., Melnik, B.S., Lobanov, M.Y., Finkelstein, A.V., and Galzitskaya, O.V. (2005). Comparison of X-ray and NMR structures: is there a systematic difference in residue contacts between X-ray- and NMR-resolved protein structures? Proteins *60,* 139-147.

Girvin, M.E., Rastogi, V.K., Abildgaard, F., Markley, J.L., and Fillingame, R.H. (1998). Solution structure of the transmembrane H+-transporting subunit c of the F1F0 ATP synthase. Biochemistry *37,* 8817-8824.

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

Gottesman, M.M., and Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. Annu. Rev. Biochem. *62,* 385-427.

Griffith, J.K., Baker, M.E., Rouch, D.A., Page, M.G., Skurray, R.A., Paulsen, I.T., Chater, K.F., Baldwin, S.A., and Henderson, P.J. (1992). Membrane transport proteins: implications of sequence comparisons. Curr. Opin. Cell Biol. *4,* 684-695.

Grinius, L., Dreguniene, G., Goldberg, E.B., Liao, C.H., and Projan, S.J. (1992). A staphylococcal multidrug resistance gene product is a member of a new protein family. Plasmid *27,* 119-129.

Grzesiek, S., Anglister, J., Hao, R., and Bax, A. (1993). Carbon-13 line narrowing by deuterium decoupling in deuterium/carbon-13/nitrogen-15 enriched proteins. Application to triple resonance 4D J connectivity of sequential amides. J. Am. Chem. Soc. *115,* 4369-4370.

Guan, L., Ehrmann, M., Yoneyama, H., and Nakae, T. (1999). Membrane topology of the xenobiotic-exporting subunit, MexB, of the MexA,B-OprM extrusion pump in Pseudomonas aeruginosa. J. Biol. Chem. *274,* 10517-10522.

Henderson, P.J., and Maiden, M.C. (1990). Homologous sugar transport proteins in Escherichia coli and their relatives in both prokaryotes and eukaryotes. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *326,* 391-410.

Higgins, C.F. (2001). ABC transporters: physiology, structure and mechanism--an overview. Res. Microbiol. *152,* 205-210.

Higgins, C.F. (1992). ABC transporters: from microorganisms to man. Annu. Rev. Cell Biol. *8, 67*-113.

Hiramatsu, K., Hanaki, H., Ino, T., Yabuta, K., Oguri, T., and Tenover, F.C. (1997). Methicillin-resistant Staphylococcus aureus clinical strain with reduced vancomycin susceptibility. J. Antimicrob. Chemother. *40,* 135-136.

Hochuli, E., Dobeli, H., and Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. J. Chromatogr. *411,* 177-184.

Howe, R.A., Bowker, K.E., Walsh, T.R., Feest, T.G., and MacGowan, A.P. (1998). Vancomycin-resistant Staphylococcus aureus. Lancet *351,* 602.

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

Jelsch, C., Teeter, M.M., Lamzin, V., Pichon-Pesme, V., Blessing, R.H., and Lecomte, C. (2000). Accurate protein crystallography at ultra-high resolution: valence electron distribution in crambin. Proc. Natl. Acad. Sci. U. S. A. *97,* 3171-3176.

Jevons, M.P. (2009). "Celbenin" - resistant Staphylococci. Br. Med. J. *1,* 124-125.

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

Kigawa, T., Muto, Y., and Yokoyama, S. (1995). Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. J. Biomol. NMR *6,* 129-134.

King, C.H., Shlaes, D.M., and Dul, M.J. (1983). Infection caused by thymidine-requiring, trimethoprim-resistant bacteria. J. Clin. Microbiol. *18,* 79-83.

Klein, E., Smith, D.L., and Laxminarayan, R. (2007). Hospitalizations and deaths caused by methicillin-resistant Staphylococcus aureus, United States, 1999-2005. Emerg. Infect. Dis. *13,* 1840-1846.

Korkhov, V.M., and Tate, C.G. (2009). An emerging consensus for the structure of EmrE. Acta Crystallogr. D Biol. Crystallogr. *65,* 186-192.

Krueger-Koplin, R.D., Sorgen, P.L., Krueger-Koplin, S.T., Rivera-Torres, I.O., Cahill, S.M., Hicks, D.B., Grinius, L., Krulwich, T.A., and Girvin, M.E. (2004). An evaluation of detergents for NMR structural studies of membrane proteins. J. Biomol. NMR *28,* 43-57.

Krulwich, T.A., Jin, J., Guffanti, A.A., and Bechhofer, H. (2001). Functions of tetracycline efflux proteins that do not involve tetracycline. J. Mol. Microbiol. Biotechnol. *3,* 237-246.

Krulwich, T.A., Lewinson, O., Padan, E., and Bibi, E. (2005). Do physiological roles foster persistence of drug/multidrug-efflux transporters? A case study. Nat. Rev. Microbiol. *3,* 566-572.

Laible, G., and Hakenbeck, R. (1991). Five independent combinations of mutations can result in low-affinity penicillin-binding protein 2x of Streptococcus pneumoniae. J. Bacteriol. *173,* 6986-6990.

Lee, C.G., Gottesman, M.M., Cardarelli, C.O., Ramachandra, M., Jeang, K.T., Ambudkar, S.V., Pastan, I., and Dey, S. (1998). HIV-1 protease inhibitors are substrates for the MDR1 multidrug transporter. Biochemistry *37,* 3594-3601.

Lewinson, O., and Bibi, E. (2001). Evidence for simultaneous binding of dissimilar substrates by the Escherichia coli multidrug transporter MdfA. Biochemistry *40,* 12612-12618.

Lewinson, O., Padan, E., and Bibi, E. (2004). Alkalitolerance: a biological function for a multidrug transporter in pH homeostasis. Proc. Natl. Acad. Sci. U. S. A. *101,* 14073-14078.

Li, X.Z., Nikaido, H., and Poole, K. (1995). Role of mexA-mexB-oprM in antibiotic efflux in Pseudomonas aeruginosa. Antimicrob. Agents Chemother. *39,* 1948-1953.

Lipari, G. and Szabo, A. (1982). Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. J. Am. Chem. Soc *104,* 4546-4559.

Lomovskaya, O., and Lewis, K. (1992). Emr, an Escherichia coli locus for multidrug resistance. Proc. Natl. Acad. Sci. U. S. A. *89,* 8938-8942.

Loscher, W., and Potschka, H. (2002). Role of multidrug transporters in pharmacoresistance to antiepileptic drugs. J. Pharmacol. Exp. Ther. *301,* 7-14.

Lowry, O.H., Rosenbrough, N.J., Farr, A.L., and Randall, R.J. (1951). Protein measurement with the Folin phenol reagent. J. Biol. Chem. *193,* 265-275.

Maiden, M.C., Davis, E.O., Baldwin, S.A., Moore, D.C., and Henderson, P.J. (1987). Mammalian and bacterial sugar transport proteins are homologous. Nature *325,* 641-643.

Marger, M.D., and Saier, M.H.,Jr. (1993). A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends Biochem. Sci. *18,* 13-20.

Mine, T., Morita, Y., Kataoka, A., Mizushima, T., and Tsuchiya, T. (1998). Evidence for chloramphenicol/H+ antiport in Cmr (MdfA) system of Escherichia coli and properties of the antiporter. J. Biochem. *124,* 187-193.

Miroux, B., and Walker, J.E. (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J. Mol. Biol. *260,* 289-298.

Mirza, O., Guan, L., Verner, G., Iwata, S., and Kaback, H.R. (2006). Structural evidence for induced fit and a mechanism for sugar/H+ symport in LacY. EMBO J. *25,* 1177-1183.

Miyake, J., Ochiai-Yanagi, S., Kasumi, T., and Takagi, T. (1978). Isolation of a membrane protein from R rubrum chromatophores and its abnormal behavior in SDS-polyacrylamide gel electrophoresis due to a high binding capacity for SDS. J. Biochem. *83,* 1679-1686.

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

Morita, Y., Kodama, K., Shiota, S., Mine, T., Kataoka, A., Mizushima, T., and Tsuchiya, T. (1998). NorM, a putative multidrug efflux protein, of Vibrio parahaemolyticus and its homolog in Escherichia coli. Antimicrob. Agents Chemother. *42,* 1778-1782.

Muchmore, S.W., Sattler, M., Liang, H., Meadows, R.P., Harlan, J.E., Yoon, H.S., Nettesheim, D., Chang, B.S., Thompson, C.B., Wong, S.L., Ng, S.L., and Fesik, S.W. (1996). X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. Nature *381,* 335-341.

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

Murray, D.S., Schumacher, M.A., and Brennan, R.G. (2004). Crystal structures of QacR-diamidine complexes reveal additional multidrug-binding modes and a novel mechanism of drug charge neutralization. J. Biol. Chem. *279,* 14365-14371.

Naroditskaya, V., Schlosser, M.J., Fang, N.Y., and Lewis, K. (1993). An E. coli gene emrD is involved in adaptation to low energy shock. Biochem. Biophys. Res. Commun. *196,* 803-809.

Nikaido, H. (1998). Antibiotic resistance caused by gram-negative multidrug efflux pumps. Clin. Infect. Dis. *27* Suppl 1, S32-41.

Ogden, S., Haggerty, D., Stoner, C.M., Kolodrubetz, D., and Schleif, R. (1980). The Escherichia coli L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and negative regulation. Proc. Natl. Acad. Sci. U. S. A. *77,* 3346-3350.

Pao, S.S., Paulsen, I.T., and Saier, M.H.,Jr. (1998). Major facilitator superfamily. Microbiol. Mol. Biol. Rev. *62,* 1-34.

Paulsen, I.T., Brown, M.H., Dunstan, S.J., and Skurray, R.A. (1995). Molecular characterization of the staphylococcal multidrug resistance export protein QacC. J. Bacteriol. *177,* 2827-2833.

Paulsen, I.T., Brown, M.H., and Skurray, R.A. (1996). Proton-dependent multidrug efflux systems. Microbiol. Rev. *60,* 575-608.

Paulsen, I.T., Littlejohn, T.G., Radstrom, P., Sundstrom, L., Skold, O., Swedberg, G., and Skurray, R.A. (1993). The 3' conserved segment of integrons contains a gene associated with multidrug resistance to antiseptics and disinfectants. Antimicrob. Agents Chemother. *37,* 761-768.

Paulsen, I.T., and Skurray, R.A. (1994). The POT family of transport proteins. Trends Biochem. Sci. *19,* 404.

Paulsen, I.T., Skurray, R.A., Tam, R., Saier, M.H.,Jr, Turner, R.J., Weiner, J.H., Goldberg, E.B., and Grinius, L.L. (1996). The SMR family: a novel family of multidrug efflux proteins involved with the efflux of lipophilic drugs. Mol. Microbiol. *19,* 1167-1175.

Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl. Acad. Sci. U. S. A. *94,* 12366-12371.

Ploy, M.C., Grelaud, C., Martin, C., de Lumley, L., and Denis, F. (1998). First clinical isolate of vancomycin-intermediate Staphylococcus aureus in a French hospital. Lancet *351,* 1212.

Poole, K., Krebes, K., McNally, C., and Neshat, S. (1993). Multiple antibiotic resistance in Pseudomonas aeruginosa: evidence for involvement of an efflux operon. J. Bacteriol. *175,* 7363-7372.

Prive, G.G. (2007). Detergents for the stabilization and crystallization of membrane proteins. Methods *41,* 388-397.

Raman, P., Cherezov, V., and Caffrey, M. (2006). The Membrane Protein Data Bank. Cell Mol. Life Sci. *63,* 36-51.

Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., and Deber, C.M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proc. Natl. Acad. Sci. U. S. A. *106,* 1760-1765.

Rivnay, B., and Metzger, H. (1982). Reconstitution of the receptor for immunoglobulin E into liposomes. Conditions for incorporation of the receptor into vesicles. J. Biol. Chem. *257,* 12800-12808.

Sagne, C., Isambert, M.F., Henry, J.P., and Gasnier, B. (1996). SDS-resistant aggregation of membrane proteins: application to the purification of the vesicular monoamine transporter. Biochem. J. *316,* 825-831.

Saier, M.H.,Jr, Beatty, J.T., Goffeau, A., Harley, K.T., Heijne, W.H., Huang, S.C., Jack, D.L., Jahn, P.S., Lew, K., Liu, J. *et al*. (1999). The major facilitator superfamily. J. Mol. Microbiol. Biotechnol. *1,* 257-279.

Saier, M.H.,Jr, and Paulsen, I.T. (2001). Phylogeny of multidrug transporters. Semin. Cell Dev. Biol. *12,* 205-213.

Saier, M.H.,Jr, Tam, R., Reizer, A., and Reizer, J. (1994). Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport. Mol. Microbiol. *11,* 841-847.

Schagger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. *166,* 368-379.

Schleif, R. (1992). DNA looping. Annu. Rev. Biochem. *61,* 199-223.

Schuldiner, S., Lebendiker, M., and Yerushalmi, H. (1997). EmrE, the smallest ion-coupled transporter, provides a unique paradigm for structure-function studies. J. Exp. Biol. *200,* 335-341.

Schwarz, D., Klammt, C., Koglin, A., Lohr, F., Schneider, B., Dotsch, V., and Bernhard, F. (2007). Preparative scale cell-free expression systems: new tools for the large scale preparation of integral membrane proteins for functional and structural studies. Methods *41,* 355-369.

Siddiqui, A., Kerb, R., Weale, M.E., Brinkmann, U., Smith, A., Goldstein, D.B., Wood, N.W., and Sisodiya, S.M. (2003). Association of multidrug resistance in epilepsy with a polymorphism in the drug-transporter gene ABCB1. N. Engl. J. Med. *348,* 1442-1448.

Sigal, N., Lewinson, O., Wolf, S.G., and Bibi, E. (2007). E. coli multidrug transporter MdfA is a monomer. Biochemistry *46,* 5200-5208.

Sigal, N., Vardy, E., Molshanski-Mor, S., Eitan, A., Pilpel, Y., Schuldiner, S., and Bibi, E. (2005). 3D model of the Escherichia coli multidrug transporter MdfA reveals an essential membrane-embedded positive charge. Biochemistry *44*, 14870-14880.

Smith, D.B., and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene *67*, 31-40.

Tate, C.G. (2006). Comparison of three structures of the multidrug transporter EmrE. Curr. Opin. Struct. Biol. *16*, 457-464.

Tenover, F.C. (2006). Mechanisms of antimicrobial resistance in bacteria. Am. J. Med. *119*, S62-70.

Torizawa, T., Shimizu, M., Taoka, M., Miyano, H., and Kainosho, M. (2004). Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. J. Biomol. NMR *30*, 311-325.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. U. S. A. 76, 4350-4354.

Ueda, K., Cardarelli, C., Gottesman, M.M., and Pastan, I. (1987). Expression of a full-length cDNA for the human "MDR1" gene confers resistance to colchicine, doxorubicin, and vinblastine. Proc. Natl. Acad. Sci. U. S. A. *84*, 3004-3008.

Uttley, A.H., George, R.C., Naidoo, J., Woodford, N., Johnson, A.P., Collins, C.H., Morrison, D., Gilfillan, A.J., Fitch, L.E., and Heptonstall, J. (1989). High-level vancomycin-resistant enterococci causing hospital infections. Epidemiol. Infect. *103*, 173-181.

Van Horn, W.D., Kim, H.J., Ellis, C.D., Hadziselimovic, A., Sulistijo, E.S., Karra, M.D., Tian, C., Sonnichsen, F.D., and Sanders, C.R. (2009). Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. Science *324*, 1726-1729.

Vardy, E., Arkin, I.T., Gottschalk, K.E., Kaback, H.R., and Schuldiner, S. (2004). Structural conservation in the major facilitator superfamily as revealed by comparative modeling. Protein Sci. *13*, 1832-1840.

Wider, G., and Wuthrich, K. (1999). NMR spectroscopy of large molecules and multimolecular assemblies in solution. Curr. Opin. Struct. Biol. *9*, 594-601.

Wise, E.M.,Jr, and Park, J.T. (1965). Penicillin: its basic site of action as an inhibitor of a peptide cross-linking reaction in cell wall mucopeptide synthesis. Proc. Natl. Acad. Sci. U. S. A. *54*, 75-81.

Yamaguchi, A., Udagawa, T., and Sawai, T. (1990). Transport of divalent cations with tetracycline as mediated by the transposon Tn10-encoded tetracycline resistance protein. J. Biol. Chem. *265,* 4809-4813.

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

Yoshida, H., Bogaki, M., Nakamura, S., Ubukata, K., and Konno, M. (1990). Nucleotide sequence and characterization of the Staphylococcus aureus norA gene, which confers resistance to quinolones. J. Bacteriol. *172,* 6942-6949.

Yu, L., Sun, C., Song, D., Shen, J., Xu, N., Gunasekera, A., Hajduk, P.J., and Olejniczak, E.T. (2005). Nuclear magnetic resonance structural studies of a potassium channel-charybdotoxin complex. Biochemistry *44,* 15834-15841.