

2016

Structure, function and regulation of integral membrane transport proteins

Abhijith Kumar Radhakrishnan Pattathil Santha
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>

 Part of the [Chemistry Commons](#)

Recommended Citation

Radhakrishnan Pattathil Santha, Abhijith Kumar, "Structure, function and regulation of integral membrane transport proteins" (2016).
Graduate Theses and Dissertations. 15996.
<https://lib.dr.iastate.edu/etd/15996>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Structure, function and regulation of integral membrane transport proteins

by

Abhijith Kumar Radhakrishnan Pattathil Santha

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemistry

Program of Study Committee:
Edward W. Yu, Major Professor
Robert S. Houk
Drena L. Dobbs
Young-Jin Lee
Wenyu Huang

Iowa State University

Ames, Iowa

2016

Copyright © Abhijith Kumar Radhakrishnan Pattathil Santha, 2016. All rights reserved.

TABLE OF CONTENTS

ABSTRACT.....	iii
CHAPTER 1. GENERAL INTRODUCTION.....	1
CHAPTER 2. CRYSTAL STRUCTURE OF THE TRANSCRIPTIONAL REGULATOR RV0678 OF <i>MYCOBACTERIUM TUBERCULOSIS</i>	11
CHAPTER 3. CRYSTAL STRUCTURE OF THE <i>CAMPYLOBACTER JEJUNI</i> CMEC OUTER MEMBRANE CHANNEL	56
CHAPTER 4. CRYSTAL STRUCTURES OF NITRITE CHANNELS NAR1.1 AND NAR1.5 OF <i>CHLAMYDOMONAS REINHARDTII</i>	81
CHAPTER 5. GENERAL CONCLUSIONS	115
ACKNOWLEDGEMENTS.....	119
APPENDIX. CURRICULUM VITAE.....	121

ABSTRACT

Integral membrane transport proteins are essential for the transport of a wide variety of substrates such as ions, drugs and metabolites across the membranes of microorganisms. They are found in diverse locations such as outer and inner membranes of archaeal, prokaryotic and eukaryotic cells, as well as the outer layers of the membrane-bound organelles, including chloroplasts, endoplasmic reticulum and mitochondria. These proteins are capable of facilitating the transport of substrates both from outside the cell to inside, and vice versa, and they can be very specific, thereby playing a significant role in substrate selection. Integral membrane transport proteins also play important roles in regulation of transport, defense against drugs and antibiotics, and hemostasis. Thus, understanding the structural framework and detailed functional mechanisms of these proteins will contribute to our knowledge of biological processes.

In this study, we have investigated the structure, function and regulation of three different types of integral membrane transport proteins. Chapter 2 focuses on elucidating the structure and function of the transcriptional regulator Rv0678 of *Mycobacterium tuberculosis*, which negatively regulates the expression of the MmpS4-MmpL4 system, which plays a key role in the biosynthesis and transport of lipids from the cytoplasm towards the periplasm. Lipids play an important role in cell wall remodeling and permeability, thereby functioning in bacterial defense against antibiotics. We crystallized Rv0678 and identified the ligand bound to the protein as 2-stearoylglycerol. We used the structure and functional studies to elucidate a possible DNA

binding mechanism for the protein. Chapter 3 discusses the crystal structure of the outer membrane channel, CmeC, of CmeABC efflux system of *Campylobacter jejuni*. CmeABC is a multidrug efflux system that pumps out quinolones and bile acid derivatives. The crystal structure revealed the structural framework of the channel and captured it in closed conformation. Chapter 4 focuses on understanding the structural basis of nitrite transport into the chloroplast in *Chlamydomonas reinhardtii*. Formate-nitrite transporter family proteins, NAR1.1 and NAR1.5, were crystallized using hanging drop vapor diffusion. Stopped flow light scattering experiments were carried out to identify nitrite as the substrate for both proteins. Site-directed mutagenesis revealed key residues that form the channel and allowed us to propose a possible nitrite transport mechanism.

CHAPTER 1. GENERAL INTRODUCTION

Bacterial infections cause millions of deaths every year [1]. Over the past century, scientists have successfully employed novel antibiotics to treat these infections and combat the spread of the bacterial pathogens [1, 2]. However, over the years, bacterial pathogens have evolved various mechanisms to help them survive the effects of the antibiotics [3, 4]. The emergence of drug resistant strains of bacterial pathogens have rendered existing antibiotics less effective and signals the need for new therapeutic ways to combat them [5-9]. Understanding the mechanisms of bacterial drug resistance is of paramount importance in this regard. Chapter 2 and 3 of this dissertation focusses on understanding the bacterial drug resistance mechanisms in *Mycobacterium tuberculosis* and *Campylobacter jejuni* using X-ray crystallography and other biophysical techniques. Chapter 4 describes the structural and functional studies of nitrite channels NAR1.1 and NAR1.5 of *Chlamydomonas reinhardtii*.

Mycobacterium tuberculosis (*Mtb*) is the causative agent for tuberculosis, a disease that infects approximately 8 million people annually [10]. The emergence of new strains of *Mtb* that are resistant to existing antibiotics has made tuberculosis treatment cumbersome [5, 9]. This precarious situation has forced scientists across the world to look for new therapeutic pathways to combat the virulence of *Mtb* inside the host cell. One of these new pathways is restricting the availability of iron, an important nutrient for the virulence of bacterial pathogens such as *Mtb* [11-13]. The only source of iron for *Mtb* is the host cell, but the availability of iron to the pathogen is often restricted by the host immune system [14, 15].

Mtb employs two classes of molecules, mycobactins and carboxymycobactins, also called siderophores, to capture iron from the host cell [16]. Recent studies have revealed two proteins, mycobacterial membrane protein small (MmpS) and mycobacterial membrane protein large (MmpL), that are involved in scavenging iron from the host cell [17]. Although mycobacteria cluster phylogenetically with gram-positive prokaryotes, they are structurally more similar to gram-negative bacteria. Unlike the resistance nodulation and cell division (RND) efflux pumps of gram-negative bacteria, which pump out a wide range of substrates, MmpL-MmpS systems do not usually participate in antibiotic efflux [18]. Instead, they are believed to participate in *Mtb* iron acquisition [17]. There is direct evidence that mycobactins, the lipophilic siderophores of mycobacteria, efficiently extract intracellular macrophage iron [16]. The metal free siderophore of mycobacteria is diffusely associated with the macrophage membrane, ready for chelation. When iron is bound, the siderophore is ready to be transported into the mycobacteria. Two of the MmpS-MmpL systems, MmpS4-MmpL4 and MmpS5-MmpL5, are the most characterized members of this family [17]. Studies using cytoplasmic iron reporter indicate that *mmpS4/S5* mutants experience iron starvation even under high iron conditions. The double mutant was also impaired in biosynthesis and secretion of siderophores [17]. MmpS4 is localized in the periplasm [19] and interacts with the periplasmic loop of the *mmpL4* protein [17]. MmpS5 also binds to the periplasmic loop of MmpL5 [17]. Structurally MmpS proteins resemble the periplasmic adaptor proteins of the tripartite efflux systems of gram-negative bacteria.

The expression of the MmpS5-MmpL5 efflux system is controlled by a transcriptional regulator, Rv0678, which helps *Mtb* adapt to changes in the environment. Rv0678 is located downstream of the *mmpS5-mmpL5* operon and belongs to the MarR family of regulators [20]. The

MarR family of regulators is found in bacteria and archaea and regulates numerous biological functions. MarR regulators possess a DNA binding domain formed by a typical winged helix-turn-helix binding motif, a dimerization domain, and a substrate binding region located between the DNA binding domain and the dimerization domain [21]. When not liganded, the MarR regulator is capable of binding to DNA. Understanding the structural details of Rv0678 will be key to understanding the mechanism of regulation and function of the MmpS5-MmpL5 siderophore efflux system and will aid in designing novel drugs to combat the highly pathogenic *Mtb*.

Campylobacter jejuni causes the acute diarrheal disease, human enterocolitis, especially in infants, and causes about 400 million infections every year [22]. Macrolide and quinolone derivatives are used as antimicrobial agents to treat the *C. jejuni* infection [23]. In a major public health concern, recent studies reveal increased resistance of the pathogen to the existing antibiotics [4]. Antibiotic efflux systems, as one of these resistance mechanisms, extrude structurally diverse antimicrobial agents out of the bacterial cell [24, 25]. Identifying the antimicrobial efflux systems of *C. jejuni* and elucidating its structural details is of significant clinical importance. Scientists have characterized an efflux pump, *cmeABC*, encoded by a three gene operon, which contributes significantly to multidrug resistance in *C. jejuni* [26-28]. Studies have shown that inactivation of *CmeABC* greatly reduces the frequency of emergence of quinolone resistant mutants, whereas its overexpression increases the frequency [29]. Sequence analysis of *cmeABC* revealed significant resemblance to multidrug efflux systems of other gram-negative bacteria. It consists of a periplasmic membrane fusion protein, *CmeA*, an inner membrane efflux pump of the resistance-nodulation-cell division (RND) superfamily, *CmeB*, and an outer membrane channel protein,

CmeC. The efflux complex spans the entire cell membrane. Understanding the structure of the individual components of CmeABC will be valuable in designing novel antibiotics.

Chlamydomonas reinhardtii is a photosynthetic single celled green alga widely distributed in soil and fresh water. It possesses a large cup-shaped chloroplast that enables it to perform photosynthesis. *C. reinhardtii* requires carbon dioxide (CO₂) and bicarbonate (HCO₃⁻), two utilizable forms of inorganic carbon (C_i) for photosynthesis [30, 31]. For aquatic photosynthetic organisms, such as *C. reinhardtii*, the supply of CO₂ can change dramatically. *C. reinhardtii* possesses a carbon concentrating mechanism (CCM) to survive in carbon limiting environments [32]. Membrane proteins belonging to different classes take part in CCM in *C. reinhardtii*.

An important step in the nitrate assimilation in *C. reinhardtii* also takes place in the chloroplast, where nitrite is reduced to ammonium, which is incorporated into the carbon skeleton [33]. Nitrite hemostasis is significant because nitrite accumulation can be toxic to most cells. Nitrite can be reduced to the nitric oxide (NO) radical, which is cytotoxic and can cause damage to DNA and iron-sulfur clusters [34-36]. Therefore, structural studies of membrane proteins involved in nitrite transport are important. .

Gene expression analysis of CCM induction in *C. reinhardtii* has revealed that multiple transporters are involved in C_i uptake [37]. *Nar1* codes for proteins belonging to the formate-nitrite transporter (FNT) family and encodes six genes; *Nar1.1*, *Nar1.2*, *Nar1.3*, *Nar1.4*, *Nar1.5* and *Nar1.6*. Out of the six FNT family proteins, NAR1.1, NAR1.2 (also called LCIA), and NAR1.5 are believed to be located at the chloroplast envelope membrane [38]. They are proposed to serve as nitrite/bicarbonate channels that conduct these anions across the chloroplast envelope membrane. These three proteins share high sequence similarity with each other and are predicted

to have a long transit peptide and six transmembrane regions [38]. Proteins belonging to the FNT family are found commonly in bacteria, but are less frequently found in eukaryotic organisms such as fungi, yeast, algae and protozoa [38]. Understanding the structure and function of these proteins will help us understand the transport mechanism of nitrite and bicarbonate across the chloroplast membrane.

Thesis organization

Chapter 1 is a brief general introduction that provides background information about the transcriptional regulator Rv0678 of *Mycobacterium tuberculosis*, the outer membrane channel CmeC of the CmeABC efflux system of *Campylobacter jejuni*, and nitrite channels, NAR1.1 and NAR1.5, of the photosynthetic green algae *Chlamydomonas reinhardtii*.

Chapter 2 is a manuscript published in *Journal of Biological Chemistry*, describing the role of Rv0678 in regulating the expression MmpS5-MmpL5 efflux system in *M. tuberculosis*. The ligand bound to the Rv0678 dimer was identified using gas chromatography–mass spectrometry (GC-MS). Electrophoretic mobility shift assay (EMSA) and fluorescence polarization studies of Rv0678 with the promoter region of the *rv0678-mmpS5* operon revealed a possible DNA binding mechanism for the MarR family regulators. Cloning of *rv0678* was done by HT Lei. Expression, purification and crystallization was carried out by N. Kumar and myself. Model building and structural refinement was done by TH Chou and CC Su.

Chapter 3 is a manuscript published in *Protein Science*, describing the crystal

structures of *C. jejuni* CmeC, an outer membrane protein of CmeABC, a multidrug efflux system belonging to RND family. The crystal structures consists of a barrel and an α -helical bundle, typical of the RND family. I carried out the cloning, expression, purification and crystallization of *cmeC* in collaboration with CC Su and N Kumar. Model building and structural refinement was done by CC Su.

Chapter 4 investigates the crystal structures of nitrite channels, NAR1.1 and NAR1.5, of *C. reinhardtii*, which belong to FNT family. The proteins are located in the chloroplast envelope membrane and conduct nitrite ion from cytoplasm to stroma. Stopped flow light scattering experiments were done to identify the substrate as well as the key residues involved in anion transport. Cloning of *nar1.1* and *nar1.5*, expression, purification and crystallization of NAR1.1 and NAR1.5 were carried out by me. Model building and structural refinement was done by TH Chou. Stopped flow light scattering experiment and data analysis was done by TH Chou and me.

Chapter 5 summarizes the conclusions and significance of these studies and outlines directions for future research.

References

1. Walsh C. 2003 Antibiotics: actions, origins, resistance. Washington, DC: ASM Press.
2. Jones DS, Podolsky SH, Greene JA. 2012. The burden of disease and the changing task of medicine. N Engl J Med 366, 2333–2338.

3. Hughes, J. M., and F. C. Tenover. 1997. Approaches to limiting emergence of antimicrobial resistance in bacteria in human populations. *Clin. Infect. Dis.* 24(Suppl. 1):S131-S135
4. Levy, S. B. 2000. Antibiotic and antiseptic resistance: impact on public health. *Pediatr. Infect. Dis. J.* 19:S120-S122.
5. Frieden T. R., Sterling T., Pablos-Mendez A., Kilburn J. O., Cauthen G. M., Dooley S.W. 1993. The emergence of drug-resistant tuberculosis in New York City. *N. Engl. J. Med.* 328, 521–556
6. Pillay M., Sturm A. W. 2007. Evolution of the extensively drug-resistant 5/LAM4/KZN strain of *Mycobacterium tuberculosis* in KwaZulu-Natal, South Africa. *Clin. Infect. Dis.* 45, 1409–1414
7. Goldman R. C., Plumley K. V., Laughon B. E. 2007. The evolution of extensively drug resistant tuberculosis (XDR-TB): history, status and issues for global control. *Infect. Disord. Drug Targets* 7, 73–91
8. Udawadia Z. F., Amale R. A., Ajbani K. K., Rodrigues C. 2012. Totally drug-resistant tuberculosis in India. *Clin. Infect. Dis.* 54, 579–581
9. Iseman M. D. 1993. Treatment of multidrug-resistant tuberculosis. *N. Engl. J. Med.* 329, 784–791
10. Maartens G., Wilkinson R. J. 2007. Tuberculosis. *Lancet* 370, 2030–2043
11. Flo, T. H., Smith K. D., Sato S., Rodriguez D. J., Holmes M. A., Strong R. K., Akira S., Aderem A. 2004. Lipocalin 2 mediates an innate immune response to bacterial infection by sequestering iron. *Nature* 432, 917–921

12. Ferreras, J. A., Ryu, J. S., Lello, F. D., Tan, D. S., Quadri, L. E. N. 2005. Small-molecule inhibition of siderophore biosynthesis in *Mycobacterium tuberculosis* and *Yersinia pestis*. *Nat. Chem. Bio.* 1, 29–32
13. Fluckinger, M., Hass, H., Merschak, P., Glasgow, B. J., Redl, B. 2004. Human tear lipocalin exhibits antimicrobial activity by scavenging microbial siderophores. *Antimicrob. Agents Chemother.* 48, 3367–3372
14. Weiss, G., Wachter, H., Fuchs, D. 1995. Linkage of cell-mediated immunity to iron metabolism. *Immunol. Today* 16, 495–500
15. Vergne, A. F., Walz, A. J., Miller, M. J. 2000. Iron chelators from mycobacteria (1954–1999) and potential therapeutic applications. *Nat. Prod. Rep.* 17, 99–116
16. Luo M., Fadeev E. A., Groves J. T. 2005. Mycobactin-mediated iron acquisition within macrophages. *Nat. Chem. Bio.* 1, 149 – 153
17. Wells R. M., Jones C. M., Xi Z., Speer A., Danilchanka O., Doornbos K. S., Sun P., Wu F., Tian C., Niederweis M. 2013. Discovery of a Siderophore Export System Essential for Virulence of *Mycobacterium tuberculosis*. *PLoS Pathog* 9(1): e1003120
18. Milano A., Pasca M. R., Provvedi R., Lucarelli A. P., Manina G., Ribeiro A. L., Manganello R., Riccardi G. 2009. Azole resistance in *Mycobacterium tuberculosis* is mediated by the MmpS5–MmpL5 efflux system. *Tuberculosis* 89, 84–90
19. Deshayes C., Bach H., Euphrasie D., Attarian R., Coureuil M., Sougakoff W., Laval F., Av-Gay Y., Daffe M., Etienne G., Reytrat J. M. 2010. MmpS4 promotes glycopeptidolipids biosynthesis and export in *Mycobacterium smegmatis*. *Mol Microbiol* 78: 989–1003

20. Ellison D. W., Miller V. L. 2006. Regulation of virulence by members of the MarR/SlyA family. *Curr. Opin. Microbiol.* 9, 153–159
21. Alekshun M. N., Levy S. B., Mealy T. R., Seaton B. A., Head J. F. 2001. The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 Å resolution. *Nat. Struct. Biol.* 8, 710–714
22. Ruiz-Palacios G. M. 2007. The health burden of *Campylobacter* infection and the impact of antimicrobial resistance: Playing chicken. *Clin Infect Dis* 44:701–703.
23. Engberg J., Aarestrup F. M., Taylor D. E., Gerner-Smidt P., Nachamkin I. 2001. Quinolone and macrolide resistance in *Campylobacter jejuni* and *C. coli*: resistance mechanisms and trends in human isolates. *Emerg. Infect. Dis.* 7:24-34
24. Coleman, K., M. Athalye, A. Clancey, M. Davison, D. J. Payne, C. R. Perry, and I. Chopra. 1994. Bacterial resistance mechanisms as therapeutic targets. *J. Antimicrob. Chemother.* 33:1091-1116
25. Dever, L. A., T. S. Dermody. 1991. Mechanisms of bacterial resistance to antibiotics. *Arch. Intern. Med.* 151:886-895.
26. Lin J., Michel L. O., Zhang Q. 2002. CmeABC functions as a multidrug efflux system in *Campylobacter jejuni*. *Antimicrob Agents Chemother* 46:2124–2131
27. Lin J., Sahino O., Michel L. O., Zhang Q. 2003. Critical role of multidrug efflux pump CmeABC in bile resistance and in vivo colonization of *Campylobacter jejuni*. *Infect Immun* 71:4250–4259
28. Pumbwe L., Piddock L. J. 2002. Identification and molecular characterisation of CmeB, a *Campylobacter jejuni* multidrug efflux pump. *FEMS Microbiol Lett* 206:185–189

29. Yan M., Sahin O., Lin J., Zhang Q. 2006. Role of the CmeABC efflux pump in the emergence of fluoroquinolone-resistant *Campylobacter* under selection pressure. *J Antimicrob Chemother* **58**:1154–1159.
30. Spalding M. H. 2008. Microalgal carbon-dioxide-concentrating mechanisms: *Chlamydomonas* inorganic carbon transporters. *J Exp Bot* **59**, 1463-1473
31. Wang Y., Stessman D. J., Spalding M. H. 2015. The CO₂ concentrating mechanism and photosynthetic carbon assimilation in limiting CO₂ : how *Chlamydomonas* works against the gradient. *Plant J* **82**, 429-448
32. Giordano M., Beardall J., Raven J. A. 2005. CO₂ concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu Rev Plant Biol* **56**, 99-131
33. Rexach J., Fernández E., Galván A. 2000. The *Chlamydomonas reinhardtii* Nar1 Gene Encodes a Chloroplast Membrane Protein Involved in Nitrite Transport. *The Plant Cell*, Vol. 12, 1441–1453
34. Gilberthorpe N. J., Poole R. K. 2008. Nitric oxide homeostasis in *Salmonella typhimurium*: Roles of respiratory nitrate reductase and flavohemoglobin. *J Biol Chem* **283**(17):11146–11154
35. Wink D. A., Kasprzak K. S., Maragos C. M., Elespuru R. K., Misra M., Dunams T. M., Cebula T. A., Koch W. H., Andrews A. W., Allen J. S. 1991. DNA deaminating ability and genotoxicity of nitric oxide and its progenitors. *Science* **254**(5034):1001–1003.
36. Hibbs J. B. Jr., Taintor R. R., Vavrin Z., Rachlin E. M. 1988. Nitric oxide: A cytotoxic activated macrophage effector molecule. *Biochem Biophys Res Commun* **157**(1):87–94.

37. Burow M. D., Chen Z. Y., Mouton T. M., Moroney J. V. 1996. Isolation of cDNA clones of genes induced upon transfer of *Chlamydomonas reinhardtii* cells to low CO₂. *Plant Mol. Biol.* 33:443-448
38. Mariscal V., Moulina P., Orselb M., Millerb A. J., Fernández E., Galvana A. Differential Regulation of the *Chlamydomonas* Nar1 Gene Family by Carbon and Nitrogen. *Protist*. Volume 157, Issue 4, 24 October 2006, Pages 421–433

**CHAPTER 2. CRYSTAL STRUCTURE OF THE TRANSCRIPTIONAL
REGULATOR RV0678 OF *MYCOBACTERIUM TUBERCULOSIS***

A paper published in J. Biol. Chem. 2014 289: 16526-. doi:10.1074/jbc.M113.538959

Abhijith Radhakrishnan^{1,ψ}, Nitin Kumar^{1,ψ}, Catherine C. Wright², Tsung-Han Chou³, Marios L. Tringides¹, Jani Reddy Bolla¹, Hsiang-Ting Lei¹, Kanagalaghatta R. Rajashankar⁴, Chih-Chia Su³, Georgiana E. Purdy², and Edward W. Yu^{1,3,ϑ}

¹From the Department of Chemistry, Iowa State University, Ames, IA 50011, USA

²Department of Molecular Microbiology and Immunology, Oregon Health and Sciences University, Portland, OR 97239, USA

³Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA

⁴NE-CAT and Department of Chemistry and Chemical Biology, Cornell University, Bldg. 436E, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne. IL 60439, USA

^ψA.R. and N.K. contributed equally to this work.

To whom correspondence should be addressed: Edward W. Yu, Department of Chemistry, Department of Physics & Astronomy, Iowa State University, Ames, IA, USA, Tel: (515) 294-4955; E-mail: ewyu@isatate.edu

Abstract

Recent work demonstrates that the Mycobacterial membrane protein large (MmpL) transporters are dedicated to the export of mycobacterial lipids for cell wall biosynthesis. An MmpL transporter frequently works with an accessory protein, belonging to the Mycobacterial membrane protein small (MmpS) family, to transport these key virulence factors. One such efflux system in *Mycobacterium tuberculosis* is the MmpS5-MmpL5 transporter. The expression of MmpS5-MmpL5 is controlled by the MarR-like transcriptional regulator Rv0678, whose open reading frame is located downstream of the *mmpS5-mmpL5* operon. To elucidate the structural basis of Rv0678 regulation, we have determined the crystal structure of this regulator, to 1.64 Å resolution, revealing a dimeric two-domain molecule with an architecture similar to members of the MarR family of transcriptional regulators. Rv0678 is distinct from other MarR regulators in that its DNA-binding and dimerization domains are clustered together. These two domains seemingly cooperate to bind an inducing ligand that we identified as 2-palmitoylglycerol, which is a fatty acid glycerol ester. The structure also suggests that the conformational change leading to substrate-mediated derepression is primarily caused by a rigid body rotational motion of the entire DNA-binding domain of the regulator towards the dimerization domain. This movement results in a conformational state that is incompatible with DNA binding. We demonstrate using electrophoretic mobility shift assays that Rv0678 binds to the *mmpS5-mmpL5*, *mmpS4-mmpL4* and the *mmpS2-mmpL2* promoters. Binding by Rv0678 was reversed upon addition of the ligand. These findings provide new insight into the mechanisms of gene regulation in the MarR family of regulators.

Introduction

Tuberculosis (TB) is one of the oldest described diseases and remains a significant global problem with more than eight million new cases reported annually (1). The World

Health Organization estimates that one-third of the world's population is infected with *M. tuberculosis*, and most of these individuals have latent TB (2). TB treatments are notoriously difficult and are compromised by the emergence of multiple drug-resistant (MDR), extensively drug-resistant (XDR) and totally drug-resistant (TDR) bacterial strains (3-7). The development of drug-resistant *M. tuberculosis* strains is a major threat that challenges global prospects for TB control.

Although mycobacteria cluster phylogenetically with Gram-positive prokaryotes, they are structurally more similar to Gram-negative bacteria. These mycobacteria are protected by an outer lipid bilayer made of mycolic acids and a cell envelope composed of non-covalently bound lipids and glycolipids. The unique structure and composition of the cell wall differentiates this highly pathogenic microorganism from other prokaryotes. The mycobacterial cell wall plays a crucial role in the host-pathogen interface on several levels (8). First, the thick, greasy cell wall acts as an effective layer of protection, providing intrinsic resistance to antibiotics and bactericidal components of the host immune response. Second, the surface-exposed polyketide and glycoconjugate lipids of the *M. tuberculosis* cell wall are associated with bacterial virulence (9-12).

The genome of *M. tuberculosis* H37Rv contains 15 genes that encode the resistance-nodulation-cell division (RND) proteins designated MmpL transporters (13, 14). Unlike the RND-type efflux pumps of Gram-negative bacteria, MmpL proteins do not usually participate in antibiotic efflux. Instead, there is strong evidence that these MmpL proteins are responsible for exporting fatty acids and lipidic elements of the cell wall (8, 9, 12, 15-17). Five *mmpL* genes are located adjacent to genes coding for proteins involved in fatty acid or polyketide synthesis, suggesting that the MmpL membrane proteins transport these key virulence factors (9, 10). Similar to RND proteins of Gram-negative bacteria, the MmpL transporters of *M. tuberculosis* are believed to work in conjunction with accessory proteins. Specifically, MmpL

transporters form complexes with the MmpS-family proteins in order to export cell wall lipid constituents (18). Five genes encoding MmpS proteins are adjacent to genes encoding MmpL proteins (8, 13). Work in the model organism *M. smegmatis* demonstrated that MmpS4 was required for bacterial sliding motility and biofilm formation (19). That the *mmpS4* and *mmpL4* mutants had similar phenotypes underscore a coordinated function for cognate MmpS-MmpL proteins.

Our efforts have focused on elucidating how *M. tuberculosis* transport systems are regulated. We previously crystallized the Rv3066 efflux regulator both in the absence and presence of bound substrate (20). Our data indicated that ligand binding triggers a rotational motion of the regulator, which in turn releases the cognate DNA and induces the expression of the Mmr efflux pump (20). We report here the crystal structure of the Rv0678 regulator, which has been proposed to control the transcriptional regulation of the MmpS5-MmpL5 transport system. Rv0678 belongs to the MarR family of regulators, which are found ubiquitously in bacteria and archaea and control various important biological processes, such as resistance to antimicrobials, sensing of oxidative stress agents and regulation of virulence factors (21). Typically, the MarR-family regulators are dimeric in form and their protein sequences are poorly conserved. However, these proteins share a common fold, consisting of a helical dimerization domain and two winged helix-turn-helix (wHtH) DNA-binding domains within the dimer (22). Our data suggest that fatty acid glycerol esters are the natural ligands of the Rv0678 regulator. Electrophoretic mobility shift assay indicates that Rv0678 binds promoters of the *mmpL2*, *mmpL4* and *mmpL5* operons. These results emphasize the importance of the Rv0678 regulator, which appears to regulate multiple MmpL transport systems.

Experimental procedures

Cloning of *rv0678*

The *rv0678* ORF from genomic DNA of *M. tuberculosis* strain H37Rv was amplified by PCR using the primers 5'-CCATGGGCAGCGTCAACGACGGGGTC-3' and 5'-GGATCCTCAGTGATGATGATGATGATGGTCGTCCTCTCCGGTTCG-3' to generate a product that encodes a Rv0678 recombinant protein with a 6xHis tag at the C-terminus. The corresponding PCR product was digested with *NcoI* and *BamHI*, extracted from the agarose gel, and inserted into pET15b as described by the manufacturer (Merck KGaA, Darmstadt, Germany). The recombinant plasmid (pET15b Ω *rv0678*) was transformed into DH5 α cells and the transformants were selected on LB agar plates containing 100 μ g/ml ampicillin. The presence of the correct *rv0678* sequence in the plasmid construct was verified by DNA sequencing.

Expression and purification of Rv0678

Briefly, the full-length Rv0678 protein containing a 6xHis tag at the C-terminus was overproduced in *E. coli* BL21(DE3) cells possessing pET15b Ω *rv0678*. Cells were grown in 6 L of Luria Broth (LB) medium with 100 μ g/ml ampicillin at 37°C. When the OD₆₀₀ reached 0.5, the culture was treated with 0.2 mM isopropyl- β -D-thiogalactopyranoside (IPTG) to induce Rv0678 expression, and cells were harvested within 3 h. The collected bacterial cells were suspended in 100 ml ice-cold buffer containing 20 mM Na-HEPES (pH 7.2) and 200 mM NaCl, 10 mM MgCl₂ and 0.2 mg DNase I (Sigma-Aldrich). The cells were then lysed with a French pressure cell. Cell debris was removed by centrifugation for 45 min at 4°C and 20,000 rev/min. The crude lysate was filtered through a 0.2 μ m membrane and was loaded onto a 5 ml Hi-Trap Ni²⁺-chelating column (GE Healthcare Biosciences, Pittsburgh, PA) pre-equilibrated with 20 mM Na-HEPES (pH 7.2) and 200 mM NaCl. To remove unbound proteins

and impurities, the column was first washed with six column volumes of buffer containing 50 mM imidazole, 250 mM NaCl, and 20 mM Na-HEPES (pH 7.2). The Rv0678 protein was then eluted with four column volume of buffer containing 300 mM imidazole, 250 mM NaCl, and 20 mM Na-HEPES (pH 7.2). The purity of the protein was judged using 12.5% SDS-PAGE stained with Coomassie Brilliant Blue. The purified protein was extensively dialyzed against buffer containing 100 mM imidazole, 250 mM NaCl, and 20 mM Na-HEPES (pH 7.5), and concentrated to 20 mg/ml.

Crystallization of Rv0678

All crystals of the 6xHis Rv0678 regulator were obtained using hanging-drop vapor diffusion. The Rv0678 crystals were grown at room temperature in 24-well plates with the following procedures. A 2 μ l protein solution containing 20 mg/ml Rv0678 protein in 20 mM Na-HEPES (pH 7.5), 250 mM NaCl and 100 mM imidazole was mixed with a 2 μ l of reservoir solution containing 28% polyethylene glycol (PEG) 1000, 0.1 M Na-acetate (pH 4.0), 0.2 M NaCl and 10% glycerol. The resultant mixture was equilibrated against 500 μ l of the reservoir solution. Crystals grew to a full size in the drops within two weeks. Typically, the dimensions of the crystals were 0.2 mm x 0.05 mm x 0.05 mm. Cryoprotection was achieved by raising the PEG 1000 concentration stepwise to 35% with a 3.5% increment in each step. Crystals of the tungsten derivative was prepared by incubating the crystals of Rv0678 in solution containing 28% PEG 1000, 0.1 M Na-acetate (pH 4.0), 0.2 M NaCl, 10% glycerol and 1 mM $(\text{NH}_4)_2\text{W}_6(\mu\text{-O})_6(\mu\text{-Cl})_6\text{Cl}_6$ for 24 hours at 25°C.

Data collection, structural determination and refinement

All diffraction data were collected at 100K at beamline 24ID-E located at the Advanced Photon Source, using an ADSC Quantum 315 CCD-based detector. Diffraction data were

processed using DENZO and scaled using SCALEPACK (23). The crystals of Rv0678 belong to the space group *P1* (Table 1). Based on the molecular weight of Rv0678 (18.34 kDa), the asymmetric unit is expected to contain four regulator molecules with a solvent content of 45.26%. Six tungsten cluster sites were identified using SHELXC and SHELXD (24) as implemented in the HKL2MAP package (25). Single isomorphous replacement with anomalous scattering (SIRAS) was employed to obtain experimental phases using the program MLPHARE (26, 27). The resulting phases were then subjected to density modification and NCS averaging using the program PARROT (28). The phases were of excellent quality and allowed for tracing of most of the molecule in PHENIX AutoBuild (29), which led to an initial model with over 90% amino acid residues containing side-chains. The remaining part of the model was manually constructed using the program Coot (30). Then, the model was refined using PHENIX (29) leaving 5% of reflections in Free-R set. Iterations of refinement using PHENIX (29) and CNS (31) and model building in Coot (30) lead to the current model, which consists of two dimers (totally 587 residues in the asymmetric unit) with excellent geometrical characteristics (Table 1).

Identification of fortuitous ligand

To identify the nature of the bound ligand in crystals of Rv0678, we used gas chromatography coupled with mass spectrometry (GC-MS). The Rv0678 crystals were extensively washed with the crystallization buffer and transferred into deionized water. The mixture was then incubated at 100°C for 5 min, then chloroform was added into the mixture to a final concentration of 80% (v/v) to denature the protein and allow for the extraction of ligand. GC-MS analysis indicated that the mass of the bound ligand was octadecanoic acid, 2-hydroxyl-1-(hydroxymethyl)ethyl ester, also called 2-palmitoylglycerol.

Virtual ligand screening using AutoDock Vina

AutoDock Vina (32) was used for virtual ligand screening of a variety of compounds. The docking area was assigned visually to cover the internal cavity of the Rv0678 dimer. A grid of $35 \text{ \AA} \times 35 \text{ \AA} \times 35 \text{ \AA}$ with 0.375 \AA spacing was calculated around the docking area for all atom types presented in the DrugBank (33) and ZINC (34) libraries using AutoGrid. The iterated local search global optimizer algorithm was employed to predict the binding free energies for these compounds.

Isothermal titration calorimetry for ligand binding – We used isothermal titration calorimetry (ITC) to determine the binding affinity of 1-stearoyl-*rac*-glycerol (an isomer of 2-palmitoylglycerol) to the purified Rv0678 regulator. Measurements were performed on a VP-Microcalorimeter (MicroCal, Northampton, MA) at $25 \text{ }^\circ\text{C}$. Before titration, the protein was thoroughly dialyzed against buffer containing 10 mM Na-phosphate pH 7.2, 100 mM NaCl and 0.001% DDM. The protein concentration was determined using the Bradford assay. The dimeric Rv0678 sample was then adjusted to a final concentration of $400 \text{ }\mu\text{M}$ and served as the titrant. The ligand solution contained $20 \text{ }\mu\text{M}$ 1-stearoyl-*rac*-glycerol, 10 mM Na-phosphate pH 7.2, 100 mM NaCl and 0.001% DDM. The protein and ligand samples were degassed before they were loaded into the cell and syringe. Binding experiments were carried out with the ligand solution (1.5 ml) in the cell and the protein solution as the injectant. Ten microliter injections of the ligand solution were used for data collection.

Injections occurred at intervals of 300 s, and the duration time of each injection was 20 s. Heat transfer ($\mu\text{cal/s}$) was measured as a function of elapsed time (s). The mean enthalpies measured from injection of the ligand in the buffer were subtracted from raw titration data before data analysis with ORIGIN software (MicroCal). Titration curves were fitted by a nonlinear least squares method to a function for the binding of a ligand to a macromolecule. Nonlinear regression fitting to the binding isotherm provided us with the equilibrium binding

constant ($K_A = 1/K_D$) and enthalpy of binding (ΔH). Based on the values of K_A , the change in free energy (ΔG) and entropy (ΔS) were calculated with the equation: $\Delta G = -RT \ln K_A = \Delta H - T \Delta S$, where T is 273 K and R is 1.9872 cal/K per mol. Calorimetry trials were also carried out in the absence of Rv0678 in the same experimental conditions. No change in heat was observed in the injections throughout the experiment.

Electrophoretic mobility shift assay

Probes were amplified from the H37Rv genome using the primers listed in Table 2. All probes were labeled with Digoxigenin using the Roche DIG Gel Shift kit. For EMSA analysis, 12 nM Dig-labeled probe and the indicated micromolar concentrations of protein were incubated for 45 minutes at room temperature in the Roche binding buffer modified by the addition of 0.25 mg/mL herring sperm DNA, and 0.75 mg/mL poly(d[I-C]). For ligand competition studies, 1-stearoyl-*rac*-glycerol (an isomer of 2-palmitoylglycerol) (Sigma-Aldrich) was resuspended in hot acetone and added to EMSA reactions at 1 μ M final concentration. Competition reactions were performed at 37°C. All reactions were resolved on a 6% native polyacrylamide gel in TBE buffer, transferred to nylon membrane and Dig-labeled DNA-protein complexes detected following the manufacturer's recommendations. Chemiluminescent signals were acquired using an ImageQuant LAS 4000 (GE).

Dye primer based DNase I footprint assay

DNase I footprinting was performed as described by Zianni et al. (35). The 296 bp Rv0678-mmpS5 probe was generated by PCR using the primers 6FAM-Rv0678-F and HEX-Rv0678-R. Gel purified, fluorescently labeled probe (0.6 pmol) was incubated with either 1 μ M Rv0678 or BSA for 30 min at room temperature in standard EMSA binding buffer. After incubation, 10 mM MgCl₂ and 5 mM CaCl₂ were added to the reaction mixture in a final

volume of 50 μ l. Then, 0.0025 U of DNase I (Thermo) was added and incubated for 5 minutes at room temperature. Digested DNA fragments were purified with QIAquick PCR Purification columns (Qiagen) and eluted in 20 μ l water. Digested DNA samples were analyzed at the Center for Genome Research and Biocomputing at Oregon State University. Purified DNA (2 mL) was mixed with HiDi formamide and GeneScan-500 LIZ size standards (Applied Biosystems) and analyzed using an Applied Biosystems 3730 DNA analyzer.

The 296 bp fragment was sequenced with the primers 6FAM-Rv0678-F and HEX-Rv0678-R, respectively, using the Thermo Sequenase Dye Primer Manual Cycle Sequencing Kit according to the manufacturer's instructions. Each reaction was diluted 5-fold in water, and 4 μ l was added to 5.98 μ l HiDi formamide and 0.02 μ l GeneScan-500 LIZ size standard. Samples were analyzed using the 3730 DNA analyzer and electropherograms aligned using the GENEMAPPER software (version 5.0, Applied Biosystems).

Site-directed mutagenesis

Site-directed point mutations on residues D90 and R92, which are expected to be critical for DNA binding, were performed to generate the single point mutants D90A and R92A. The primers used for these mutations are listed in Table 3. All oligonucleotides were purchased from (Integrated DNA Technologies, Inc., Coralville, IA) in a salt-free grade.

Fluorescence polarization assay for DNA binding

Fluorescence polarization assays were used to determine the affinity for DNA binding by Rv0678 and its mutants. Both the 26-bp oligodeoxynucleotide and fluorescein labeled oligodeoxynucleotide were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). These oligodeoxynucleotides contain the consensus 18-bp putative promoter DNA sequence (TTTCAGAGTACAGTGAAA) for Rv0678. The sequences of the

oligodeoxynucleotides were 5'-CAGATTCAGAGTACAGTGAACTTG-3' and 5'-F-CAAGTTTCACTGTACTCTGAAATCTG-3', where F denotes the fluorescein which was covalently attached to the 5' end of the oligodeoxynucleotide by a hexamethylene linker. The 26-bp fluoresceinated ds-DNA was prepared by annealing these two oligodeoxynucleotides together. Fluorescence polarization experiment was done using a DNA binding solution containing 10 mM Na-phosphate (pH 7.2), 100 mM NaCl, 5 nM fluoresceinated DNA, and 1 μ g of poly(dI-dC) as non-specific DNA. The protein solution containing 2,500 nM dimeric Rv0678 or Rv0678 mutant and 5 nM fluoresceinated DNA was titrated into the DNA binding solution until the millipolarization (mP) become unchanged. All measurements were performed at 25°C using a PerkinElmer LS55 spectrofluorometer equipped with a Hamamatsu R928 photomultiplier. The excitation wavelength was 490 nm, and the fluorescence polarization signal (in ΔP) was measured at 525 nm. Each titration point recorded was an average of 15 measurements. Data were analyzed using the equation, $P = \{(P_{\text{bound}} - P_{\text{free}})[\text{protein}]/(K_D + [\text{protein}])\} + P_{\text{free}}$, where P is the polarization measured at a given total protein concentration, P_{free} is the initial polarization of free fluorescein-labeled DNA, P_{bound} is the maximum polarization of specifically bound DNA, and [protein] is the protein concentration. The titration experiments were repeated for three times to obtained the average K_D value. Curve fitting was accomplished using the program ORIGIN (OriginLab Corporation, Northampton, MA).

Results and discussion

Overall structure of Rv0678

M. tuberculosis Rv0678 belongs to the MarR family of regulators. It possesses 165 amino acids, sharing 14% and 15% protein sequence identity with MarR (22) and OhrR (36) (Fig. 1). The crystal structure of Rv0678 was determined to a resolution of 1.64 Å using single

isomorphous replacement with anomalous scattering (SIRAS) (Table 1). Four molecules of Rv0678 are found in the asymmetric unit, which assemble as two independent dimers (Fig. 2). Superimposition of these two dimers gives an RMSD of 0.8 Å over 271 C^α atoms, indicating that their conformations are nearly identical to each other.

The structure of Rv0678 (Fig. 3) is quite distinct in comparison with the known structures of the MarR-family regulators (22, 36-39). Each subunit of Rv0678 is composed of six α -helices and two β -strands: α 1 (17-31), α 2 (36-47), α 3 (55-62), α 4 (66-79), β 1 (82-85), β 2 (94-97), α 5 (101-127) and α 6 (132-160) (Fig. 1). The monomer is L-shaped with the shorter side forming a DNA-binding domain. However, the longer side contributes to an extended long arm, creating a dimerization domain for the regulator. Residues 34-99, which include α 2, α 3, α 4, β 1 and β 2, are responsible for constructing the DNA-binding domain. The dimerization domain of Rv0678 is generated by residues 16-32 and 101-160, which cover α 1, α 5 and α 6 of the protomer. Each protomer of Rv0678 is approximately 55 Å tall, 35 Å wide and 35 Å thick.

As a member of the MarR family of regulators, the DNA-binding domain of Rv0678 features a typical winged helix-turn-helix (wHtH) binding motif. The two anti-parallel β 1 and β 2 strands are found to generate a β -hairpin structure, which also forms the wing of the DNA-binding domain. The crystal structure of the OhrR-DNA complex (36) showed that this β -hairpin directly participates to contact the double-stranded DNA and is critical for repressor-operator interactions. Another important component of the wHtH motif for DNA recognition is helix α 4. In the OhrR-DNA complex (36), the corresponding α -helix is found to bind within the deep major groove of the B-DNA. Protein sequence alignment suggests that Rv0678 contains three conserved amino acids common among members of the MarR family. These three residues, R84, D90 and R92, are located within the DNA-binding domain of the regulator (Fig. 1), and are likely important for protein-DNA interactions. Among them, D90 and R92

are positioned within the β -hairpin of the wing. The corresponding amino acids located at the winged loop region of the ST1710 regulator play a major role in regulator-promoter interactions (39).

The Rv0678 crystal structure reveals that helices $\alpha 1$, $\alpha 5$ and $\alpha 6$ are involved in the formation of the dimer. Specifically, helices $\alpha 5$, $\alpha 6$, $\alpha 5'$ and $\alpha 6'$ (where the prime denotes the next subunit) form intertwined helical bundles and constitute the dimerization domain. Helices $\alpha 6$ and $\alpha 6'$ are oriented in an anti-parallel fashion and form the scaffold of the dimer (Fig. 3). Extensive hydrophobic interactions are observed at the interface between the two subunits of the regulator. In addition, Y147 and Y159' and their identical counter pair perform aromatic stacking interactions, securing the dimeric organization. Additional salt bridges between R32 and E115' and between E106 and R109' (as well as their counter pairs) stabilize the binding.

Perhaps the most striking difference between the structures of Rv0678 and other MarR-family members is the relative orientation of the DNA-binding and dimerization domains. The structures of MarR (22), OhrR (36) and MexR (37, 38) suggest that helices $\alpha 4$ and $\alpha 4'$ orient approximately perpendicular to the pseudo two-fold axis of the dimeric regulators. However, our crystal structure of Rv0678 depicts these two helices more or less in parallel with the dimer's pseudo two-fold axis. Similar orientation of helix $\alpha 4$ has also been found in the structure of the *Vibrio cholerae* AphA transcriptional activator (40). This conformation is not compatible and does not allow the regulator to interact with the B-form DNA. To bind its cognate DNA, the Rv0678 regulator must undergo a large conformational movement that reorients the DNA-binding domain such that the positions of helices $\alpha 4$ and $\alpha 4'$ can be matched with the two consecutive major groove of the promoter DNA. Based on the OhrR-DNA (36) and ST1710-DNA (39) crystal complexes, we predict that the entire DNA-binding domain of Rv0678, including $\alpha 2$, $\alpha 3$, $\alpha 4$, $\beta 1$ and $\beta 2$, has to rotate downward by $\sim 70^\circ$ with respect to $\alpha 5$

of the dimerization domain before DNA binding (Fig. 4). If this is the case, then the loop region between $\beta 2$ and $\alpha 5$ form the hinge for this rotational motion.

Rv0678 was liganded

Unexpectedly, a large extra electron density was found at the interface between the DNA-binding and dimerization domains of Rv0678, indicating the existence of a fortuitous bound ligand co-purified and co-crystallized with the regulator (Fig. 5). Thus, this region is also a substrate-binding site. To identify the unknown bound ligand, gas chromatography coupled with mass spectrometry (GC-MS) was applied to investigate the Rv0678 crystals (Fig. 6). The result suggests that the fortuitous ligand is 2-palmitoylglycerol, also called octadecanoic acid, 2-hydroxyl-1-(hydroxymethyl)ethyl ester, which contains 21 carbons with the molecular formula $C_{21}H_{42}O_4$. That this fatty acid glycerol ester is co-purified with the Rv0678 regulator suggests that fatty acid glycerol esters may be the natural substrates for this protein.

The propanetriol of the bound 2-palmitoylglycerol is completely buried within the dimer interface, leaving the tail portion of its elongated octadecanoate hydrophobic carbon chain oriented at the entry point of this binding site. This orientation facilitates the contribution of R32 and E106' to form two hydrogen bonds with the glycerol head group of the fatty acid. The backbone oxygen of F79' also participates to create the third hydrogen bond with this glycerol head group. In addition, the carbonyl oxygen of the octadecanoate group contributes to make another hydrogen bond with R109, securing the binding. Interestingly, Rv0678 further anchors the bound fatty acid molecule through hydrophobic interactions with residues F79, F79' and F81'. Therefore, the binding of 2-palmitoylglycerol in Rv0678 is extensive: Within 4.5 Å of the bound fatty acid glycerol ester, 20 amino acids contact this molecule (Table 4). It should be noted that residues F79, F79' and F81' belong to helices $\alpha 4$ and $\alpha 4'$. In the OhrR-

DNA structure (36), the corresponding $\alpha 4$ and $\alpha 4'$ helices were buried within the two consecutive major grooves, directly contacting the promoter DNA. Thus, we suspect that helices $\alpha 4$ and $\alpha 4'$ have dual responsibilities in the Rv0678 regulator. They form the DNA-binding site for operator DNA as well as substrate-binding site for inducing ligands.

In the second Rv0678 dimer of the asymmetric unit, it is also found that a 2-palmitoylglycerol molecule is bound within the corresponding substrate-binding site. Residues contributed to form this binding site are nearly identical but with a slightly different subset of amino acids in comparison with those of the first Rv0678 dimer described above (Table 4).

Virtual ligand library screening

Virtual ligand screening was then performed to elucidate the nature of protein-ligand interactions in the Rv0678 regulator. The 2-palmitoylglycerol binding site was chosen as a substrate binding cavity for this docking study. AutoDock Vina (32) was used to screen small molecules listed in the DrugBank (33) and ZINC (34) libraries. Vina utilizes the iterated local search global optimizer algorithm, which results in predicted binding free energies for these compounds ranging from -13.8 to +20 kcal/mol. Of the 70,000 screened compounds, it is predicted that the best substrate for Rv0678 is the heterocyclic compound diethyl-[(5E)-5-(6,8,9,10-tetrahydro-5H-benzo[c]xanthen-11-ylmethylene)-7,8-dihydro-6H-xanthen-3-yl]. Table 5 lists the top three substrates, which have the lowest predicted binding free energies, for the Rv0678 regulator.

As the crystal structure of Rv0678 depicts that a fatty acid glycerol ester is bound within the substrate binding site of this regulator, Vina (32) was also used to examine if these fatty acids are able to interact with Rv0678. As a positive control, the molecule 2-palmitoylglycerol was docked into the substrate-binding site of this regulator, resulting in a predicted binding free energy of -7.6 kcal/mol. Vina was then used to screen for 2,500 different fatty acids.

Based on the lowest predicted binding free energies, the top three compounds in this class was selected and listed in Table 6, where 18-[8-chloro-1-(hydroxymethyl)-6-phenyl-4H-[1,2,4]triazolo[4,3-a][1,4]benzodiazepin-4-yl]octadecanoic acid is the best compound for Rv0678 binding among these fatty acids.

Rv0678-ligand interaction

The binding affinity of 1-stearoyl-*rac*-glycerol for the Rv0678 regulator was then determined using ITC, which obtained a binding affinity constant, K_A , of $4.9 \pm 0.4 \times 10^5 \text{ M}^{-1}$. The titration is characterized by a negative enthalpic contribution, which gives rise to a hyperbolic binding curve (Fig. 7). The thermodynamic parameters of binding of 1-stearoyl-*rac*-glycerol to Rv0678 display enthalpic (ΔH) and entropic (ΔS) contributions of $-1.0 \pm 0.1 \text{ kcal/mol}$ and $22.5 \text{ cal}\cdot\text{mol}\cdot\text{deg}^{-1}$, respectively. Interestingly, the molar ratio for this binding reaction based on ITC is one Rv0678 dimer per ligand. This ligand-binding experiment confirms that Rv0678 is capable of recognizing fatty acid glycerol esters.

Electrophoretic mobility shift assay

To demonstrate direct transcriptional regulation, we performed EMSAs using a probe corresponding to the intergenic region between *mmpS5* and *rv0678* (Fig. 8a). This probe shifted in a concentration-dependent manner (Fig. 8b). This result is consistent with previous reports of altered *mmpS5/mmpL5* gene expression in *M. bovis* BCG spontaneous *rv0678* mutants (13). Preliminary CHIPSeq data from the TB Systems Biology Consortium suggests Rv0678 regulates expression of additional genes (41). We designed additional probes to experimentally demonstrate binding of Rv0678 to the promoter regions of *mmpS2-mmpL2*, *mmpS4-mmpL4* and *rv0991-0992*. Probes are depicted schematically in Fig. 8a. We also saw concentration-dependent binding of Rv0678 to these two probes (Fig. 8b). As a control,

EMSAs were performed in the presence of non-labeled probes. Release of Dig-labeled probe was observed consistent with specific binding of Rv0678 to the *rv0678-mmpS5*, *rv0505-mmpS2* and *mmpL4* probes (Fig. 8c). Using the sequence of the six probes that shifted, we identified a putative consensus binding sequence for Rv0678 using the MEME algorithm (42) (Fig. 8e). Rv0678 co-crystallized with a ligand whose binding renders the protein unable to bind DNA. Addition of 1-stearoyl-*rac*-glycerol (an isomer of 2-palmitoylglycerol) to the EMSA reaction buffer reduced Rv0678 binding to a target promoter probe (Fig. 8c).

Dye primer based DNase I footprint assay

To further refine the binding site of Rv0678 in the *rv0678-mmpS5* intergenic region, DNase I footprint assay was performed on the Rv0678-*mmpS5* probe using established methods (35). Electropherograms in Figure 9 show the DNA sequence bound by Rv0678. The control protein BSA did not result in DNA protection at the same concentration. Interestingly, the region bound by Rv0678 includes the start codon of the *rv0678* gene (underlined nucleotides, Fig. 9b). The bound sequence contains a potential inverted repeat motif (GAACGTCACAGATTTCA ... N₈ ... TGAAACTTGTGAGCGTCAAC).

Rv0678-DNA interaction

Fluorescence polarization-based assay was carried out to study the interaction between Rv0678 and the 26-bp DNA containing the 18-bp putative promoter DNA sequence (TTTCAGAGTACAGTGAAA). Our footprint assay has suggested that this promoter DNA sequence was protected by the Rv0678 regulator. Fig. 10a illustrates the binding isotherm of Rv0678 in the presence of 5 nM fluoresceinated DNA. The titration experiment indicated that this regulator binds the 26-bp promoter DNA with a dissociation constant, K_D , of 19.6 ± 3.0

nM. The binding data also indicate that Rv0678 binds its cognate DNA with a stoichiometry of one Rv0678 dimer per ds-DNA.

In addition, fluorescence polarization was used to determine the binding affinities of this 26-bp DNA by the Rv0678 mutants D90A and R92A. These two residues are located within the β -hairpin of the wHtH motif of the N-terminal DNA-binding domain. In ST1710, the corresponding two residues are critical for regulator-promoter interactions. Interestingly, our measurements indicate that the K_D values of the D90A-DNA and R92A-DNA complexes are 113.3 ± 16.8 and 86.0 ± 7.4 nM (Fig. 10b and c), revealing the DNA binding affinities for these mutants are significantly weaker than that of the native Rv0678 regulator. Like ST1710, our experimental results suggest that residues D90 and R92 are important for DNA recognition.

With the rising incidence of drug resistant strains of *M. tuberculosis*, it is increasingly important to understand the molecular mechanisms underlying virulence and drug resistance of this pathogen. This knowledge will inform the development of new strategies to combat TB. In this report, we describe the crystal structure the Rv0678 transcriptional regulator, which controls the expression level of the MmpS5-MmpL5, MmpS4-MmpL4 and MmpS2-MmpL2 transport systems. MmpS4 and MmpS5 contribute to siderophore export, but the substrate of MmpL2 is not known (15). Fortuitously, the structure of Rv0678 was resolved in complex with a 2-palmitoylglycerol molecule, suggesting that fatty acid glycerol esters are the natural substrates for the Rv0678 transcriptional regulator. Further work is required to demonstrate if this ligand is structurally related to the substrate of either efflux system, or how its availability changes in different environments and mycobacterial growth phases. The crystal structure of the 2-palmitoylglycerol-Rv0678 complex probably provides a snapshot of the ligand-binding state of this regulator, whereby both the DNA-binding and dimerization domains are recruited to participate for ligand binding. In this case, the DNA-binding domain must bend upward and shift towards the dimerization domain to accommodate the bound ligand. As crystallized, the

regulator is incompatible with the operator DNA. When the inducing ligand is removed from the ligand-binding site, freeing helices $\alpha 4$ and $\alpha 4'$ to rotate downward and shift away from the dimerization domain, this conformational state should be compatible with the B-DNA and allow for DNA binding.

References

1. Maartens, G., and Wilkinson R. J. (2007) Tuberculosis. *The Lancet* **370**, 2030-2043.
2. World Health Organization (2010) Fact sheet no. 104: tuberculosis. Available at <http://www.who.int/mediacentre/factsheets/fs104/en/index.html>.
3. Frieden, T. R., Sterling, T., Pablos-Mendez, A., Kilburn, J. O., Cauthen, G. M., and Dooley, S.W. (1993) The emergence of drug-resistant tuberculosis in New York City. *N. Engl. J. Med.* **328**, 521-556.
4. Pillay, M., and Sturm, A. W. (2007) Evolution of the extensively drug-resistant F15/LAM4/KZN strain of Mycobacterium tuberculosis in KwaZulu-Natal, South Africa. *Clin. Infect. Dis.* **45**, 1409-1414.
5. Goldman, R. C., Plumley, K. V., and Laughon, B. E. (2007) The evolution of extensively drug resistant tuberculosis (XDR-TB): history, status and issues for global control. *Infect. Disord. Drug Targets* **7**, 73-91.
6. Udawadia, Z. F., Amale, R. A., and Rodrigues, C. (2012) Totally drug-resistant tuberculosis in India. *Clin. Infect. Dis.* **54**, 579-581.
7. Iseman M. D. (1993) Treatment of multidrug-resistant tuberculosis. *N. Engl. J. Med.* **329**, 784-791.
8. Domenech, P., Reed, M. B., and Barry, C. E. III. (2005). Contribution of the Mycobacterium tuberculosis MmpL Protein Family to Virulence and Drug Resistance. *Infect Immun.* **73**, 3492-3501.

9. Jain, M., and Cox, J. S. (2005). Interaction between polyketide synthase and transporter suggests coupled synthesis and export of virulence lipid in *M. tuberculosis*. *PLoS Pathog.* **1**, 12-19.
10. Cox, J. S., Chen, B., McNeil, M., and Jacob, W. R. Jr. (1999) Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* **402**, 79-83.
11. Brennan, P. J., and Nikaido, H. (1995) The envelope of mycobacteria. *Annu. Rev. Biochem.* **64**, 29-63.
12. Converse, S. E., Mougous, J. D., Leavell, M. D., Leary, J. A., Bertozzi, C. R., and Cox, J. S. (2003) MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. *Proc. Natl. Acad. Sci. USA* **100**, 6121-6126.
13. Milano, A., Pasca, M. R., Provvedi, R., Lucarelli, A. N., Manina, G., Ribeiro, A. L., Manganelli, R., and Riccardi, G. (2009) Azole resistance in *Mycobacterium tuberculosis* is mediated by the MmpS5–MmpL5 efflux system. *Tuberculosis* **89**, 84-90.
14. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E. *et al.* (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544.
15. Wells, R. M., Jones, C. M., Xi, Z., Speer, A., Danilchanka, O. *et al.* (2013) Discovery of a siderophore export system essential for virulence of *Mycobacterium tuberculosis* *PLoS Pathog.* **9**, 1-14.
16. Pacheco, S., Hsu, F.-F., Powers, K. M., and Purdy, G. E. (2013) MmpL11 transports mycolic acid-containing lipids to the mycobacterial cell wall and contributes to biofilm formation in *M. smegmatis*. *J. Biol. Chem.* **288**, 24213-24222.
17. Jacobs, W.R ., Cox, J. S., Chen, B., and McNeil, M. (1999). Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* **402**, 79-83.

18. Tekaia, F., Gordon, S. V., Garnier, T., Brosch, R., Barrell, B. G., and Cole, S. T. (1999) Analysis of the proteome of *Mycobacterium tuberculosis* in silico. *Tuber. Lung. Dis.* **79**, 329–342.
19. Deshayes, C., Bach, H. Euphrasie, D., Attarian, R., Coureuil, M, Sougakoff, W. Laval, F., Av-Gay, Y., Daffé, M. Etienne, G. and Reyrat, J. M. (2010) MmpS4 promotes glycopeptidolipids biosynthesis and export in *Mycobacterium smegmatis*. *Mol. Microbiol.* **78**, 989-1003.
20. Bolla, J. R., Do, S. V., Long, F., Dai, L., Su, C.-C., Lei, H. T., Chen, X., Gerkey, J. E., Murphy, D. C., Rajashankar, K. R., Zhang, Q., and Yu, E.W. (2012) Structural and functional analysis of the transcriptional regulator Rv3066 of *Mycobacterium tuberculosis*. *Nucleic Acids Res.* **40**, 9340-9355.
21. Ellison, D. W., and Miller, V. L. (2006) Regulation of virulence by members of the MarR/SlyA family. *Curr. Opin. Microbiol.* **9**, 153-159.
22. Alekshun, M. N., Levy, S. B., Mealy, T. R., Seaton, B. A., and Head, J. F. (2001) The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3Å resolution. *Nat. Struct. Biol.* **8**, 710-714.
23. Otwinowski, Z., and Minor, M. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
24. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr.* **D58**, 1772–1779.
25. Pape, T., and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843-844.
26. Otwinowski, Z. (1991) ML-PHARE, CCP4 Proc. 80-88 (Daresbury Laboratory, Warrington, UK).

27. Collaborative Computational Project No. 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.* **D50**, 760-763.
28. Cowtan, K. (2010) Recent developments in classical density modification. *Acta Crystallogr.* **D66**, 470-478.
29. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCroy, A. J., Moriarty, N. W. et al. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr.* **58**, 1948-1954.
30. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr.* **D60**, 2126.
31. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905-921.
32. Trott, O., and Olson, A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comp. Chem.* **31**, 455-461.
33. Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., and Wishart, D. S. (2011) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucl. Acids Res.* **39**, D1035-D1041.
34. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., and Coleman, R.G. (2012) ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **52**, 1757-1768.

35. Zianni, M., Tessanne, K., Merighi, M., Laguna, R. and Tabita, F.R. (2006) Identification of the DNA bases of a DNase I footprint by the use of dye primer sequencing on an automated capillary DNA analysis instrument. *J. Biomolecular Techniques* **17**, 103-113.
36. Hong, M., Fuangthong, M., Helmann, J. D., and Brennan, R. G. (2005) Structure of an OhrR-OhrA Operator Complex Reveals the DNA Binding Mechanism of the MarR Family. *Mol. Cell* **20**, 131-141.
37. Lim, D., Poole, K., and Strynadka, N. C. J. (2002) Crystal structure of the MexR repressor of the *mexRAB-oprM* multidrug efflux operon of *Pseudomonas aeruginosa*. *J. Biol. Chem.* **277**, 29253-29259.
38. Wilke, M. S., Heller, M., Creagh, A. L., Haynes, C. A., McIntosh, L. P., Poole, K., and Strynadka, N. C. J. (2008) The crystal structure of MexR from *Pseudomonas aeruginosa* in complex with its antirepressor ArmR. *Proc. Natl. Acad. Sci. USA* **39**, 14832-14837.
39. De Silva, R. S., Kovachikova, G., Lin, W., Taylor, R. K., Skorupski, K., and Kull, F. J. (2005) Crystal structure of the virulence gene activator AhpA from *Vibrio cholerae* reveals it is a novel member of the winged helix transcription factor superfamily. *J. Biol. Chem.* **280**, 13779-13783.
40. Kumarevel, T., Tanaka, T., Umehara, T., and Yokoyama, S. (2009) ST1710–DNA complex crystal structure reveals the DNA binding mechanism of the MarR family of regulators. *Nucl. Acids Res.* **37**, 4723-4735.
41. Reddy, T. B., Riley, R., Wymore, F., Montgomery, P., Decaprio, D., Engels, R., Gellesch, M., Hubble, J., Jen, D., Jin H, Koehrsen, M., Larson, L., Mao, M., Nitzberg, M., Sisk, P., Stolte, C., Weiner, B., White, J., Zachariah, Z. K., Sherlock, G., Galagan, J. E., Ball, C. A., and Schoolnik, G. K. (2009) TB database: an integrated platform for tuberculosis research. *Nucl. Acids Res.* **37**, D499-D508.

42. Bailey, T. L., and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28-36.

Acknowledgements

This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by an award GM103403 from the National Institutes of General Medical Sciences. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. We are grateful to Louis Messerle at University of Iowa for providing us the $(\text{NH}_4)_2\text{W}_6(\mu\text{-O})_6(\mu\text{-Cl})_6\text{Cl}_6$ complex used in this study.

Footnotes

This work was supported by NIH Grants R01AI087840 (G.E.P.) and R01GM086431 (E.W.Y.). Coordinates and structural factors for the structure of Rv0678 have been deposited at the RCSB Protein Data Bank with an accession code 4NB5.

Figures and captions

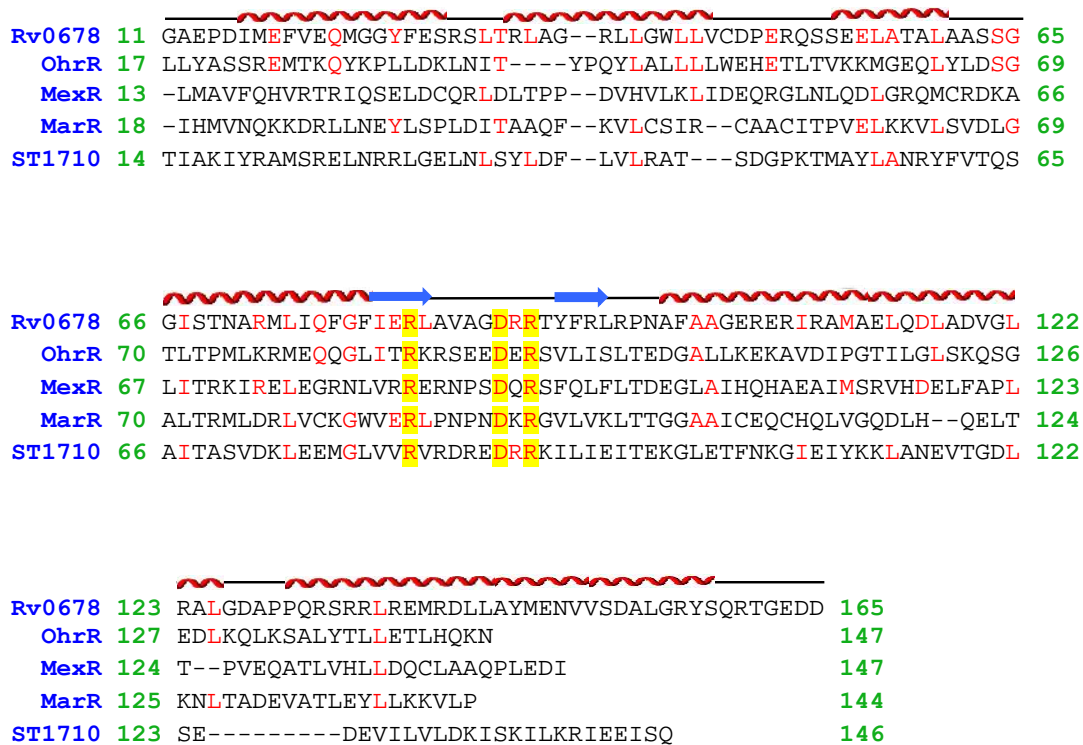


Figure 1. Protein sequence alignment of the MarR family of regulators. Alignment of the amino acid sequences of *Mycobacterium tuberculosis* Rv0678, *Bacillus subtilis* OhrR, *Pseudomonas aeruginosa* MexR, *Escherichia coli* MarR and *Sulfolobus tokodaii* ST1710. The alignment is done using FFAS03. The topology of *M. tuberculosis* Rv0678 is shown at the top. The three conserved amino acids are highlighted with yellow bars.

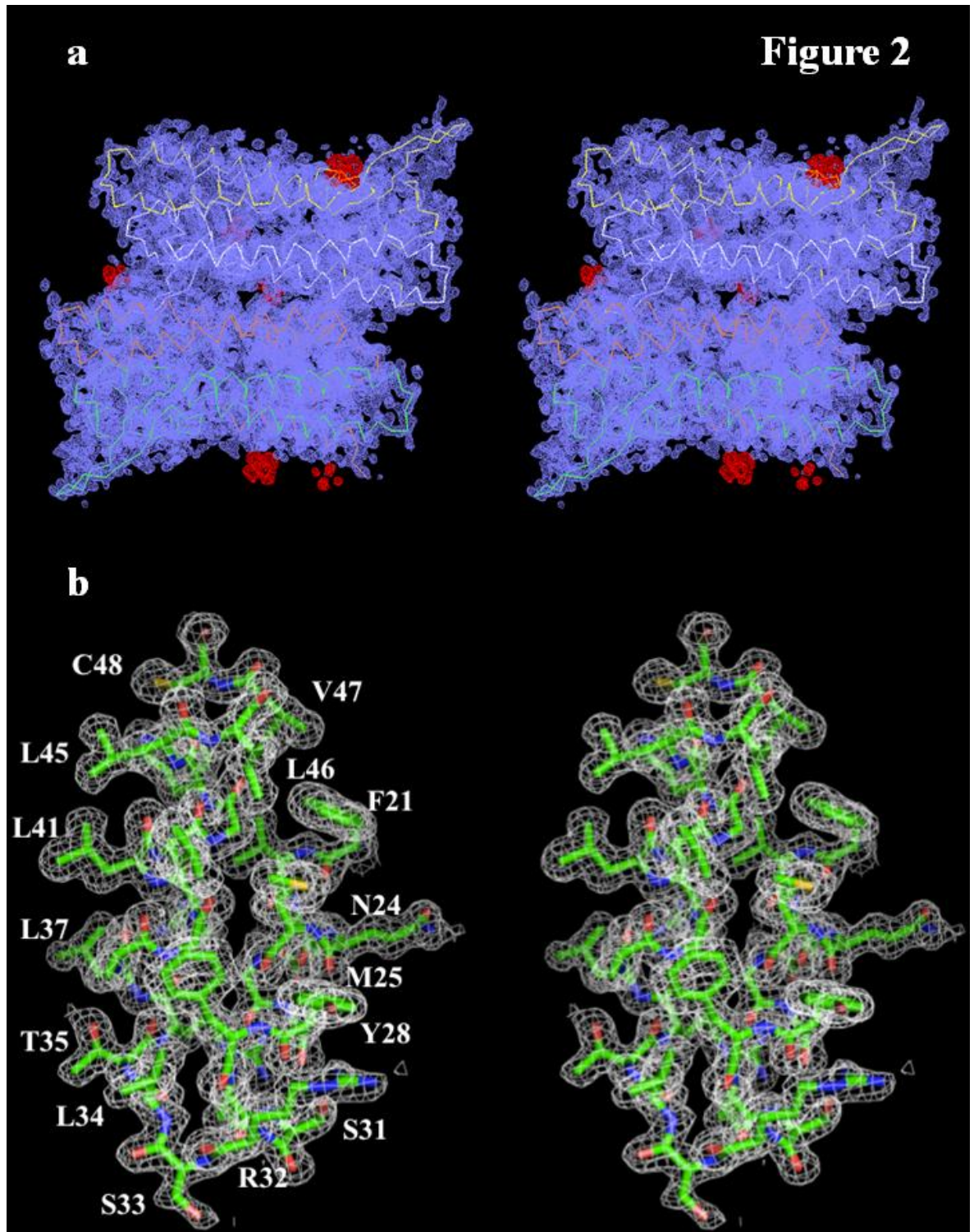
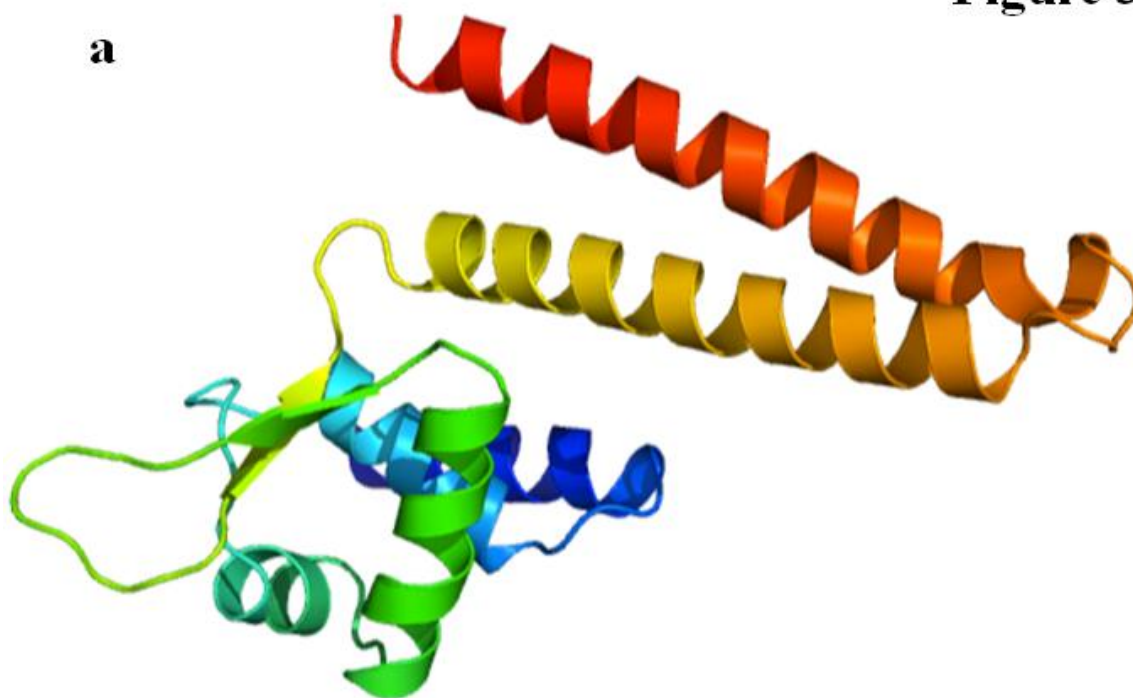


Figure 2. Stereo view of the experimental electron density maps of Rv0678 at a resolution of 1.64 Å. (a) The electron density maps are contoured at 1.2 σ . The C α traces of the two

Rv0678 dimers in the asymmetric unit are in yellow, light blue, tv red and lime green.

Anomalous signals of the six $W_6(\mu-O)_6(\mu-Cl)_6Cl_6^{2-}$ cluster sites (contoured at 4σ) found in the asymmetric unit are colored red. (b) Representative section of electron density in the vicinity of helices $\alpha 1$ and $\alpha 2$. The solvent-flattened electron density (50-1.64 Å) is contoured at 1.2σ and superimposed with the final refined model (green, carbon; red, oxygen; blue nitrogen; yellow, sulfur).

Figure 3



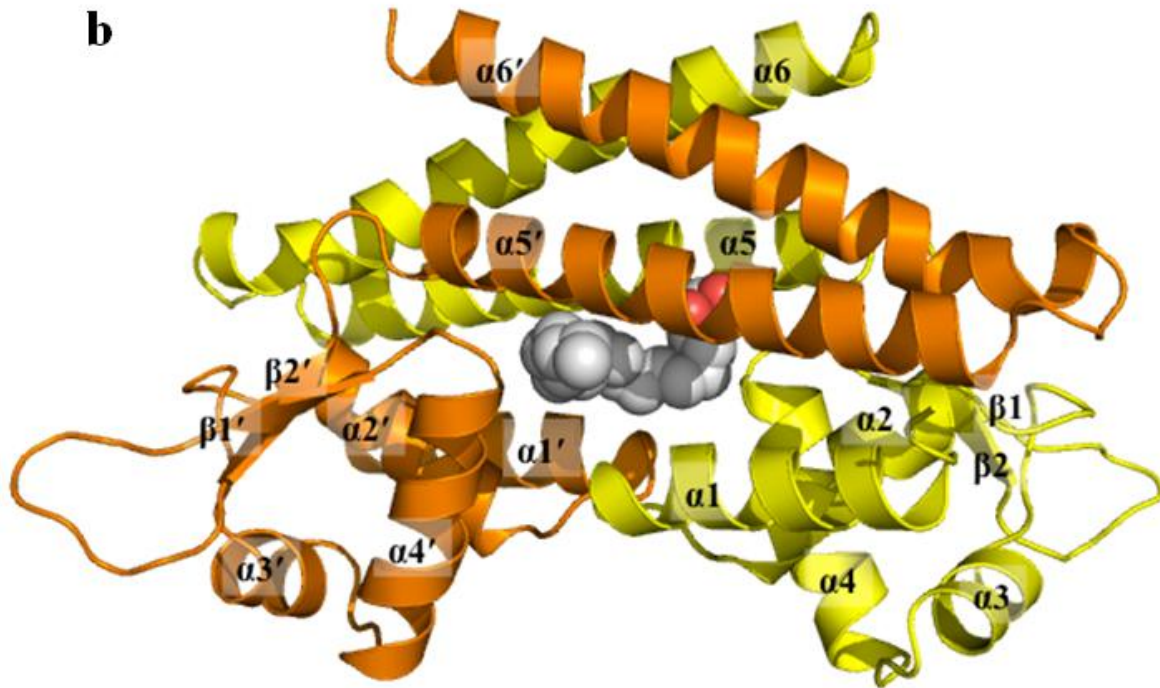


Figure 3. Structure of the *M. tuberculosis* Rv0678 regulator. (a) Ribbon diagram of a protomer of Rv0678. The molecule is colored using a rainbow gradient from the N-terminus (blue) to the C-terminus (red). (b) Ribbon diagram of the Rv0678 dimer. Each subunit of Rv0678 is labeled with a different color (yellow and orange). The bound 2-palmitoylglycerol within the dimer is shown in sphere form (gray, carbon; red, oxygen). The figure was prepared using PyMOL (<http://www.pymol.sourceforge.net>).

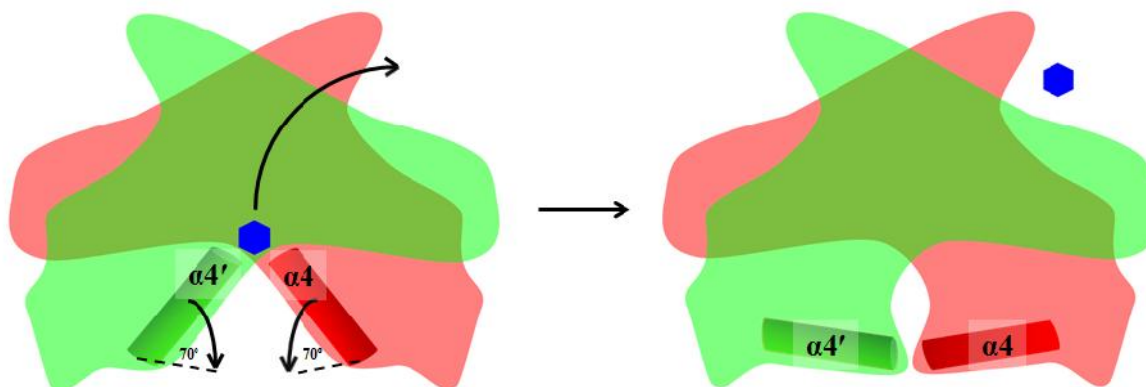
Figure 4

Figure 4. Rigid-body rotation of the DNA-binding domain of Rv0678. This is a schematic representation illustrating the conformational change of Rv0678 between the ligand bound and unbound structures. Helices $\alpha 4$ and $\alpha 4'$ of the DNA-binding domain are indicated. The ligand is colored blue.

Figure 5

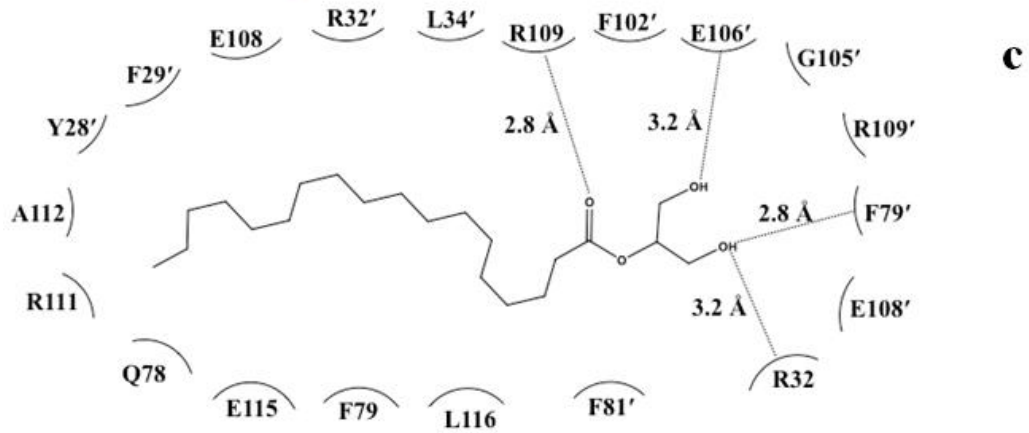
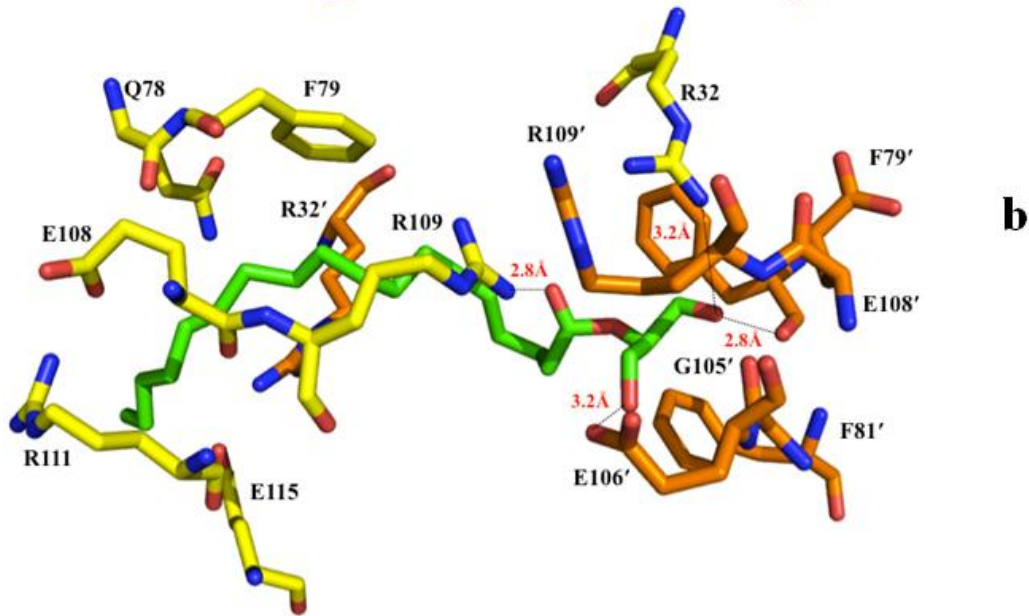
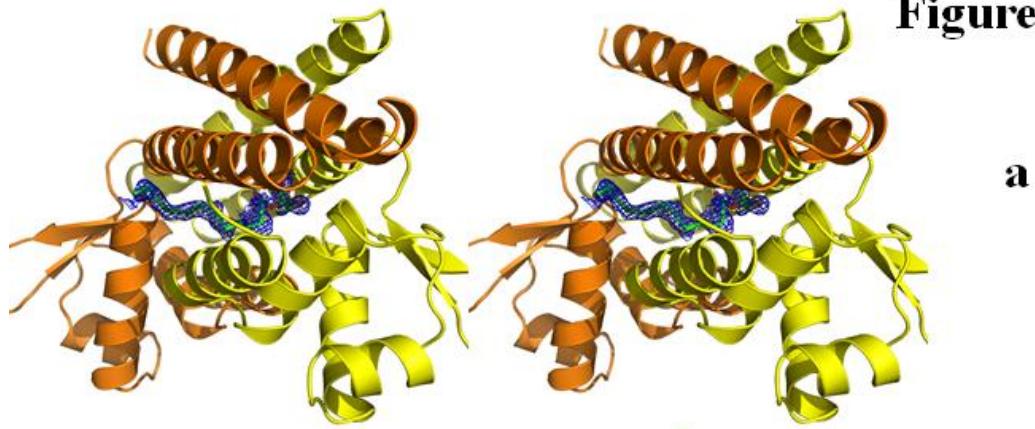


Figure 5. Simulated annealing electron density maps and the 2-palmitoylglycerol binding site. (a) Stereo view of the simulated annealing electron density map of the bound 2-palmitoylglycerol within the Rv0678 dimer (the orientation corresponds to the side view of Fig. 1b). The bound 2-palmitoylglycerol is shown as a stick model (green, carbon; red, oxygen). The simulated annealing $2F_o-F_c$ electron density map is contoured at 1.2σ (blue mesh). The left and right subunits of Rv0678 are shown as orange and yellow ribbons. (b) The 2-palmitoylglycerol binding site. Amino acid residues within 3.9 \AA from the bound 2-palmitoylglycerol (green, carbon; red, oxygen) are included. The side chains of selected residues from the right subunit of Rv0678 in Figure 1b are shown as yellow sticks (yellow, carbon; blue, nitrogen; red, oxygen). Residues from the next subunit of Rv0678 are shown as orange sticks (orange, carbon; blue, nitrogen; red, oxygen). (c) Schematic representation of the Rv0678 and 2-palmitoylglycerol interactions. Amino acid residues within 4.5 \AA from the bound 2-palmitoylglycerol are included. Dotted lines depict the hydrogen bonds. The hydrogen-bonded distances are also indicated in this figure.

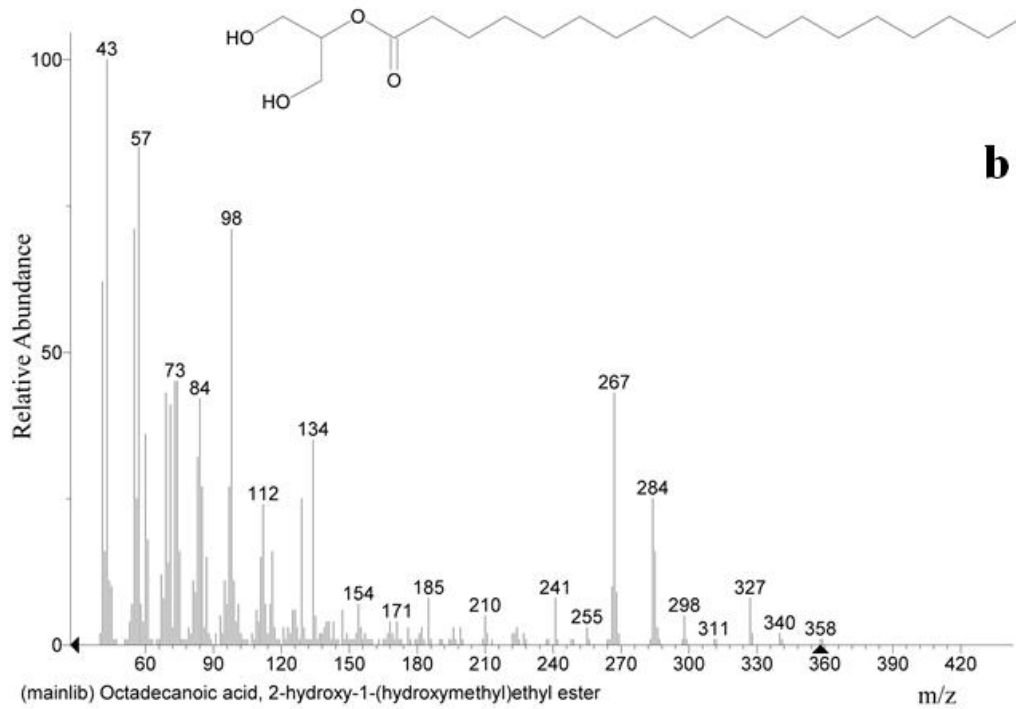
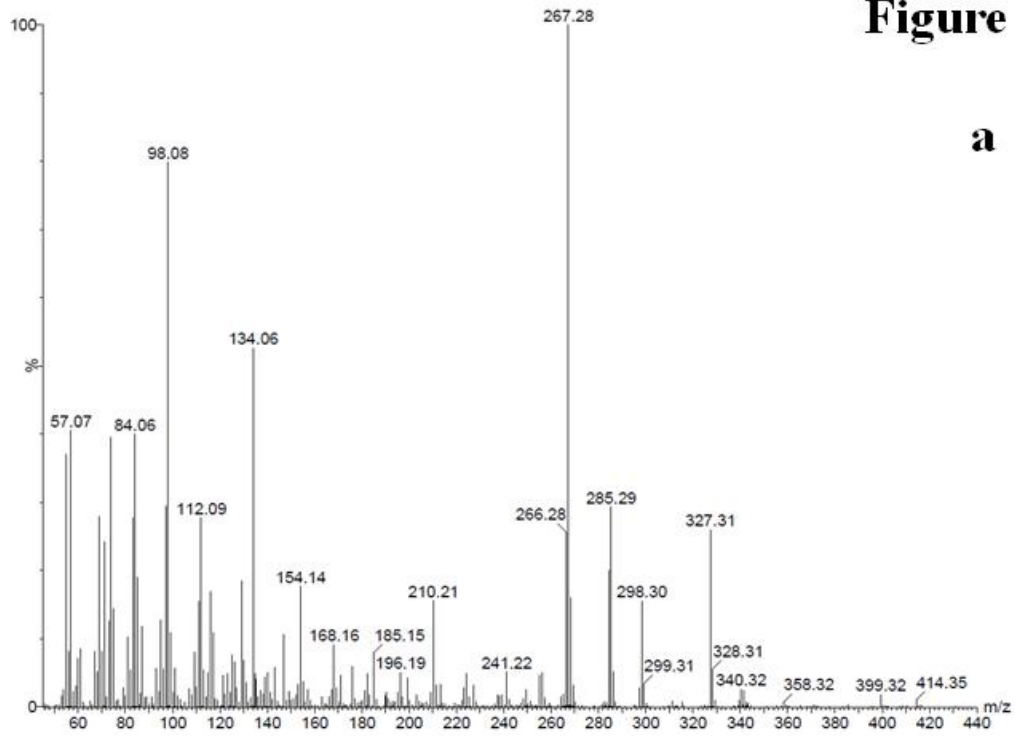
Figure 6

Figure 6. Identification of the fortuitous ligand by GC-MS. (a) Electron ionization (EI) spectrum of the strongest GC peak at 14.45 min. (b) GC-MS spectrum of octadecanoic acid,

2-hydroxyl-1-(hydroxymethyl)ethyl ester from the internal GC-MS library. The ligand was identified as 2-palmitoylglycerol.

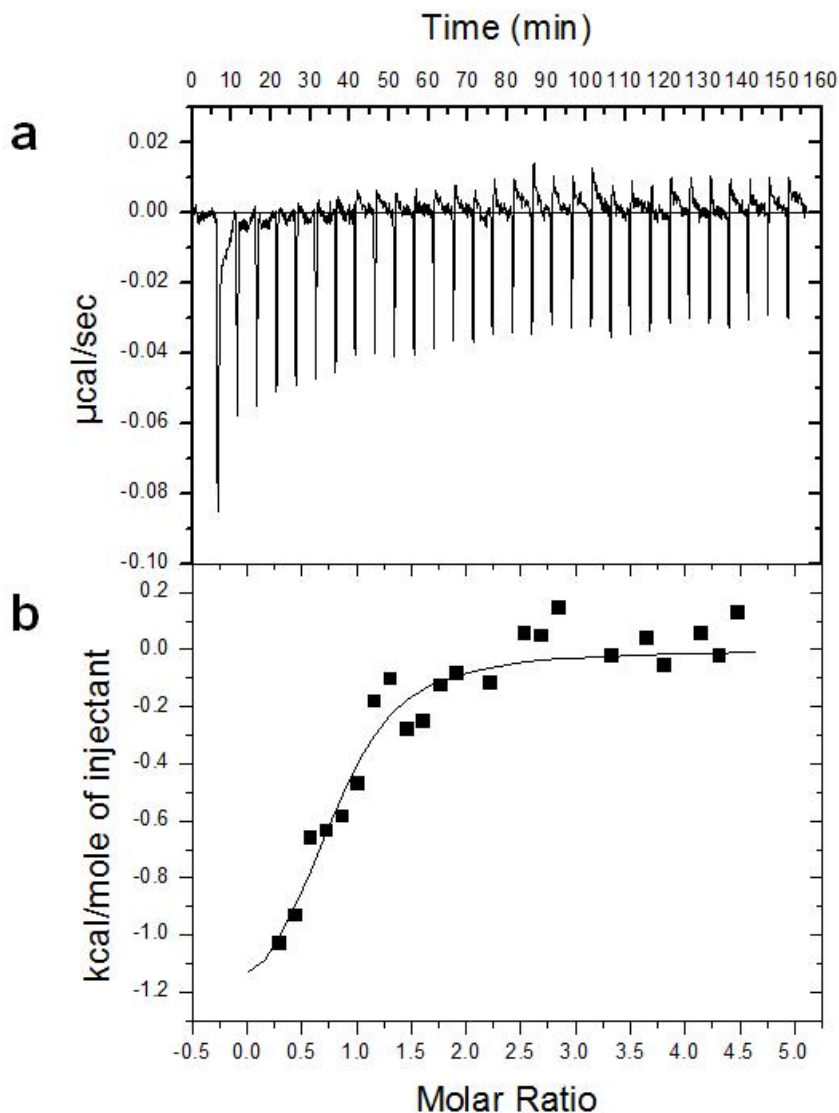


Figure 7

Figure 7. Representative isothermal titration calorimetry for the binding of 1-stearoyl-*rac*-glycerol to Rv0678. (a) Each peak corresponds to the injection of 10 μl of 400 μM dimeric Rv0678 in buffer containing 10 mM Na-phosphate (pH 7.2), 100 mM NaCl and 0.001% DDM into the reaction containing 20 μM 1-stearoyl-*rac*-glycerol in the same buffer. (b)

Cumulative heat of reaction is displayed as a function of the injection number. The solid line is the least-square fit to the experimental data, giving a K_A of $4.9 \pm 0.4 \times 10^5 \text{ M}^{-1}$.

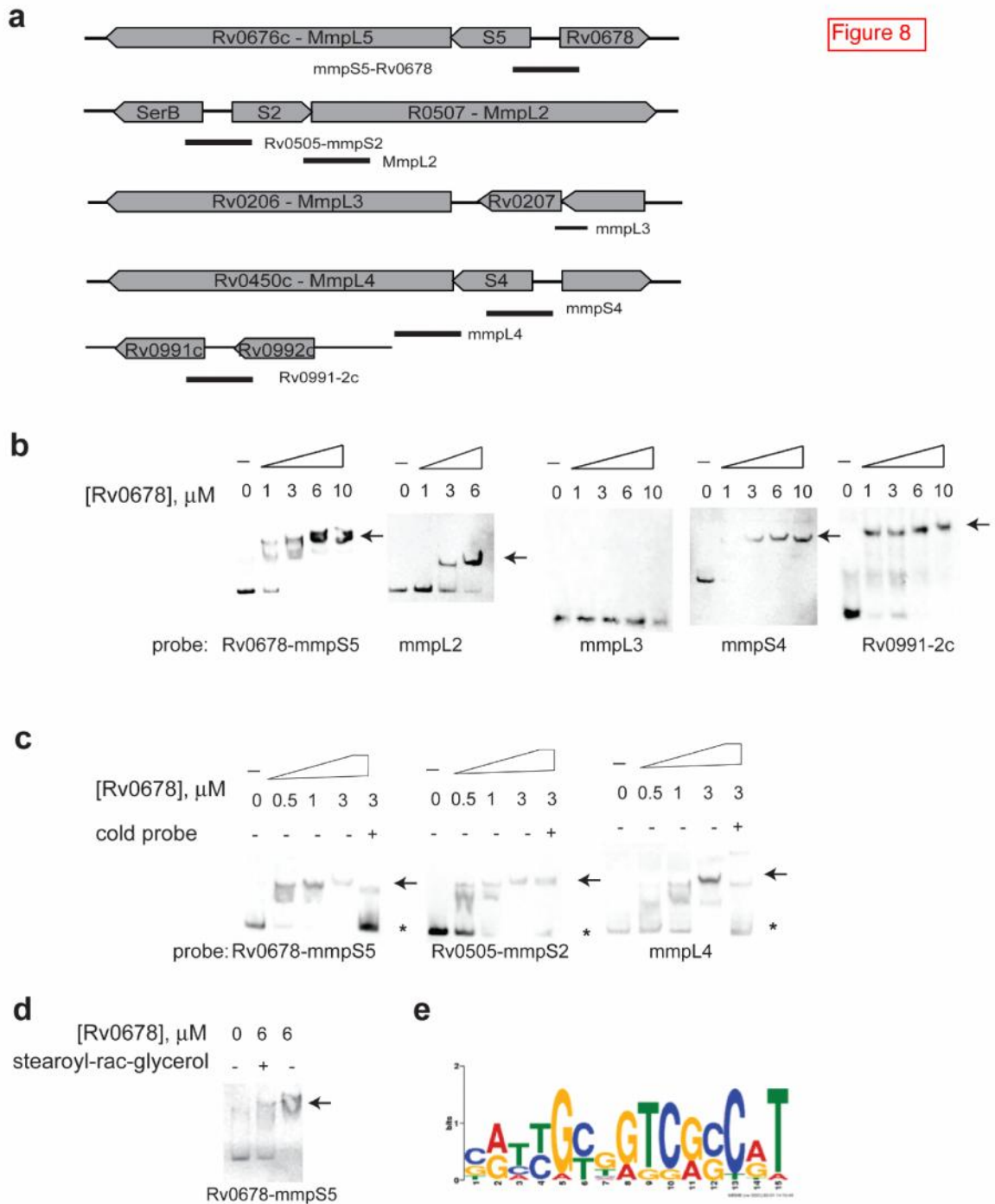


Figure 8

Figure 8. Rv0678 binds to promoter regions of *mmpS2-mmpL2*, *mmpS4-mmpL4*, *mmpS5* and *rv0991-2c*. (a) A schematic depicting the DNA probes used in electrophoretic mobility shift assays (EMSAs) to examine the promoter and intragenic regions of the *mmpS2-mmpL2*, *mmpL3*, *mmpS4-mmpL4*, *mmpS5-mmpL5* and *rv0991-2c* genes. (b) EMSAs were performed using 12 nM Dig-labeled probe and the indicated micromolar concentrations of protein. An arrow denotes the shifted probes. (c) To demonstrate specificity, EMSAs were performed in the presence of non-labeled (“cold”) probe. Reactions were performed with 6 nM Dig-labeled probe, the indicated micromolar concentrations of protein, and 0.6 μ M cold probe. The asterisk notes the accumulation of free Dig-labeled probe. (d) EMSAs were performed using 12 μ M Dig-labeled probe and 6 μ M Rv0678 in the presence or absence of 1 μ M 1-stearoyl-*rac*-glycerol as indicated above the blot. (e) The sequence of the probes bound by Rv0678 in panel b and c were compared using the motif-based sequence analysis tool MEME, yielding a putative Rv0678 binding motif.

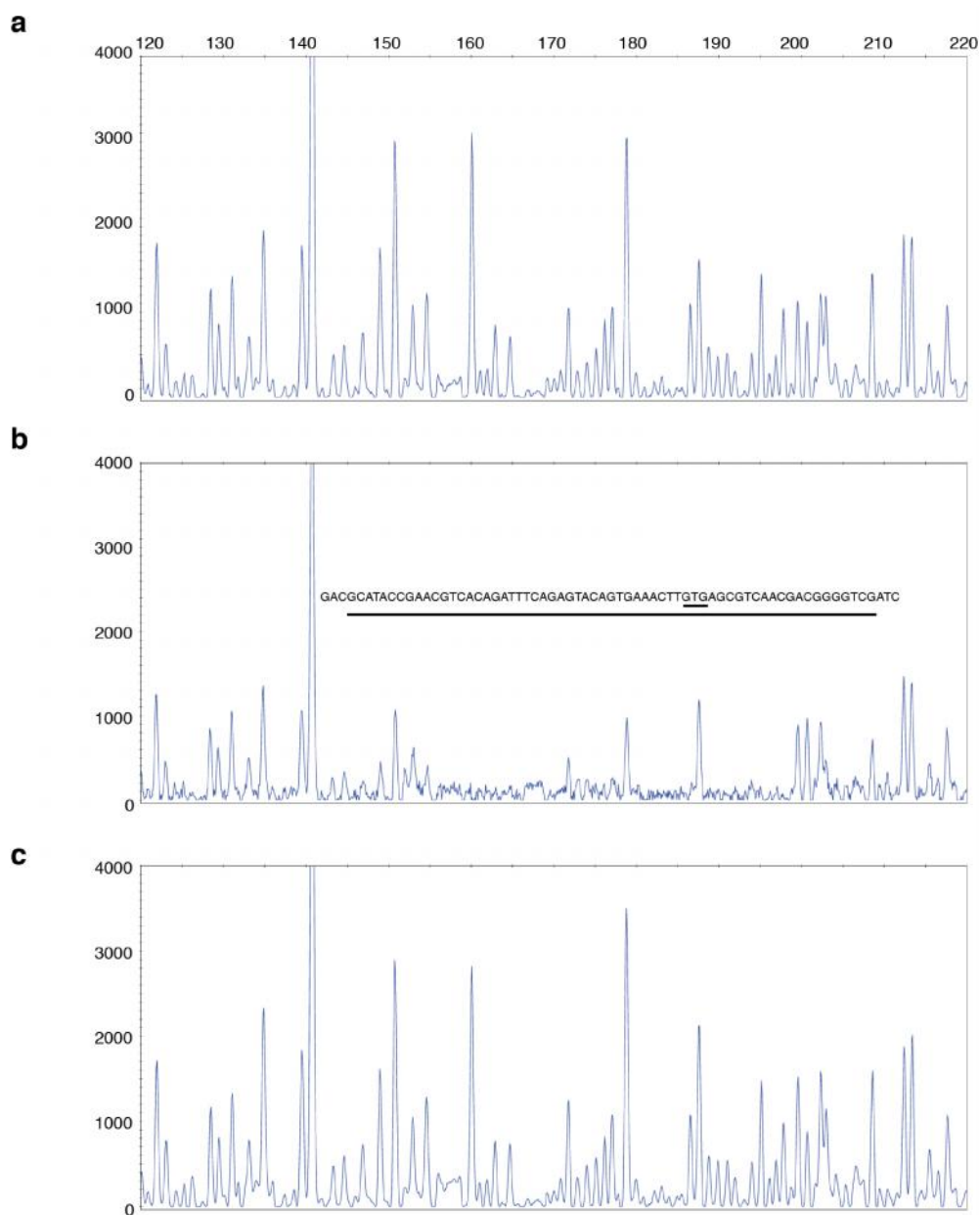


Figure 9. Direct binding of Rv0678 to the *rv0678-mmpS5* intergenic region by dye primer based DNase I footprint assay. Electropherograms indicating the protection pattern of the Rv0678-*mmpS5* probe after digestion with DNase I following incubation with (a) 0, (b) 1 μ M Rv0678 and (c) 1 μ M BSA are shown. The protected DNA sequence is indicated above the electropherogram in panel b, the predicted start codon of *rv0678* is underlined.

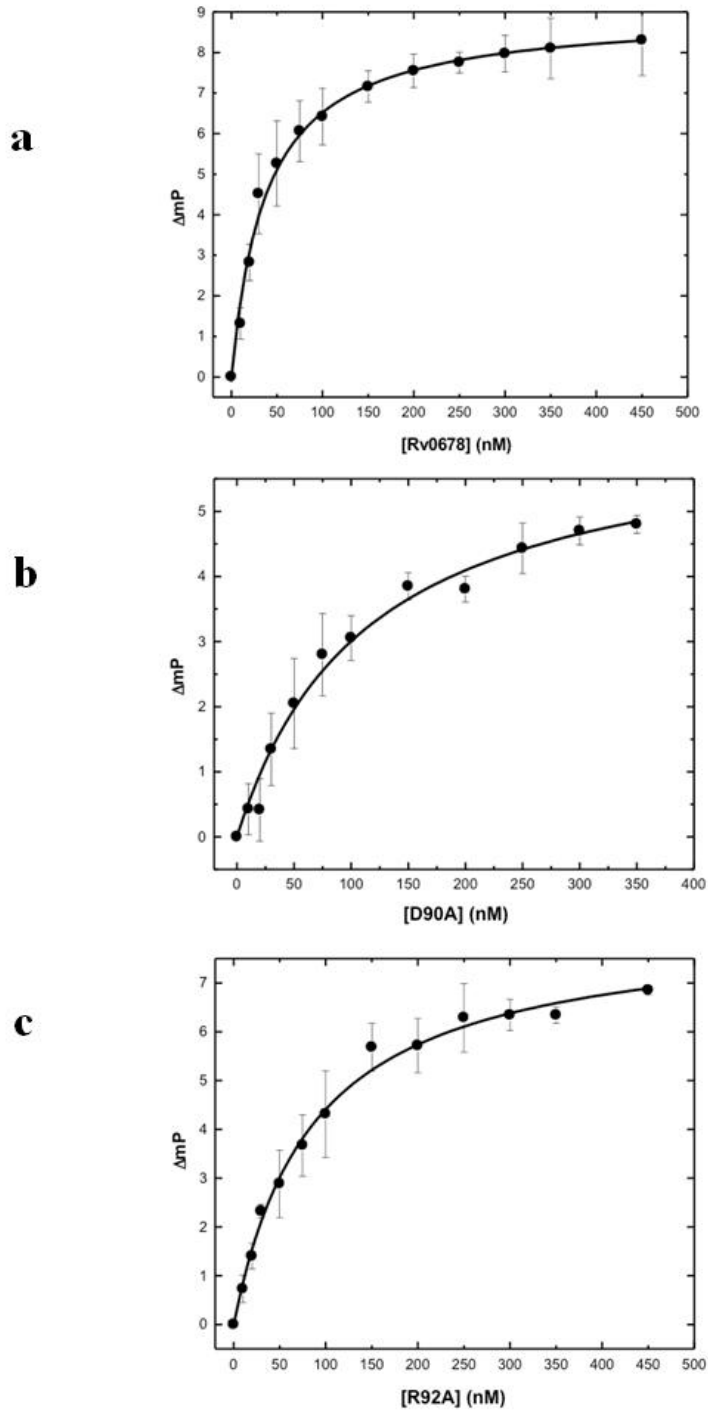


Figure 10

Figure 10. Representative fluorescence polarization of Rv0678. (a) The binding isotherm of Rv0678 with the 26-bp DNA containing the 18-bp promoter sequence, showing a K_D of 19.6 ± 3.0 nM. (b) The binding isotherm of mutant D90A with the 26-bp DNA, showing a K_D of

113.3 ± 16.8 nM. (c) The binding isotherm of mutant R92A with the 26-bp DNA, showing a K_D of 86.0 ± 7.4 nM. Fluorescence polarization is defined by the equation, $FP = (V - H) / (V + H)$, where FP equals polarization, V equals the vertical component of the emitted light, and H equals the horizontal component of the emitted light of a fluorophore when excited by vertical plane polarized light. FP is a dimensionless entity and is not dependent on the intensity of the emitted light or on the concentration of the fluorophore. mP is related to FP, where 1 mP equals one thousandth of a FP.

Table 1. Data collection, phasing and structural refinement statistics of Rv0678.

Data set	Rv0678	$W_6(\mu-O)_6(\mu-Cl)_6Cl_6^{2-}$ derivative
Data Collection		
Wavelength (Å)	0.98	0.98
Space group	<i>P1</i>	<i>P1</i>
Resolution (Å)	50 – 1.64 (1.70-1.64)	50 – 1.90 (1.97-1.90)
Cell constants (Å)		
a	54.54	54.75
b	57.24	57.49
c	61.44	61.42
α, β, γ (°)	82.2, 68.4, 72.2	82.3, 68.5, 72.4
Molecules in ASU	4	4
Redundancy	2.0 (2.0)	1.9 (1.8)
Total reflections	326,940	512,196
Unique reflections	80,449	52,208
Completeness (%)	97.5 (95.6)	88.4 (90.1)
R_{sym} (%)	4.4 (39.5)	9.1 (35.3)
$I / \sigma(I)$	17.46 (2.2)	14.29 (3.4)
Phasing		
Number of sites		6
Phasing power (acentric)		1.71
R_{cullis} (acentric)		0.70
Figure of merit (acentric)		0.66
Refinement		
Resolution (Å)	50 – 1.64	
R_{work}	16.28	
R_{free}	19.44	
Average B-factor (Å ²)	23.85	
RMSD bond lengths (Å)	0.011	
RMSD bond angles (°)	1.253	
Ramachandran plot		
most favoured (%)	96.7	
additional allowed (%)	3.3	
generously allowed (%)	0	
disallowed (%)	0	

Table 2. Primers.

Probe	Primer 1	Primer 2
Rv0678	CTTCGGAACCAAAGAAAGTG	CCAACCGAGTCAAACCTCCTG
Rv0505	GAACACGAGGGTGAGGATG	GCGTCGTCTCGACCGTGAC
Rv0991-2	GAGCTGGTTGACTTCTCGG	CAATGCGGTCGGCGTGTTG

Table 3. Primers for site-directed mutagenesis.

D90A-forward	5'-CGCCTGGCAGTCGCTGGTGTCTCGTCGCACGTATTTTCGTC-3'
D90A-reverse	5'-GACGAAAATACGTGCGACGAGCACCAGCGACTGCCAGGCG-3'
R92A-forward	5'-GCAGTCGCTGGTGATCGTGCCACGTATTTTCGTCTGCGC-3'
R92A-reverse	5'-GCGCAGACGAAAATACGTGGCACGATCACCAGCGACTGC-3'

Table 4. Rv0678-ligand contacts.

	Dimer 1	Dimer 2
Residue-ligand contacts	Distance (Å)	Distance (Å)
R32	3.2*	-
Q78	3.9	3.7
F79	3.8	4.2
E108	3.4	3.2
R109	2.8*	3.2*
R111	3.4	3.5
A112	4.0	3.6
M113	-	4.4
E115	3.0	2.9
L116	4.4	3.7
L144	-	4.4
L145	-	4.0
Y28'	4.0	3.9
F29'	4.4	4.3
R32'	3.5	3.6
L34'	4.2	3.6
F79'	2.8*	-
F81'	3.4	3.5
F102'	4.5	2.3
A103'	-	4.4
G105'	2.9	3.0

Table 4. Continued.

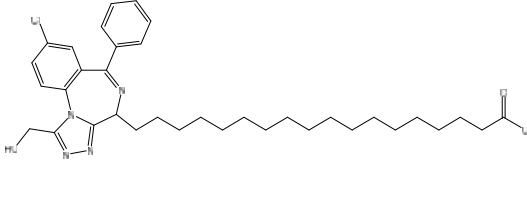
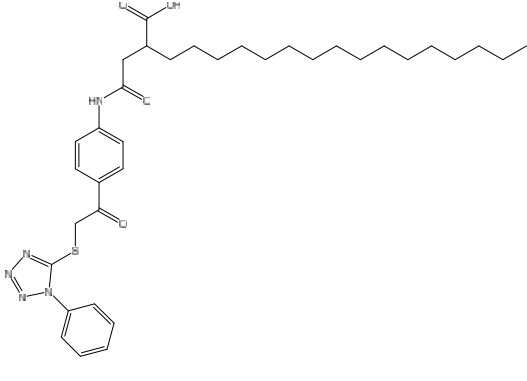
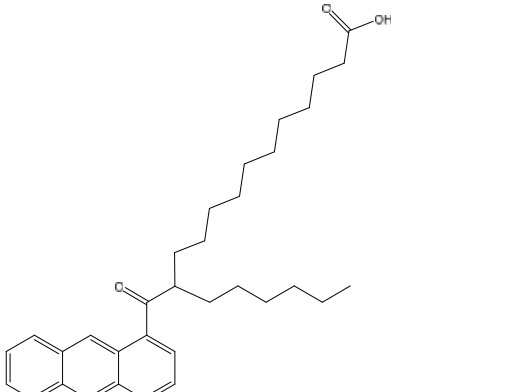
E106'	3.2*	3.1*
E108'	3.9	-
R109'	3.8	3.5

* Hydrogen bond distance
Contacts within 4.5Å are listed

Table 5. Top three ligands for the Rv0678 regulator.

Ligand	Structure of ligand	Binding affinity (kcal/mol)
Diethyl-[(5E)-5-(6,8,9,10-tetrahydro-5H-benzo[c]xanthen-11-ylmethylene)-7,8-dihydro-6H-xanthen-3-yl]		-13.8
(5E)-N,N-diethyl-5-(6,8,9,10-tetrahydrochromeno[3,2-c]chromen-11-ylmethylene)-7,8-dihydro-6H-xanthen		-13.5
Dibenzylfluorescein		-13.4

Table 6. Top three fatty acids for the Rv0678 regulator.

Fatty acid	Structure of fatty acid	Binding affinity (kcal/mol)
18-[8-chloro-1-(hydroxymethyl)-6-phenyl-4H-[1,2,4]triazolo[4,3-a][1,4]benzodiazepin-4-yl]octadecanoic acid		-10.6
2-[2-oxo-2-[4-[2-(1-phenyltetrazol-5-yl)sulfanylacetyl]anilino]ethyl]octadecanoic acid		-10.5
12-(anthracene-1-carbonyl)octadecanoic acid		-10.4

**CHAPTER 3. CRYSTAL STRUCTURE OF THE *CAMPYLOBACTER JEJUNI* CMEC
OUTER MEMBRANE CHANNEL**

A paper published in Protein Science, 23: 954–961. doi:10.1002/pro.2478

Chih-Chia Su^{1,ψ}, Abhijith Radhakrishnan^{2,ψ}, Nitin Kumar^{2,ψ}, Feng Long¹, Jani Reddy Bolla²,
Hsiang-Ting Lei², Jared A. Delmar¹, Sylvia V. Do³, Tsung-Han Chou¹, Kanagalaghatta R.
Rajashankar⁴, Qijing Zhang⁵ and Edward W. Yu^{1,2,3*}

¹Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA

²Department of Chemistry, Iowa State University, Ames, IA 50011, USA

³Bioinformatics and Computational Biology Interdepartmental Graduate Program, Iowa State
University, Ames, IA 50011, USA

⁴NE-CAT and Department of Chemistry and Chemical Biology, Cornell University, Bldg.
436E, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne. IL 60439, USA

⁵Department of Veterinary Microbiology, College of Veterinary Medicine, Iowa State
University, Ames, IA 50011, USA

^ψC.S., A.R. and N.K. contributed equally to this work.

* To whom correspondence should be addressed. E-mail: ewyu@iastate.edu

Abstract

As one of the world's most prevalent enteric pathogens, *Campylobacter jejuni* is a major causative agent of human enterocolitis and is responsible for more than 400 million cases of diarrhea each year. The impact of this pathogen on children is of particular significance. *Campylobacter* has developed resistance to many antimicrobial agents via multidrug efflux machinery. The CmeABC tripartite multidrug efflux pump, belonging to the resistance-nodulation-cell division (RND) superfamily, plays a major role in drug resistant phenotypes of *C. jejuni*. This efflux complex spans the entire cell envelope of *C. jejuni* and mediates resistance to various antibiotics and toxic compounds. We here report the crystal structure of *C. jejuni* CmeC, the outer membrane component of the CmeABC tripartite multidrug efflux system. The structure reveals a possible mechanism for substrate export.

Introduction

Campylobacter jejuni is a major causative agent of human enterocolitis and is responsible for more than 400 million cases of diarrhea each year worldwide.¹ *Campylobacter* infection may also trigger an autoimmune response, which is associated with the development of Guillain-Barre syndrome, an acute flaccid paralysis caused by degeneration of the peripheral nervous system.² *C. jejuni* is widely distributed in the intestinal tracts of animals and is transmitted to humans via contaminated food, water, or raw milk. For antibiotic treatment of human campylobacteriosis, fluoroquinolones and macrolides are frequently prescribed. Unfortunately, *Campylobacter* has developed resistance to these antimicrobials, especially fluoroquinolones.³⁻⁵ Antimicrobial resistance in *Campylobacter* has become a major concern for public health. Resistance of *Campylobacter* to antibiotics is mediated by multiple mechanisms,⁶ including synthesis of antibiotic-

inactivating enzymes, alteration or protection of antibiotic targets and active extrusion of drugs from *Campylobacter* cells via multidrug efflux pumps. While the first two mechanisms typically confer resistance to a specific class of drugs, multidrug efflux pumps in *Campylobacter* contribute to both intrinsic and acquired resistance to a broad range of antimicrobials and toxic compounds. Acquisition of antibiotic resistance by *Campylobacter* not only compromises the effectiveness of clinical treatment, but also affect the course of the clinical diseases.

According to the genomic sequence of NCTC 11168, *C. jejuni* harbors 13 putative antibiotic efflux transporters of the resistance-nodulation-cell division (RND), ATP-binding cassette (ABC), multidrug and toxic compound extrusion (MATE), major facilitator (MF) and small multidrug resistance (SMR) families.⁷ In Gram-negative pathogens, efflux systems belonging to the hydrophobic and amphiphilic efflux RND (HAE-RND) play major roles in the intrinsic and acquired tolerance of antibiotics and toxic compounds. Among them, the *Campylobacter* multidrug efflux system CmeABC,⁸⁻¹⁰ a HAE-RND-type efflux pump,¹¹ is the primary antibiotic efflux system and is the best functionally characterized transporter in *Campylobacter*. The tripartite CmeABC complex includes the outer membrane channel CmeC, the inner membrane drug transporter CmeB, and the periplasmic membrane fusion protein CmeA. CmeABC contributes significantly to the intrinsic and acquired resistance of *Campylobacter* to structurally diverse antimicrobials, including fluoroquinolones and macrolides, by reducing the accumulation of drugs in *Campylobacter* cells.^{8,10,12} It has been found that CmeABC functions synergistically with target mutations in conferring and maintaining high-level resistance to fluoroquinolones and macrolides.^{12,13} This efflux pump also plays an important role in the emergence of fluoroquinolone-resistant *Campylobacter* under selection pressure.¹⁴ Inactivation of CmeABC reduced the frequency of emergence of fluoroquinolone-resistant mutants, while overexpression of CmeABC increased this

frequency.¹⁴ The contributing effect of CmeABC is attributed to the fact that many of the spontaneous *gyrA* mutants cannot survive the selection by ciprofloxacin in the absence of CmeABC.

In addition to conferring antibiotic resistance, CmeABC has an important role in bile resistance. As an enteric pathogen, *C. jejuni* must possess means to adapt in the animal intestinal tract, where bile acids are commonly present. Indeed, it has been suggested that bile resistance is the natural function of this HAE-RND-type efflux pump.⁹

Currently, two crystal structures of the HAE-RND efflux pumps have been resolved by crystallography. These efflux pumps are the *Escherichia coli* AcrB¹⁵⁻²⁰ and *Pseudomonas aeruginosa* MexB²¹ multidrug transporters. The crystal structures of other components of these tripartite complex systems have also been determined. These include the outer membrane channels, *E. coli* TolC²² and *P. aeruginosa* OprM,²³ as well as the periplasmic membrane fusion proteins, *E. coli* AcrA²⁴ and *P. aeruginosa* MexA.²⁵⁻²⁷

Thus far, no structural information is available for any protein component of the CmeABC tripartite complex system. To elucidate the mechanism used by the *C. jejuni* CmeABC efflux system for multidrug extrusion, we here describe the crystal structure of the CmeC outer membrane channel. The structure reveals that this the interior surface of this channel is closed at the outer membrane surface as well as the periplasmic tip, raising the possibility that this channel forms two gates for substrate export.

Results and discussion

Overall structure of the *C. jejuni* CmeC outer membrane channel

We cloned, expressed, purified and crystallized the full-length CmeC channel protein that contains a 6xHis tag at the C-terminus. We obtained crystals of this membrane protein using vapor diffusion. Data collection and refinement statistics are summarized in Table 1.

The diffraction data were indexed to the space group $C222_1$. The crystal structure of CmeC was then determined to a resolution of 2.37 Å using single anomalous dispersion. In the asymmetric unit, three CmeC protomers were found to form a trimer. Figure 1 illustrates the electron density maps of this CmeC trimer. The final structure consists of 100% of the protein sequence, and is refined to R_{work} and R_{free} of 21.0% and 24.3%, respectively (Table 1).

CmeC assembles as a trimer of 492 residues per protomer. Like other known structures of outer membrane channel proteins, including TolC²² and OprM,²³ the appearance of the CmeC trimer is that of a cannon, forming a 130 Å long tunnel to export antimicrobials such as fluoroquinolones and macrolides. Each monomer contributes four β-strands and six α-helices to form the transmembrane β-barrel and periplasmic α-helical domains (Figs. 2 and 3). In addition, the CmeC trimer contains an equatorial domain, which is made up of 12 short α-helices (four from each protomer), encircling the mid-section of the periplasmic α-helical domain. The periplasmic tunnel of CmeC is nearly 100 Å long with an outermost diameter of 35 Å at the tip of the tunnel.

Like other outer membrane channel protein structures, including TolC,²² OprM²³ and CusC,^{28,29} CmeC possesses a typical charge distribution found in this protein family. That is, the interior surface of the CmeC channel is strikingly electronegative (Fig. 4). However, the charge distribution of the exterior surface of a typical outer membrane channel is quite random and does not form extensive positively or negatively charge patches. CmeC is distinct in that its exterior surface is also predominately occupied by acidic residues (Fig. 4).

The N-terminal cysteine residue

It is interesting to note that the N-terminal end of CmeC forms an elongated loop. This loop extends from the membrane surface and leads down to the equatorial domain in the

periplasm (Fig. 2). The first N-terminal residue of CmeC is a cysteine. Like OprM²³ and CusC,²⁸ the crystal structure of CmeC suggests that this residue is covalently linked to the lipid elements at the inner leaflet of the outer membrane. This cysteine residue is believed to play an important role in protein-membrane interaction and critical for the insertion of this channel protein into the outer membrane.²⁹ It has been shown that a single-point mutation on the corresponding cysteine residue in the CusC channel is able to unfold the entire transmembrane β -barrels, disallowing the channel protein to anchor to the outer membrane.²⁹ As phylogenetic tree of homologous suggests that *C. jejuni* CmeC and *E. coli* CusC are derived from the same branch, it is expected that a mutation on this N-terminal cysteine residue will have a drastic effect on the structure and function of the CmeC channel.

CmeC forms two interior gates

The structure of CmeC indicates that this channel is at its closed conformational state as this channel is closed at both ends. The interior of the CmeC outer membrane β -barrel is partially occluded (Fig. 5). Residues 96-108 appear to form a flexible loop between strands S3 and S4. This loop may be responsible for the opening and closing of the top end of the β -barrel. In the CmeC structure, the three R104 residues from different protomers are found to interact with each other through hydrogen bonds (Fig. 6). Thus, R104 seems to form a gate at the upper part of the β -barrel and control the passage of antimicrobial molecules through the outer membrane.

The periplasmic end of the α -barrel of CmeC is also partially occluded (Fig. 5). The α -helices at this end, inner H7/H8 and outer H3/H4, are densely packed through coiled-coil interactions. The gate is formed by a layer of charged and polar residues, including Q412, D413, E416 and N420 (Fig. 6). Among these residues, Q412 and E416' of the next subunit interact through hydrogen bonds to form the gate's central core. Thus, the CmeC channel

protein seems to form two gates: a negatively charged gate located at the tip of the α -helical periplasmic domain, and a positively charged gate found at the top portion of the interior of the β -barrel outer membrane domain. During antimicrobial extrusion, these two gates may dilate sequentially to allow passage of these antimicrobial molecules.

The exterior intra- and inter-protomer grooves

The outermost surface of the periplasmic domain of the CmeC trimer forms three intra-protomer and three inter-protomer grooves. Like the MtrE channel,^{30,31} these CmeC grooves are likely to provide interaction sites for the CmeA membrane fusion protein. If this is the case, then the α -helical coiled-coil domain of the CmeA fusion protein is likely to fit into these grooves and contact CmeC to form a complex to function. In turn, this CmeA-CmeC interaction could control the opening and closing of the CmeC outer membrane channel, similar to the case of the *Neisseria gonorrhoeae* MtrCDE efflux system.^{30,31}

It is well established that overexpression of RND multidrug efflux pumps leads to a resistant phenotype in pathogenic organisms. Because of the fact that these multidrug efflux pumps are able to respond to a wide spectrum of substrates, pathogenic bacteria that overexpress them can be selected for by many different agents.³² Thus, it is very important to understand the molecular mechanism as well as detailed structural information of these efflux pumps in order to combat infectious diseases. The pathogen *Campylobacter* has developed resistance to many antimicrobial agents. The availability of the crystal structure of the CmeC outer membrane channel may allow us to rationally design agents that block the function of this channel protein and eventually heighten the sensitivity of this pathogen to antimicrobials.

Materials and methods

Cloning, expression and purification of the outer membrane CmeC channel

The full-length CmeC membrane protein containing a 6xHis tag at the C-terminus was overproduced in *E. coli* C43(DE3)/pBAD22b Ω cmeC cells. Cells were grown in 12 L of LB medium with 100 μ g/ml ampicillin at 37°C. When the OD₆₀₀ reached 0.5, the culture was cooled down to 25°C and then treated with 0.2% arabinose to induce *cmeC* expression. Cells were harvested after shaking for 16 h at 25°C. The collected bacteria were resuspended in buffer containing 20 mM Na-HEPES (pH 7.5), 300 mM NaCl and 1 mM phenylmethanesulfonyl fluoride (PMSF), and then disrupted with a French pressure cell. The membrane fraction was collected by ultracentrifugation, followed by a pre-extraction procedure by incubating in buffer containing 0.5% sodium lauroyl sarcosinate, 20 mM Na-HEPES (pH 7.5) and 50 mM NaCl for 0.5 h at room temperature. The membrane was collected and washed twice with buffer containing 20 mM HEPES-NaOH (pH 7.5) and 50 mM NaCl. The membrane protein was then solubilized in 2 % (w/v) n-dodecyl β -D-maltoside (DDM). Insoluble material was removed by ultracentrifugation at 100,000 x g. The extracted protein was purified with a Ni²⁺-affinity column. The purity of the CmeC protein (>95%) was judged using 12% SDS-PAGE stained with Coomassie Brilliant Blue. The purified protein was then dialyzed and concentrated to 15 mg/ml in buffer containing 20 mM Na-HEPES (pH 7.5), 200 mM NaCl and 0.05% DDM.

For 6xHis selenomethionyl-substituted (SeMet)-CmeC, the protein was expressed in *E. coli* C43(DE3) cells possessing pBAD22b Ω cmeC. In brief, a 10 ml LB broth overnight culture containing *E. coli* C43(DE3)/pBAD22b Ω cmeC cells was transferred into 120 ml of LB broth containing 100 μ g/ml ampicillin and grown at 37°C. When the OD₆₀₀ value reached 1.2, cells were harvested by centrifugation at 6000 rev/min for 10 min, and then washed two times with 20 ml of M9 minimal salts solution. The cells were re-suspended in 120 ml of M9

media and then transferred into a 12 L pre-warmed M9 solution containing 100 µg/ml ampicillin. The cell culture was incubated at 37°C with shaking. When the OD₆₀₀ reached 0.4, the culture was cooled down to 25°C and 100 mg/l of lysine, phenylalanine and threonine, 50 mg/l isoleucine, leucine and valine, and 60 mg/l of L-selenomethionine were added. The culture was induced with 0.2% arabinose after 15 min. Cells were then harvested within 16 h after induction. Further procedures for the preparation, pre-extraction and purification of the outer membrane fraction as well as purification of the SeMet-CmeC protein were identical to those for the native CmeC channel. The purity of the SeMet-CmeC protein (>95%) was judged using 10% SDS-PAGE stained with Coomassie Brilliant Blue. The purified protein was then dialyzed and concentrated to 15 mg/ml in buffer containing 20 mM Na-HEPES (pH 7.5), 200 mM NaCl and 0.05% DDM.

Crystallization of CmeC

Crystals of the 6xHis CmeC were obtained using sitting-drop vapor diffusion. The CmeC crystals were grown at room temperature in 24-well plates with the following procedures. A 2 µl protein solution containing 15 mg/ml CmeC protein in 20 mM Na-HEPES (pH 7.5), 200 mM NaCl and 0.05% (w/v) DDM was mixed with a 2 µl of reservoir solution containing 18% PEG 400, 0.1 M sodium acetate (pH 4.0), 0.3 M (NH₄)₂SO₄ and 2% tetraethylene glycol monoethyl ether (C₈E₄). The resultant mixture was equilibrated against 500 µl of the reservoir solution. Crystals of CmeC grew to a full size in the drops within two weeks. Typically, the dimensions of the crystals were 0.2 mm x 0.2 mm x 0.2 mm. Cryoprotection was achieved by raising the PEG 400 concentration stepwise to 30% with a 4% increment in each step. The procedures for growing and preparing the SeMet-CmeC crystals were identical to those for the native CmeC crystals.

Data collection, structural determination and refinement

All diffraction data were collected at 100K at beamline 24ID-C located at the Advanced Photon Source, using an ADSC Quantum 315 CCD-based detector. Diffraction data were processed using DENZO and scaled using SCALEPACK.³³

Crystals of the CmeC outer membrane channel belong to the space group $C222_1$ (Table 1). Based on the molecular weight of CmeC (54.26 kDa), three molecules per asymmetric unit with a solvent content of 80.6% were expected. While the structure of CmeC could be determined using molecular replacement, the structure resulted in a relatively high refinement R-factor of 49.8%. To aid modeling of the conformation of CmeC in the least biased manner, single anomalous dispersion phasing using the program PHASER³⁴ was employed to obtain experimental phases in addition to the phases from the structural model of OprM.²³ Phases were then subjected to density modification and phase extension to 2.37 Å-resolution using the program RESOLVE.³⁵ The resulting phases were of excellent quality, which enabled us to trace the conformation of CmeC using the program COOT.³⁶ The full-length CmeC protein contains five methionine residues and all of these five selenium sites per CmeC molecule (15 selenium sites per asymmetric unit) were identified. The SeMet data not only augmented the experimental phases, but also helped in tracing the molecules by anomalous difference Fourier maps where we could ascertain the proper registry of SeMet residues. After tracing the initial model manually using the program Coot,³⁶ the model was refined against the native data at 2.37 Å-resolution using TLS refinement techniques adopting a single TLS body as implemented in PHENIX³⁷ leaving 5% of reflections in Free-R set. Iterations of refinement using PHENIX³⁷ and CNS³⁸ and model building in Coot³⁶ lead to the current CmeC model, which consists of 1,419 residues in the asymmetric unit with excellent geometrical characteristics (Table 1).

Accession codes

Atomic coordinates and structure factors have been deposited in the Protein Data Bank under the accession code 4MT4.

Acknowledgements

This work is supported by NIH Grants R01DK063008 (Q.Z.) and R01GM086431 (E.W.Y.). This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by an award GM103403 from the National Institutes of General Medical Sciences. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

References

1. Ruiz-Palacios GM (2007) The health burden of *Campylobacter* infection and the impact of antimicrobial resistance: playing chicken. *Clin Infect Dis* 44:701-703.
2. van Doorn PA, Ruts L, Jacobs BC (2008) Clinical features, pathogenesis, and treatment of Guillain-Barre syndrome. *Lancet Neurol* 7:939-950.
3. Engberg J, Aarestrup FM, Taylor DE, Gerner-Smidt P, Nachamkin I (2001) Quinolone and macrolide resistance in *Campylobacter jejuni* and *C. coli*: resistance mechanisms and trends in human isolates. *Emerg Infect Dis* 7:24-34.
4. Gibreel A, Taylor DE (2006) Macrolide resistance in *Campylobacter jejuni* and *Campylobacter coli*. *J Antimicrob Chemother* 58:243-255.

5. Luangtongkum T, Jeon B, Han J, Plummer P, Logue CM, Zhang Q (2009) Antibiotic resistance in *Campylobacter*: emergence, transmission and persistence. *Future Microbiol* 4:189-200.
6. Zhang Q, Plummer P (2008) Mechanisms of antibiotic resistance in *Campylobacter*, in *Campylobacter* (Nachamkin I, Szymanski CM, Blaser MJ Eds.) 3rd ed., pp 263-276.
7. Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream MA, Rutherford KM, van Vliet AH, Whitehead S, Barrell BG (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 403:665-668.
8. Lin J, Michel LO, Zhang Q (2002) CmeABC Functions as a Multidrug Efflux System in *Campylobacter jejuni*. *Antimicrob Agents Chemother* 46:2124-2131.
9. Lin J, Sahino O, Michel LO, Zhang Q (2003) Critical role of multidrug efflux pump CmeABC in bile resistance and *in vivo* colonization of *Campylobacter jejuni*. *Infect Immun* 71:4250-4259.
10. Pumbwe L, Piddock LJ (2002) Identification and molecular characterisation of CmeB, a *Campylobacter jejuni* multidrug efflux pump. *FEMS Microbiol Lett* 206:185-189.
11. Tseng TT, Gratwick KS, Kollman J, Park D, Nies DH, Goffeau A, Saier MH Jr (1999) The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development protein. *J Mol Microbiol Biotechnol* 1:107-125.
12. Luo N, Sahin O, Lin J, Michel LO, Zhang Q (2003) *In vivo* selection of *Campylobacter* isolates with high levels of fluoroquinolone resistance associated with

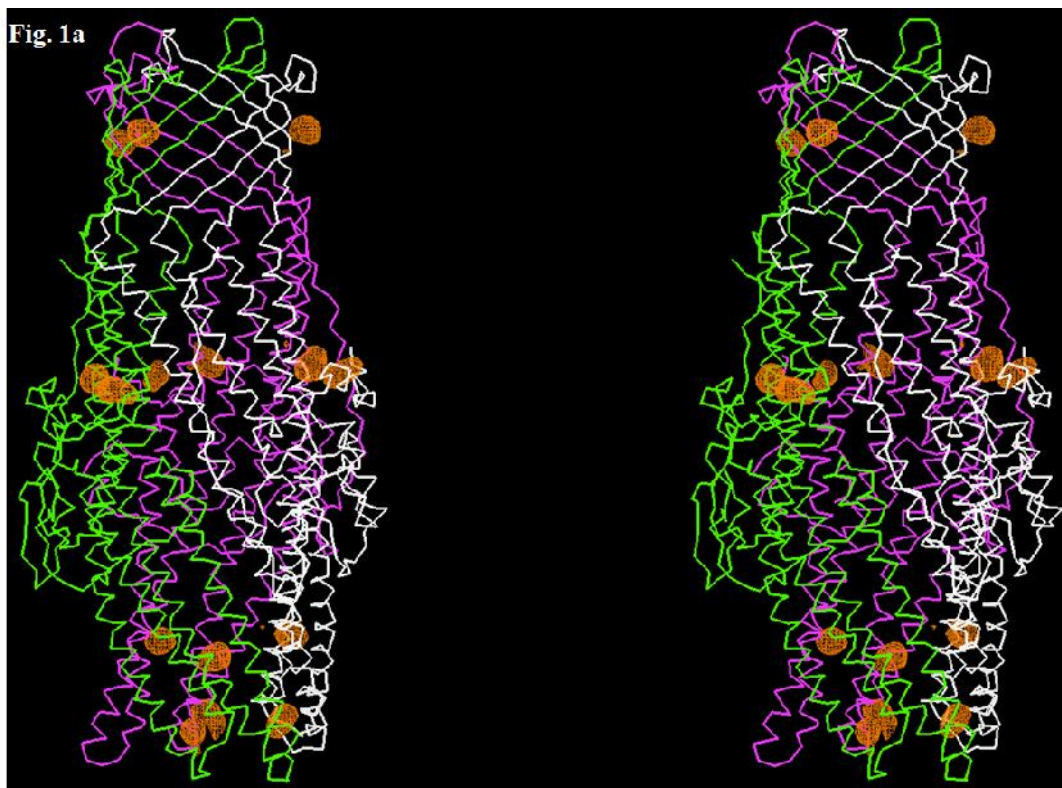
- gyrA mutations and the function of the CmeABC efflux pump. *Antimicrob Agents Chemother* 47:390-394.
13. Ge B, McDermott PF, White DG, Meng J (2005) Role of efflux pumps and topoisomerase mutations in fluoroquinolone resistance in *Campylobacter jejuni* and *Campylobacter coli*. *Antimicrob Agents Chemother* 49:3347-3354.
 14. Yan M, Sahin O, Lin J, Zhang Q (2006) Role of the CmeABC efflux pump in the emergence of fluoroquinolone-resistant *Campylobacter* under selection pressure. *J Antimicrob Chemother* 58:1154-1159.
 15. Murakami S, Nakashima R, Yamashita E, Yamaguchi A (2002) Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* 419:587-593.
 16. Murakami S, Nakashima R, Yamashita E, Matsumoto T, Yamaguchi A (2006) Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* 443:173-179.
 17. Seeger MA, Schiefner A, Eicher T, Verrey F, Dietrichs K, Pos KM (2006) Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science* 313:1295-1298.
 18. Sennhauser G, Amstutz P, Briand C, Storchengegger O, Grütter MG (2007) Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS Biol* 5:e7.
 19. Yu EW, McDermott G, Zgruskaya HI, Nikaido H, Koshland DE Jr (2003) Structural basis of multiple drug binding capacity of the AcrB multidrug efflux pump. *Science* 300:976-980.
 20. Yu EW, Aires JR, McDermott G, Nikaido H (2005) A periplasmic-drug binding site of the AcrB multidrug efflux pump: a crystallographic and site-directed mutagenesis study. *J Bacteriol* 187:6804-6815.

21. Sennhauser G, Bukowska MA, Briand C, Grütter MG (2009) Crystal structure of the multidrug exporter MexB from *Pseudomonas aeruginosa*. *J Mol Biol* 389:134-145.
22. Koronakis V, Sharff A, Koronakis E, Luisi B, Hughes C (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 405:914-919.
23. Akama H, Kanemaki M, Yoshimura M, Tsukihara T, Kashiwag T, Yoneyama H, Narita S, Nakagawa A, Nakae T (2004) Crystal structure of the drug discharge outer membrane protein, OprM, of *Pseudomonas aeruginosa*. *J Biol Chem* 279:52816-52819.
24. Mikolosko J, Bobyk K, Zgurskaya HI, Ghosh P (2006) Conformational flexibility in the multidrug efflux system protein AcrA. *Structure* 14:577-587.
25. Akama H, Matsuura T, Kashiwag S, Yoneyama H, Narita S, Tsukihara T, Nakagawa A, Nakae T (2004) Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in *Pseudomonas aeruginosa*. *J Biol Chem* 279:25939-25942.
26. Higgins MK, Bokma E, Koronakis E, Hughes C, Koronakis V (2004) Structure of the periplasmic component of a bacterial drug efflux pump. *Proc Natl Acad Sci USA* 101:9994-9999.
27. Symmons M, Bokma E, Koronakis E, Hughes C, Koronakis V (2009) The assembled structure of a complete tripartite bacterial multidrug efflux pump. *Proc Natl Acad Sci USA* 106:7173-7178.
28. Kulathila R, Kulathila R, Indic M, van den Berg B (2011) Crystal structure of *Escherichia coli* CusC, the outer membrane component of a heavy-metal efflux pump. *PLoS One* 6:e15610.

29. Lei HT, Bolla JR, Bishop NR, Su CC, Yu EW (2014) Crystal structures of CusC review conformational changes accompanying folding and transmembrane channel formation. *J Mol Biol* 426:403-411.
30. Janganan TK, Zhang L, Bavro VN, Matak-Vinkovic D, Barrera NP, Burton MF, Steel PG, Robinson CV, Borges-Walmsley MI, Walmsley AR (2011) Opening of the outer membrane protein channel in tripartite efflux pumps is induced by interaction with the membrane fusion partner. *J Biol Chem* 286:5484-5493.
31. Janganan TK, Barvro VN, Zhang L, Borges-Walmsley MI, Walmsley AR (2013) Tripartite efflux pumps: energy is required for dissociation, but not assembly or opening of the outer membrane channel of the pump. *Mol Microbiol* 88:590-602.
32. Piddock LJ (2006) Multidrug-resistance efflux pumps? not just for resistance. *Clin Microbiol Rev* 19:382-402.
33. Otwinowski Z, Minor M (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307-326.
34. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) *Phaser* crystallographic software. *J Appl Crystallogr* 40:658-674.
35. Terwilliger TC (2001) Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallogr D* 57:1755-1762.
36. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126.
37. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger, TR, McCroy AJ, Moriarty NW et al (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr* 58:1948-1954.
38. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren

GL (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D54:905-921.

Figures and captions



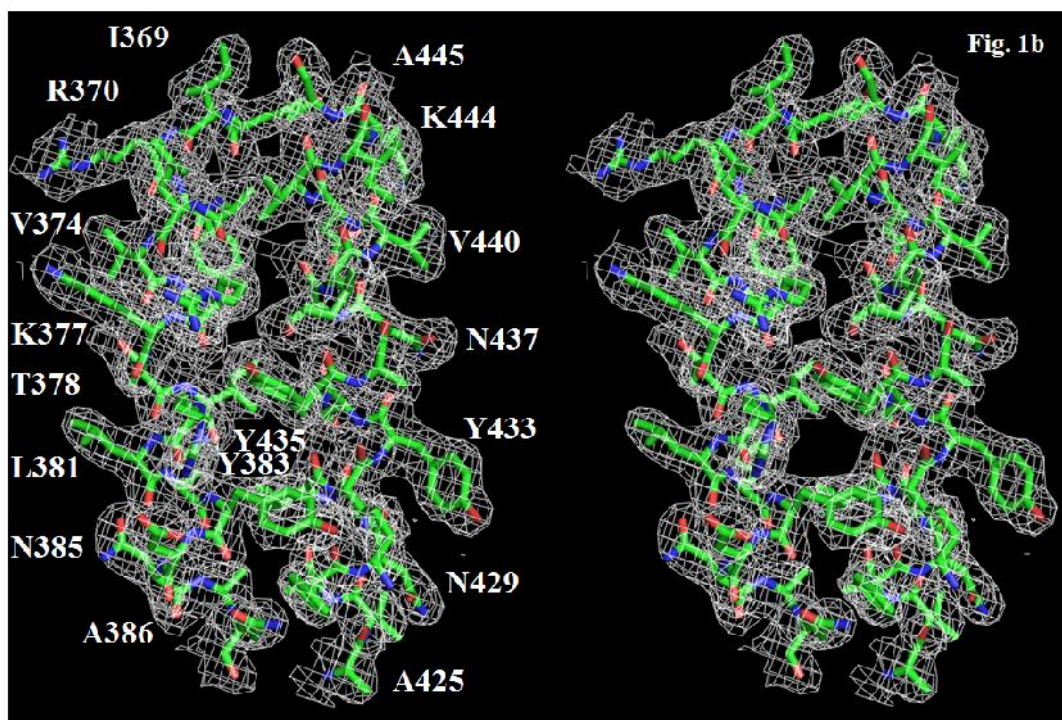


Fig. 1. Stereo view of the experimental electron density map of the CmeC channel at a resolution of 2.37 Å. (a) Anomalous maps of the 15 selenium sites (contoured at 4 σ). The selenium sites corresponding to the four methionines from each protomer of the CmeC trimer are in orange. The C α traces of CmeC are in green, magenta and white, respectively. (b) Representative section of the electron density at the interface between H8 and H9 of the periplasmic domain of CmeC. The electron density (colored white) is contoured at the 1.2 σ level and superimposed with the final refined model (green, carbon; red, oxygen; blue, nitrogen).

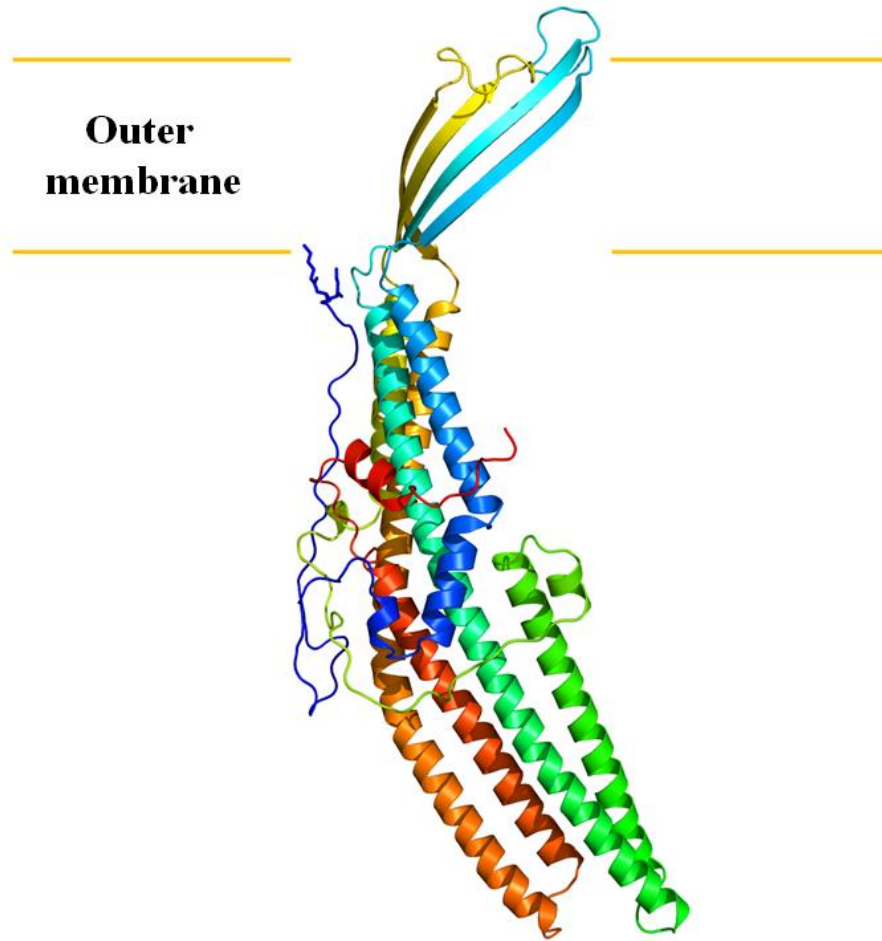
Fig. 2a

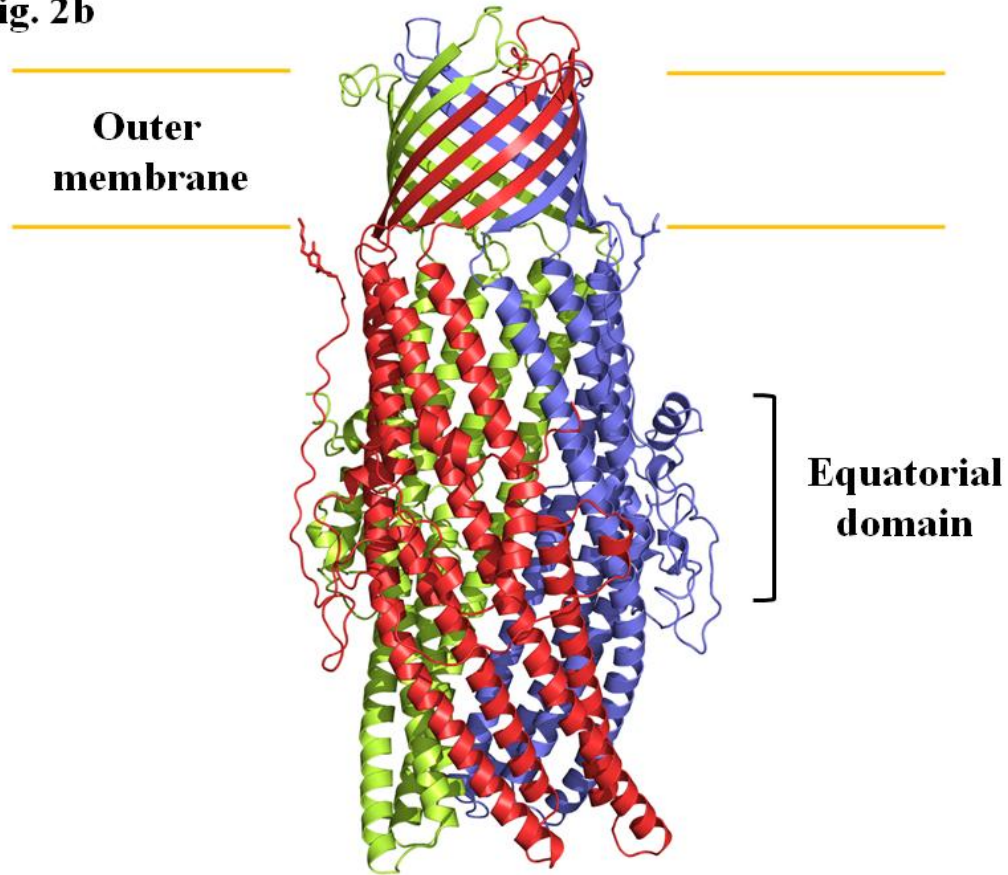
Fig. 2b

Fig. 2. Structure of the *C. jejuni* CmeC channel protein. (a) Ribbon diagram of a protomer of CmeC viewed in the membrane plane. The molecule is colored using a rainbow gradient from the N-terminus (blue) to the C-terminus (red). (b) Ribbon diagram of the CmeC trimer viewed in the membrane plane. Each subunit of CmeC is labeled with a different color. The CmeC protomer is acylated (in sticks) through the first N-terminal cysteine residue to anchor onto the outer membrane.

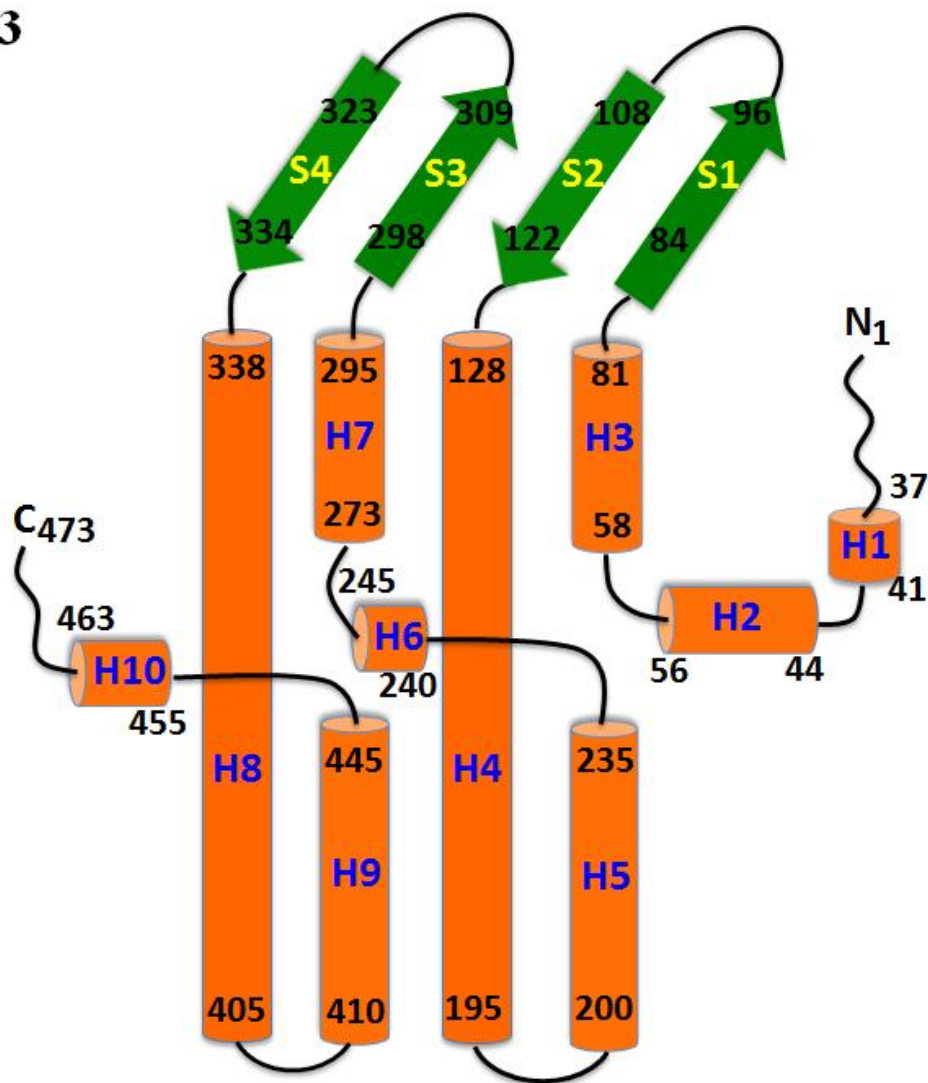
Fig. 3

Fig. 3. Secondary structural topology of the CmeC monomer. The topology was constructed based on the crystal structure of CmeC. The α -helices and β -strands are colored orange and green, respectively.

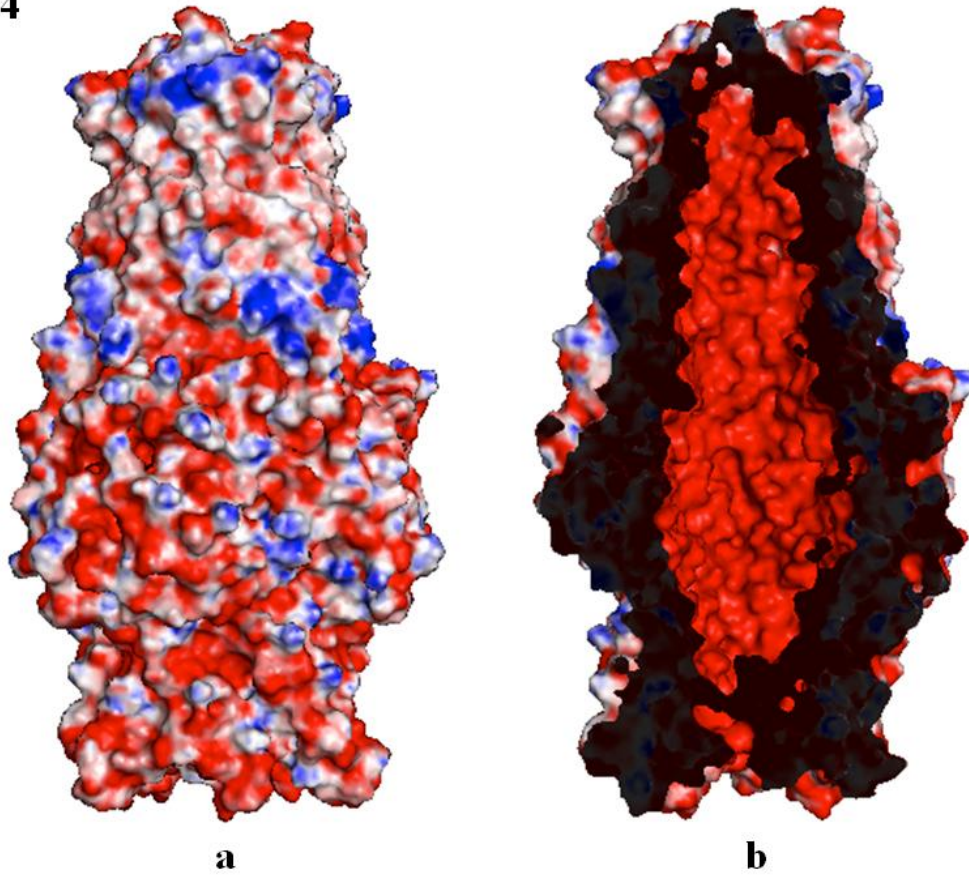
Fig. 4

Fig. 4. Electrostatic surface potentials of CmeC. Surface representations of the (a) outside and (b) inside the CmeC channel colored by charge (red; negative -15 kT/e, blue; positive +15 kT/e).

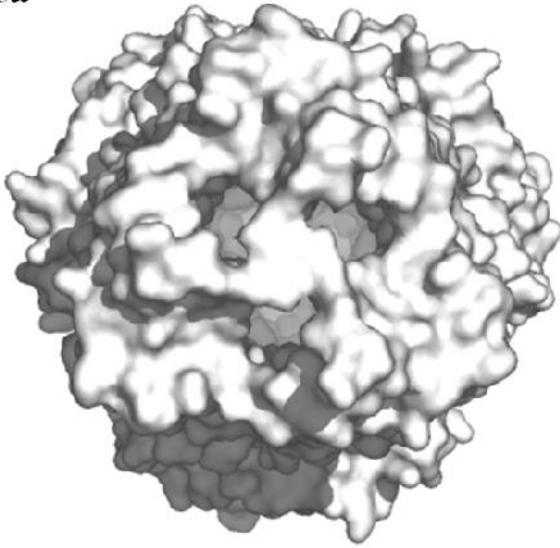
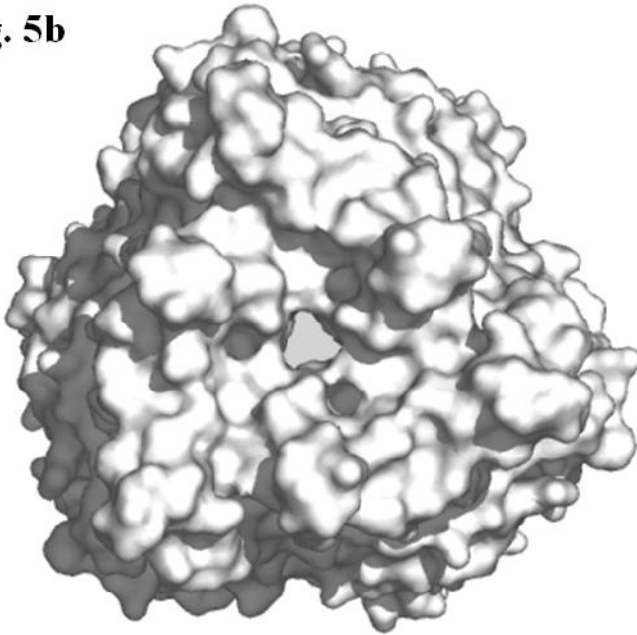
Fig. 5a**Fig. 5b**

Fig. 5. Surface representations of the trimeric CmeC channel. The views from both the (a) extracellular and (b) periplasmic sides suggest that the CmeC channel is in its closed form.

Fig. 6a

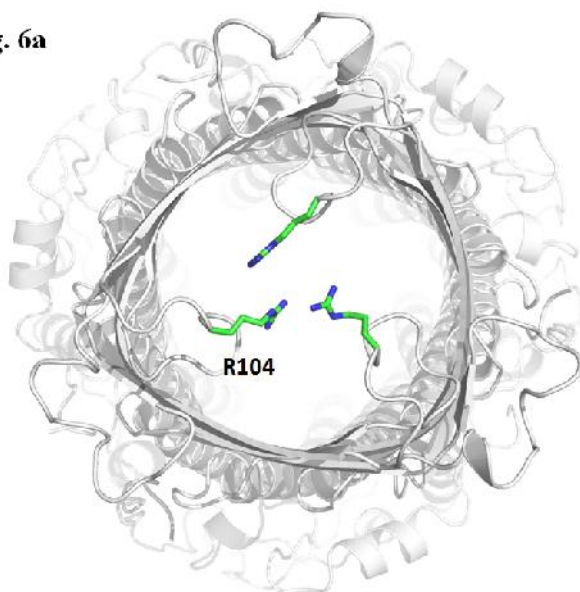


Fig. 6b

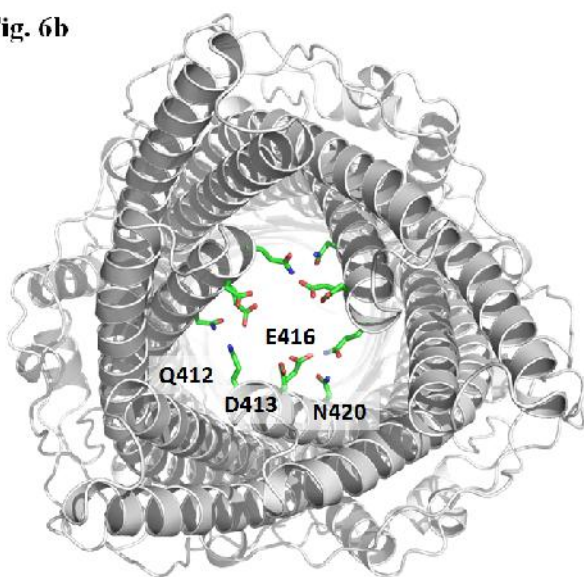


Fig. 6. The interior of the CmeC channel. (a) Extracellular view of the CmeC trimer. The three R104 residues of the trimer are found to interact and block the channel. (b) Periplasmic view of the CmeC trimer. The charged and polar residues Q412, D413, E416 and N420 are found to interact with each other and block this end of the channel.

Table 1. Data collection, phasing and refinement statistics of CmeC.

Data set	CmeC	SeMet-CmeC (peak)
Data Collection		
Wavelength (Å)	0.98	0.98
Space group	C222 ₁	C222 ₁
Resolution (Å)	50 – 2.37 (2.46-2.37)	50 – 3.10 (3.21-3.10)
Cell constants (Å)		
a	92.38	91.95
b	147.35	147.20
c	420.43	419.13
α, β, γ (°)	90, 90, 90	90, 90, 90
Molecules in ASU	3	3
Redundancy	3.3 (3.0)	3.0 (3.0)
Total reflections	2911,371	1203,938
Unique reflections	112,815	51,984
Completeness (%)	97.7 (93.6)	90.7 (83.4)
R _{sym} (%)	10.9 (42.8)	10.5 (44.9)
I/ σ (I)	12.8 (1.3)	12.9 (2.3)
Phasing		
Number of sites		15
Figure of merit (acentric/centric)		0.75/0.62
Refinement		
Resolution (Å)	50-2.37	
R _{work}	20.97	
R _{free}	24.32	
rms deviation from ideal		
bond lengths (Å)	0.002	
bond angles (°)	0.492	
Ramachandran plot		
most favoured (%)	95.7	
additional allowed (%)	4.3	
generously allowed (%)	0	
disallowed (%)	0	

**CHAPTER 4. CRYSTAL STRUCTURE OF NITRITE CHANNELS NAR1.1 and
NAR1.5 OF *CHLAMYDOMONAS REINHARDTII***

(A manuscript in preparation)

Abhijith Radhakrishnan^{1,ψ}, Tsung-Han Chou^{2,ψ}, Nitin Kumar¹, Jani Reddy Bolla³, Chih-Chia
Su², Martin H. Spalding³, and Edward W. Yu^{1,2*}

¹Department of Chemistry, Iowa State University, Ames, IA 50011, USA

²Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA

³Department of Genetics, Developmental and Cell Biology, Iowa State University, Ames, IA
50011, USA

^ψA.R. and T.C. contributed equally to this work.

* To whom correspondence should be addressed. E-mail: ewyu@iastate.edu

Introduction

Nitrogen is one of the essential elements for cell metabolism in all the living matter. The major sources of nitrogen for photosynthetic organisms such as green algae and plants are nitrate and ammonium which are found abundantly in soil and surface of the ocean. The uptake and conversion of nitrate found in minerals to organic nitrogen containing compound inside the cell plays a huge role in sustaining life on earth. The nitrate assimilation in photosynthetic eukaryotes undergo five successive stages. Nitrate is transported into the cell by plasma membrane nitrate transporters, then a cytosolic enzyme Nitrate Reductase (NR) reduces nitrate to nitrite. Subsequently, plastidic nitrite transporters permit the entry of nitrite into the chloroplast, where the Nitrite Reductase (NiR) enzyme catalyzes the reduction of nitrite to ammonium [1, 2, 3]. Ultimately, ammonium is incorporated into the carbon framework through the glutamine synthase/glutamine oxoglutarate transferase or glutamate synthase (GS/GOGAT) cycle [4, 5].

In the photosynthetic green algae *Chlamydomonas reinhardtii*, three families of transporters NRT1, NRT2 and NAR1, are known to take part in nitrate assimilation [2, 3, 6, 7, 8, 9, 10]. A unique NRT1 transporter belong to NRT1/PTR family also called NPF [11] can be found in *C. reinhardtii* genome [2]. The *NRT2* genes can be found in plants, algae, fungi, yeast and bacteria [7, 8, 13, 14]. The NRT2 transporters are members of Nitrate Nitrite Porter (NNP) family which belongs to MFS superfamily [10, 12] and have been marked as high affinity nitrate/nitrite transporters (HANNiT). Some of the NRT2 transporters require additional component NAR2 to be fully functional [8]. Both NRT1 and NRT2 proteins are located at plasma membrane [3, 8, 10].

C. reinhardtii genome encodes six genes in *NAR1* family (*NAR1.1-NAR1.6*) [20, 21]. *NAR1.1*, *NAR1.2* and *NAR1.5* possess putative chloroplast transit peptides which indicate a

possible involvement in nitrite transport to chloroplast [21]. The NAR1 proteins correspond to Formate/Nitrate Transporter (FNT) family which has been extensively identified in bacterial genomes. The structural and functional studies of bacterial homologs have shown that FNT-like proteins are capable of transporting various anions [15, 16, 17, 18, 19]. But their structure and function are not extensively studied in eukaryotes. *Nar1.1* was found clustered with other genes which are essential to nitrate assimilation and controlled by the nitrate-specific regulatory gene *Nit2* [22]. The previous studies have shown the expression of *NAR1.1* was nitrate induced and ammonium repressed [21]. Studies also have revealed that NAR1.1 is located at chloroplast envelope and is permeable to nitrite [22]. Moreover, it is suggested that lack of the NAR1.1 leads to a metabolic imbalance of nitrate incorporation to the chloroplast [23]. NAR1.2, however, is a chloroplast bicarbonate transporter which plays an important role in CO₂-concentrating mechanism (CCM) in *C. reinhardtii*. NAR1.2, or limiting CO₂ inducible A (LCIA) allow the movement of bicarbonate into chloroplast envelope under very low CO₂ [24, 25]. As for NAR1.5, a plastidic or mitochondrial transit peptide was predicted but the actual localization is still unknown. The expression of *NAR1.5* did not exhibit a sustained response to nitrogen or carbon sources but showed a nitrate-dependent up-regulation [21]. Here we report the crystal structures of plastidic nitrite transporters NAR1.1 and NAR1.5 to a resolution of 2.4 Å and 3.3 Å respectively. *In-vitro* light scattering experiments have revealed that both NAR1.1 and NAR1.5 conduct nitrite ions and have identified important residues playing key roles in nitrite recognition and transport.

Results and discussion

Structure of NAR1.5

Nar1.5 gene of the *Nar1* multigene family of *Chlamydomonas reinhardtii* encodes for a protein with 343 amino acid residues. Residues 1-78 of NAR1.5 were identified to form a

transit peptide and were removed before making the recombinant protein construct. After the removal of the transit peptide forming residues the NAR1.5 was expressed in *E. coli* and was purified and crystallized in two space groups, *P21* and *C2*. NAR1.5 crystals diffracted to a resolution of 3.6 Å in *P21* space group and 3.2 Å in *C2* (Table S1). The structure in *P21* was solved by molecular replacement accompanied by the single-wavelength anomalous dispersion (SAD) (Figure S1a, S1b). The structure in *C2* was also determined based on the *P21* structure and led to the final model at 3.2 Å (Figure S1c). The asymmetric unit of NAR1.5 has five molecules organized to form a homopentamer (Figure S1a). The homopentamer of NAR1.5 reveals a pentagonal cylindrical structure (Figure 1a, 1b) with a dimension about 80 Å in diameter and 47 Å in height. A wineglass-like pore is located at the center of the cylinder with 8.0 Å radius at the widest region and 2.5 Å at the narrowest region close to the stromal side (Figure 1a). The central pore is surrounded by hydrophobic residues which is in agreement with the structure of *stFocA* and *stNirC*, indicating that substrate transport through this pore is unlikely. The charge distribution on the water accessible surfaces of NAR1.5 is highly positive at the center of each protomer which is observed in all the FNT family structures whose structures are available in PDB (Figure 1c, 1d). However, in contrast to *stFocA* which possess a highly charged periplasmic and cytoplasmic surface, NAR1.5 has neutral periplasmic and cytoplasmic surfaces hinting at a distinct functional role and transporting mechanism. NAR1.5 pentamer is formed by five identical residues with a pairwise r.m.s.d. of 0.73 Å -0.85 Å. There are ten α -helical segments in each monomer and eight of them contribute to form six transmembrane helices (TMs 1-6); TM1 (104-131), TM2 (140-179), TM3 (184-212), TM4 (233-255), TM5 (260-295) and TM6 (301-331). TM2 and TM5 are split by loop regions L2 (161-166) and L5 (277-283) into two α -helical segments TM2a, TM2b and TM5a, TM5b respectively (Figure 2a). There are two short hydrophilic α -helices α_1 (85-100) at the N-terminal and α_2 (219-229) between TM3 and TM4. NAR1.5 is

localized in the chloroplast envelope membrane which separate the organelles of the chloroplast from the cytoplasm. NAR1.5 folds into the chloroplast membrane in such a way that the N-terminal and C-terminal ends are on the same side facing the stroma. The other side faces the cytoplasm. Each monomer of the NAR1.5 homopentamer possesses a channel that is predominantly hydrophobic (Figure 3d). The polar residues forming the pore are contributed by L2 and L5 loops. The channel has its entrance in the cytoplasmic side and extends towards the stromal side and helps in the movement of substrates from the cytoplasm towards stroma. The channel has a funnel shaped entrance in the cytoplasmic side formed by hydrophilic residues of L5. The pore forming residues are L129, K144, F151, L155, T159, L165, T167, G168, L224, K228, V247, V251, F274, V275, H281 and A284. HOLE was used to calculate the pore diameter along the axial passage (Figure 3f). The pore also has a central vestibule which is separated from the cytoplasmic and stromal side by two constriction sites; constriction site-1 (CS1) close to the cytoplasmic side and constriction site (CS2) close to the stromal side. CS1 is the narrowest region of the channel and is formed by residues F151, F274 and H281. H281 lies on L5 whereas F151 and F274 are from TM2a and TM5a. These three residues are highly conserved in FNT family proteins. CS2 is formed by residues L155, L165 and V251. L165 is a part of L2 while L155 belongs to TM2a and V251 belongs TM4. The location of L2 and L5 suggests that the residues forming these loops play a key role in substrate recognition and regulation of their movements. L2 and L5 being flexible means that their positioning plays a significant role in functioning of NAR1.5 as a channel. The electrostatic surface potential shows that the inner lining of the channel is highly positive which augurs well with the fact that the channel transports anions (figure 3e).

Structure of NAR1.1

Nar1.1, another gene of the *Nar1* multigene family of *Chlamydomonas reinhardtii* encodes for a protein with 357 amino acid residues. Residues 1-86 of NAR1.1 forms a transit peptide and was not included for making the recombinant protein construct. After the removal of the transit peptide forming residues, the proteins shared 50% sequence identity NAR1.5. We purified and crystallized NAR1.1 after expressing the recombinant protein in *E. coli*. The crystals diffracted X-rays to low resolutions in $P2_12_12$ space group (Table S1). Subsequently 13 amino acid residues were deleted from the C-terminal of NAR1.1 and the crystals diffracted to 2.3 Å in $P3_121$ space group. The structure of NAR1.1 was solved by molecular replacement using the NAR1.5 structure (Figure S2a). Unlike NAR1.5, the asymmetric unit of NAR1.1 has ten molecules organized as two homopentamers (Figure S2a, S2b). The homopentamer of NAR1.1 also reveals a pentagonal cylindrical structure with a dimension about 80 Å in diameter and 47 Å in height (Figure 2b). A wineglass-like pore is located at the center of the cylinder with 8.0 Å radius at the widest region and 2.4 Å at the narrowest region near the stromal side. The central pore is surrounded by hydrophobic residues and is filled with five elongated electron densities which might represent the hydrophobic chain of β -D-glucopyranoside. These results are in agreement with the structure of *stFocA* and *stNirC*, indicating that substrate transport through this pore is unlikely. Each monomer of the pentameric NAR1.1 are identical with a pairwise r.m.s.d. of 0.20 Å -0.32 Å. There are ten α -helical segments in each monomer and eight of them contribute to form six transmembrane helices; TM1 (112-143), TM2 (148-187), TM3 (192-220), TM4 (241-264), TM5 (268-303) and TM6 (309-340). TM2 and TM5 are split by loop regions L2 (169-174) and L5 (285-290) into two α -helical segments (Figure 2a). There are two short hydrophilic α -helices; α 1 (93-109) at the N-terminal and α 2 (227-238) between TM3 and TM4. NAR1.1 is localized in the chloroplast envelope membrane which separate the organelles of the

chloroplast from the cytoplasm. NAR1.1 folds into the membrane in a way that the N-terminal and C-terminal ends are on the same side facing the stroma. The other side faces the cytoplasm. Each monomer has a channel that is predominantly hydrophobic. The polar residues forming the pore are contributed by L2 and L5 loops. The channel has its entrance in the cytoplasmic side and extends towards the stromal side and helps in the movement of substrates from the cytoplasm towards stroma (Figure 3a). The channel has a funnel shaped entrance in the cytoplasmic side formed by hydrophilic residues of L5. The pore forming residues are K107, I126, F159, L163, T166, L167, A171, L173, M232, K236, V255, V259, V279, F282, I283, H289, A292 and F296. HOLE was used to calculate the pore diameter along the axial passage (Figure 3a, 3b, 3c). The pore also has a central vestibule which is separated from the cytoplasmic and stromal side by two constriction sites; constriction site-1 (CS1) close to the cytoplasmic side and constriction site (CS2) close to the stromal side. CS1 is the narrowest region of the channel and is formed by residues F159, F282 and H289 (Figure 4b). H289 lies on L5 whereas F159 and F282 are from TM2a and TM5a. These three residues are highly conserved in FNT family proteins. CS2 is formed by residues L163, L173 and V255. L173 is a part of L2 while L163 belongs to TM2a and V255 belongs TM4. The location of L2 and L5 suggests that the residues forming these loops play a key role in substrate recognition and regulation of their movements. L2 and L5 being flexible means that their positioning plays a significant role in functioning of NAR1.1 as a channel. The electrostatic surface potential shows that the inner lining of the channel is highly positive which augurs well with the fact that the channel transports anions (Figure 2c, 2d). The higher resolution of NAR1.1 allowed us to determine the interatomic distance between the residues that form the constriction sites.

Stopped flow light scattering experiments

In-vitro studies using stopped flow light scattering experiments were carried out on proteoliposomes to elucidate the substrate transport mechanism of NAR1.1 and NAR1.5. Previous studies suggest that both these proteins are involved in nitrite (NO_2^-) assimilation and regulation. Our results have shown that wild type of both NAR1.1 and NAR1.5 transport NO_2^- (Figure 5a). We have also identified three possible gating regions in NAR1.1 monomer. Two of these gating regions are salt triads formed by Lys-Glu-Asn residues and are related by pseudo-two-fold symmetry (Figure 4a). The salt triads are located at the entrance (cytoplasmic side) and exit (stromal side) of the channel suggesting that the triads might play a role in substrate recognition. The salt triad close to the cytoplasmic side is formed by the residues K236, E288 and N293 whereas the salt triad close to stromal side is formed by K107, E172 and N177 (Figure 4b). Interestingly, K107 and K236 lies on α_1 and α_2 respectively. We have observed that out of the three residues forming the salt triad only the lysine residues are found at 4.0 Å vicinity of the channel. This suggest that the only positively charged residue of the triad, lysine probably participates in the recognition of the negatively charged substrate at the entrance of the channel. To substantiate our claim we have done stopped flow light scattering studies using proteoliposomes by mutating the K107 residue to alanine (A) and K236 to glutamic acid (E), one at a time. The mutations were chosen in such a way that the salt triad formation is broken. The results of the light scattering experiments showed that the both the mutant proteins could not transport the nitrite ion (NO_2^-) across the liposome whereas the wildtype proteins aided the movement of NO_2^- (Figure 5b). The *in-vitro* studies prove that lysine residues in the salt triads is an important residue assisting in the movement of NO_2^- across the channel and establishes the salt triads as possible gating regions. The residues forming the salt triad are identical in both NAR1.1 and NAR1.5 which is consistent with the HSC. The residues forming the salt triad at the entrance

of the channel are highly conserved in all the FNT family proteins whose structures are known but the residues that form the other salt triad are less conserved. This suggests that the salt triad at the entrance of the channel (cytoplasmic side) is the possible recognition site where substrate selection occurs which is in agreement with the functional role of this protein in aiding the movement of NO_2^- from cytoplasm to stroma.

The narrowest region of the channel is formed by the constriction site near the cytoplasmic side CS1 with a pore radius of 0.95 Å (chain A of NAR1.1) (Figure 4a). CS1 is formed by two hydrophobic residues F159, F282 and a charged residue H289 (Figure 4b). All the three residues are highly conserved in the member of FNT family. Very small pore size at CS1 suggested that NO_2^- cannot pass through this narrowest pore region without a possible electrostatic interaction with H289. The other two residues being hydrophobic cannot form electrostatic interactions. It is also interesting to note that H289 is the only charged residue found in the entire middle region of the channel. This indicates that H289 might be playing a key role in the movement of NO_2^- towards the stroma. To study its impact we mutated H289 to hydrophobic residue alanine (A). Surprisingly, stopped flow light scattering studies carried out using the alanine mutant (H289A) failed to transport NO_2^- (Figure 5b). Alanine being a smaller residue meant that the mutated protein had a bigger pore radius at CS1. But there was no evidence of substrate transport. This strengthened our theory that charged H289 residue interacts with the NO_2^- and helps it move past the narrowest region of the channel, CS1. This indicates that Phe-Phe-His forms a gate close to the cytoplasmic side of the protein and acts as a filter to select NO_2^- . However its exact modus operandi is still unknown.

Conclusions

Nar1 multigene family of *Chlamydomonas reinhardtii* encodes six proteins belonging to the FNT family. Two of these proteins, NAR1.1 and NAR1.5 are found in the chloroplast

envelope membrane and aids in the transport of anions from cytoplasm towards stroma. Based on the crystal structures and *in-vitro* functional studies of NAR1.1 and NAR1.5 we have identified key residues and elucidated a transport mechanism for FNT family of proteins. The crystals structure revealed a homopentamer with each monomer forming a channel facilitating substrate passage. The channel, spans the entire monomer with a funnel shaped entrance at the cytoplasmic side and two constriction sites in the middle region. Two salt bridges formed by identical residues Lys-Glu-Asn guard the entrance and the exit of the channel. *In-vitro* studies using stopped flow light scattering studies on wildtype proteins have revealed NO_2^- as the substrate for NAR1.1 and NAR1.5. The single point mutagenesis studies have revealed that the lysine residue of the salt triad is essential for the functioning of the protein as a NO_2^- channel. This validated our hypothesis that the positively charged lysine interacts with the negatively charged NO_2^- and facilitates substrate recognition at the entrance and exit of the channel. Getting past the salt triad at the cytoplasmic side, NO_2^- reach the narrowest region of the channel, constriction site CS1 formed by Phe-Phe-His residues. Crystal structure of NAR1.1 indicates that the pore radius at CS1 is too small for the NO_2^- to pass through. Our results showed that mutating the histidine residue to alanine leaves the channel incapable of transporting NO_2^- . This unambiguously proves that positively charged histidine residue is essential for the functioning of the channel and interacts with the negatively charged NO_2^- and helps it move past the narrowest region of the pore thus unravelling the anion transport mechanism of FNT family proteins.

Materials and methods

Cloning, expression and purification of *C. reinhardtii* NAR1.5

After removing the first 78 residues that forms the chloroplast transit peptide, the full-length NAR1.5 (residues 79-343) membrane protein containing 6xHis tag at the N-terminus

was overproduced in *E. coli* BL21-CodonPlus *acrB ompF* cells possessing the pET15b expression vector. Cells were grown in 12 l of Luria broth medium with 100 µg/ml ampicillin at 37 °C. When the OD_{600 nm} reached 0.8, the culture was treated with 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) to induce the expression of NAR1.5, and the cells were harvested after 6 hr. The collected bacteria were suspended in a buffer containing 20 mM Na-HEPES (pH 7.5), 300mM NaCl, 20mM ethylenediaminetetraacetic acid (EDTA) and 10% glycerol and then disrupted with a French pressure cell. The membrane fraction was then collected and washed with a buffer containing 20 mM Na-HEPES (pH 7.5) and 50mM NaCl. The membrane protein was then solubilized in a buffer containing 20 mM Na-HEPES (pH 7.5), 20mM Imidazole-HCl (pH 7.2), 50mM NaCl and 2% (w/v) n-dodecyl-β-D-maltoside (DDM). Insoluble material was removed by ultracentrifugation at 100,000 × g. The extracted protein was purified with a Ni²⁺-affinity column. The purified protein was dialyzed and concentrated to 10 mg/ml in a buffer containing 20 mM Na-HEPES (pH 7.5), 100mM NaCl and 0.05% DDM. The 6xHis tag at the N terminus was then cleaved by adding 2 U of thrombin (GE Healthcare) per mg of purified NAR1.5 at room temperature for 20 hr. The protein was subsequently passed through a Ni²⁺-affinity column to remove the free 6xHis tag. A final purification step was performed using a G200 size exclusion column loaded with buffer solution containing 20 mM Tris-HCl (pH 7.5), 100mM NaCl and 1.1% n-octyl-β-D-glucopyranoside (β-OG). The purity of the NAR1.5 protein (>95%) was judged using 15% SDS-PAGE stained with Coomassie Brilliant Blue. The purified protein was then concentrated to 10 mg/ml in a buffer containing 20 mM Tris-HCl (pH 7.5), 100mM NaCl and 1.1% β-OG.

For 6xHis selenomethionyl-substituted (SeMet)-NAR1.5 protein expression, a 10 ml LB broth overnight culture containing *E. coli* BL21-CodonPlus *acrB ompF*/pET15b cells were transferred into 120 ml of LB broth containing 100 µg/ml ampicillin and grown at 37

°C. When OD_{600 nm} value reached 0.5, cells were harvested by centrifugation at 6,000 rev/min for 10 min and then washed two times with 20 ml of M9 minimal salts solution. The cells were re-suspended in 120 ml of M9 media and then transferred into a 12 l pre-warmed M9 solution containing 100 µg/ml ampicillin. The cell culture was incubated at 37 °C with shaking. When OD_{600 nm} reached 0.4, 100 mg/l of lysine, phenylalanine, and threonine, 50 mg/l isoleucine, leucine, and valine, and 60 mg/l of L-selenomethionine were added. The culture was induced with 0.2 mM IPTG after 15 min. Cells were then harvested 6 hr after induction. The procedures for purifying SeMet-NAR1.5 were identical to those of the native protein.

Cloning, expression and purification of *C. reinhardtii* NAR1.1

After removing the first 86 residues that forms the chloroplast transit peptide, the truncated NAR1.1 (residues 87-344) – with 13 amino acids removed from the C-terminus - membrane protein containing 6xHis tag at the N-terminus was overproduced in *E. coli* BL21-CodonPlus *acrB ompF* cells possessing the pET15b expression vector. Cells were grown in 12 l of Luria broth medium with 100 µg/ml ampicillin at 37 °C. When the OD_{600 nm} reached 0.8, the culture was treated with 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) to induce the expression of NAR1.1, and the cells were harvested after 6 hr. The collected bacteria were suspended in a buffer containing 20 mM Na-HEPES (pH 7.5), 300mM NaCl, 20mM ethylenediaminetetraacetic acid (EDTA) and 10% glycerol and then disrupted with a French pressure cell. The membrane fraction was then collected and washed with a buffer containing 20 mM Na-HEPES (pH 7.5) and 50mM NaCl. The membrane protein was then solubilized in a buffer containing 20 mM Na-HEPES (pH 7.5), 20mM Imidazole-HCl (pH 7.2), 50mM NaCl and 2% (w/v) n-dodecyl-β-D-maltoside (DDM). Insoluble material was removed by ultracentrifugation at 100,000 × g. The extracted protein was purified with a

Ni²⁺-affinity column. The purified protein was dialyzed and concentrated to 10 mg/ml in a buffer containing 20 mM Na-HEPES (pH 7.5), 100mM NaCl and 0.05% DDM. The 6xHis tag at the N terminus was then cleaved by adding 2 U of thrombin (GE Healthcare) per mg of purified NAR1.1 at room temperature for 20 hr. The protein was subsequently passed through a Ni²⁺-affinity column to remove the free 6xHis tag. A final purification step was performed using a G200 size exclusion column loaded with buffer solution containing 20 mM Tris-HCl (pH 7.5), 100mM NaCl and 1.1% n-octyl-β-D-glucopyranoside (β-OG). The purity of the NAR1.1 protein (>95%) was judged using 15% SDS-PAGE stained with Coomassie Brilliant Blue. The purified protein was then concentrated to 10 mg/ml in a buffer containing 20 mM Tris-HCl (pH 7.5), 100mM NaCl and 1.1% β-OG.

Crystallization of *C. reinhardtii* NAR1.5

Crystals of the NAR1.5 protein were obtained using hanging-drop vapor diffusion method. The NAR1.5 crystals were grown at 25 °C in 24-well plates using the following procedure. A 0.5 μl protein solution containing 10 mg/ml NAR1.5 in 20 mM Na-HEPES (pH 7.5), 100mM NaCl and 1.1% (w/v) β-OG was mixed with a 0.5 μl of reservoir solution containing 35% polyethylene glycol (PEG) 400, 0.05 M Tris-HCl (pH 8.2), 0.05 M Na₂SO₄, 0.05 M Li₂SO₄ and 0.005 M NaNO₃. The resultant mixture was equilibrated against 500 μl of the reservoir solution at 25 °C. The crystallization conditions for SeMet-NAR1.5 were the same as those for native NAR1.5. Crystals of both NAR1.5 and SeMet-NAR1.5 grew to a full size in drops within two weeks. Typically, the dimensions of the crystals were 0.2 mm x 0.2 mm x 0.4 mm. The condition was already cryo-protectant.

Crystallization of *C. reinhardtii* NAR1.1

Crystals of the NAR1.1 protein were obtained using hanging-drop vapor diffusion method. The NAR1.1 crystals were grown at 25 °C in 24-well plates using the following procedure. A 0.5 µl protein solution containing 10 mg/ml NAR1.5 in 20 mM Na-HEPES (pH 7.5), 100mM NaCl and 1.1% (w/v) -OG was mixed with a 0.5 µl of reservoir solution containing 43% polyethylene glycol (PEG) 400, 0.1 M Na-MES (pH 6.5), 0.06 M NaCl, 0.05 M Li₂SO₄. The resultant mixture was equilibrated against 500 µl of the reservoir solution at 25 °C. Crystals of both NAR1.1 grew to a full size in drops within a month. Typically, the dimensions of the crystals were 0.2 mm x 0.2 mm x 0.3 mm. The condition was already cryo-protectant.

Data Collection, Structural Determination, and Refinement

All diffraction data were collected at 100 K at Beamline 24ID-C located at the Advanced Photon Source, using a Pilatus 6M detector (Dectris Ltd., Switzerland). Diffraction data were processed using DENZO and scaled using SCALEPACK [26]. The NAR1.5 crystals belonged to the space group C2 (Table S1). Five NAR1.5 molecules were expected in the asymmetric unit with a solvent content of 53.96% based on the molecular weight of NAR1.5 (27.1 kDa). The experimental phases were solved by molecular replacement with FocA (PDBID: 3Q7K) as the template. The crystals of SeMet-NAR1.5 belong to space group P2₁2₁2 with a content of 62.50%. Within the asymmetric unit of SeMet-NAR1.5, 37 selenium sites were identified using SHELXC and SHELXD [27] as implemented in the HKL2MAP package [28]. The mature NAR1.5 protein contains eight methionine residues; all were identified as SeMet sites in each protomer. Single anomalous dispersion (SAD) was employed to obtain experimental phases using the program PHENIX [30]. The SeMet sites were then subjected to trace the molecules by anomalous difference Fourier maps, where we could assure the proper registry of SeMet residues. The initial model, which consisted of

pentameric molecules, was constructed manually using program Coot [29]. Then the model was refined using PHENIX [30], leaving 5% of reflections in the Free-R set. Iterations of refinement using PHENIX [30] and CNS [31] and model building in Coot [29] lead to the current model, which consists single pentameric structure with a total number of 1,254 residues.

The NAR1.1 crystal belongs to the space group $P3_121$, resulting in two NAR1.1 pentamers (26.0 kDa) in the asymmetric unit with a solvent content of 61.81%. The structure of NAR1.1 was determined using molecular replacement. The NAR1.5 pentamer model was used as a template. The resulting phases were then subjected to density modification using the program RESOLVE [32]. The model was constructed manually by Coot [29] then refined using PHENIX [30], leaving 5% of reflections in the Free-R set. Iterations of refinement using PHENIX [30] and CNS [31] and model building in Coot [29] lead to the current model of the NAR1.1 pentamers, consisting of a total number of 2,536 residues (Table S1).

Site-Directed Mutagenesis

We performed site-directed mutagenesis on residues K107, K236 and H289 of NAR1.1 to generate single point mutants K107A, K236E and H289A. All oligonucleotides were purchased from Integrated DNA Technologies in a salt-free grade.

Reconstitution and stopped flow light scattering assay

The purified wild-type NAR1.5, wild-type NAR1.1, NAR1.1 mutants K107A, K236E and H289A proteins were reconstituted into liposomes made of *E. coli* polar lipid and egg-yolk phosphatidylcholine (Avanti Polar Lipids) in a molar ratio of 3:1. In brief, 50 μg of the purified protein (wild-type or its mutant) was added to 5 mg of lipids dispersed in the reconstitution buffer containing 20 mM Na-HEPES (pH 7.5) and 100mM NaCl. The protein

samples were completely incorporated into liposomes as judged by using 15% SDS–PAGE stained with Coomassie blue. Control liposomes were prepared by following the same procedure without the addition of protein. The diameter of the proteoliposomes were measured as 368 ± 60 nm using Malvern Zetasizer Nano ZS.

To measure the nitrite permeability of target proteins and to identify the key residues involved in the nitrite transport, the liposomes or proteoliposomes were rapidly mixed with 100mM sodium nitrite at equal volume. The change in vesicle size was detected by recording the light scattering signal at 440 nm. The above experiments were performed on a stopped-flow apparatus (LS55; PerkinElmer). On the basis of a published protocol [33], changes in light scattering could be fitted by two exponentials in the equation $Y = [a(1 - e^{-kt}) - b]e^{-\mu t} + c$, where t is time and Y is the signal of light scattering. The first time-constant (k) corresponds to the rapid water efflux. The second time-constant (μ) corresponds to the relative rates of swelling again due to nitrite conduction¹. The data were analyzed with Origin.

Acknowledgements

This work was supported by an NIH Grant R01AI114629 (E.W.Y.). This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by an award GM103403 from the National Institutes of General Medical Sciences. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

References

1. Guerrero, M. G., Vega, J. M., and Losada, M. (1981). The assimilatory nitrate-reducing system and its regulation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 32, 169–204. doi: 10.1146/annurev.pp.32.060181.001125
2. Fernandez, E., and Galvan, A. (2007). Inorganic nitrogen assimilation in *Chlamydomonas*. *J. Exp. Bot.* 58, 2279–2287. doi: 10.1093/jxb/erm106
3. Fernandez, E., and Galvan, A. (2008). Nitrate assimilation in *Chlamydomonas*. *Eukaryotic Cell* 7, 555–559. doi: 10.1128/EC.00431-07
4. Mifflin, B. J., and Lea, P. J. (1975). Glutamine and asparagine as nitrogen donors for reductant-dependent glutamate synthesis in pea roots. *Biochem. J.* 149, 403–409. doi: 10.1042/bj1490403
5. Crawford, N.M. (1995). Nitrate: Nutrient and signal for plant growth. *Plant Cell* 7, 859–868.
6. Crawford, N. M., and Glass, A. (1998). Molecular and physiological aspects of nitrate uptake in plants. *Trends Plant Sci.* 3, 389–395 doi: 10.1016/S1360- 1385(98)01311-9
7. Forde, B. G. (2000). Nitrate transporters in plants: structure, function and regulation. *Biochem. Biophys. Acta* 1465, 219–235. doi: 10.1016/S0005- 2736(00)00140-1
8. Galvan, A., and Fernández, E. (2001). Eukaryotic nitrate and nitrite transporters. *Cell. Mol. Life Sci.* 58, 225–233. doi: 10.1007/PL00000850
9. Forde, B. G., and Cole, J. A. (2003). Nitrate finds a place in the sun. *Plant Physiol.* 131, 395–400. doi: 10.1104/pp.016139
10. Sanz-Luque, E.; Chamizo-Ampudia, A.; Llamas, A.; Galván, A.; Fernández, E. Understanding nitrate assimilation and its regulation in microalgae. *Front. Plant Sci.* 2015, 6, 899.

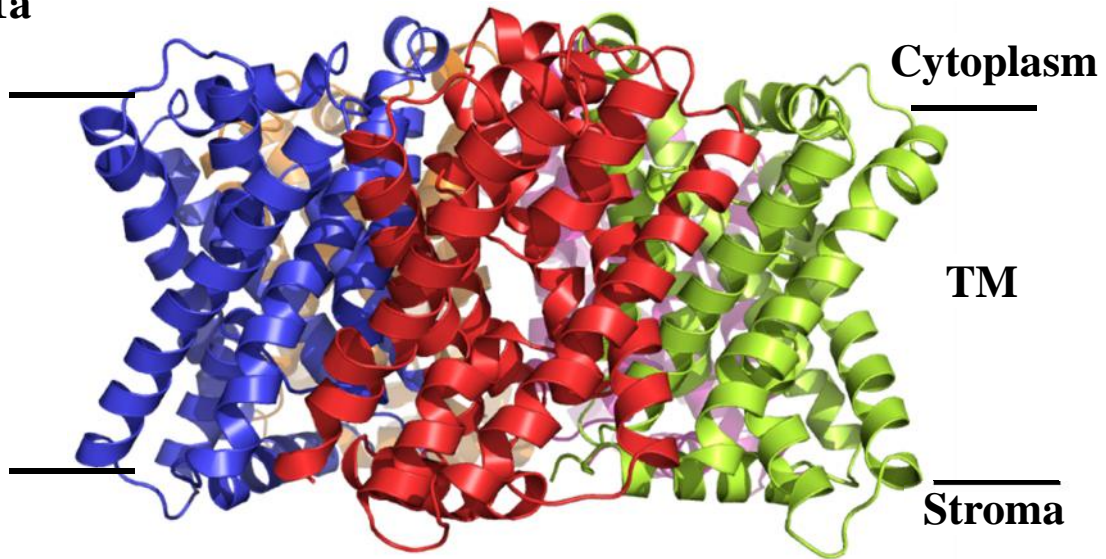
11. L eran, S., Varala, K., Boyer, J.-C., Chiurazzi, M., Crawford, N. M., Daniel- Vedele, F., et al. (2014). A unified nomenclature of NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family members in plants. *Trends Plant Sci.* 19, 5–9. doi: 10.1016/j.tplants.2013.08.008
12. Pao, S. S., Paulsen, I. T., and Saier, M. H. (1998). Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* 62, 1–34.
13. Orsel, M., A. Krapp, and F. Daniel-Vedele. (2002). Analysis of the NRT2 nitrate transporter family in Arabidopsis. Structure and gene expression. *Plant Physiol.* 129:886–896.
14. Tsujimoto, R., H. Yamazaki, S. Maeda, and T. Omata. (2007). Distinct roles of nitrate and nitrite in regulation of expression of the nitrate transport genes in the moss *Physcomitrella patens*. *Plant Cell Physiol.* 48:484–497.
15. Wang Y, et al. (2009) Structure of the formate transporter FocA reveals a pentameric aquaporin-like channel. *Nature* 462(7272):467–472.
16. Waight, A. B., Love, J. & Wang, D. N. Structure and mechanism of a pentameric formate channel. *Nature Struct. Mol. Biol.* 17, 31–37 (2010).
17. L  W, et al. (2011) pH-dependent gating in a FocA formate channel. *Science* 332 (6027):352–354.
18. L  W, Schwarzer NJ, Du J, Gerbig-Smentek E, Andrade SL, Einsle O. Structural and functional characterization of the nitrite channel NirC from *Salmonella typhimurium*. *Proc Natl Acad Sci USA.* 2012;109:18395–400. doi: 10.1073/pnas.1210793109.
19. B.K. Czyzewski, D.N. Wang Identification and characterization of a bacterial hydrosulphide ion channel *Nature*, 483 (2012), pp. 494–497

20. Galva'n A, Rexach J, Mariscal V, Ferná'ndez E (2002) Nitrite transport to the chloroplast in *Chlamydomonas*: molecular evidences for a regulated process. *J Exp Bot* 53: 845—853
21. Mariscal V, Moulin P, Orsel M, Miller AJ, Ferná'ndez E, Galva'n A. 2006. Differential regulation of the *Chlamydomonas* NAR1 gene family by carbon and nitrogen. *Protist* 157, 421–433.
22. Rexach J., Ferná'ndez E. & Galva'n A. (2000) The *Chlamydomonas reinhardtii* Nar1 gene encodes a chloroplast membrane protein involved in nitrite transport. *Plant Cell* 12, 1441–1454. Rexach
23. Mariscal V, Rexach J, Ferná'ndez E, Galva'n A. The plastidic nitrite transporter NAR1.1 improves nitrate use efficiency for growth in *Chlamydomonas*. *Plant, Cell and Environment* 2004;27:1321-1328.
24. Wang, Y., and Spalding, M. H. (2014). Acclimation to very low CO₂: contribution of limiting CO₂ inducible proteins, LCIB and LCIA, to inorganic carbon uptake in *Chlamydomonas reinhardtii*. *Plant Physiol.* 166, 2040–2050. doi: 10.1104/pp.114.248294
25. Yamano, T., Sato, E., Iguchi, H., Fukuda, Y., and Fukuzawa, H. (2015). Characterization of cooperative bicarbonate uptake into chloroplast stroma in the green alga *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7315–7320. doi: 10.1073/pnas.1501659112
26. Otwinowski, Z., and Minor, M. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326
27. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* 58, 1772–1779

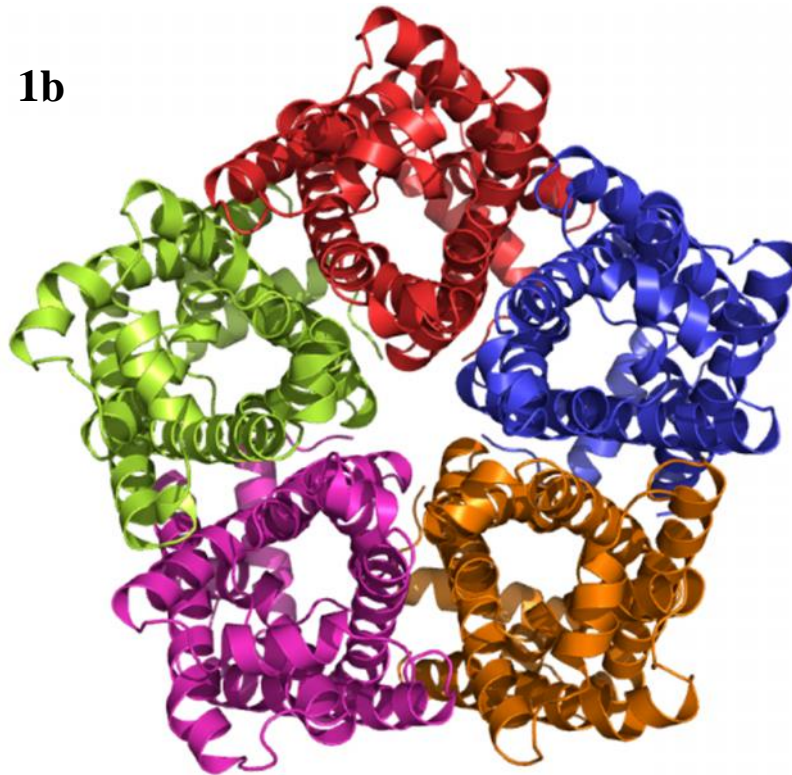
28. Pape, T., and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* 37, 843–844 37.
29. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2126–2132
30. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., Mc-Coy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 58, 1948–1954
31. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 54, 905–921 43.
32. Terwilliger, T. C. (2001) Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallogr. DBiol. Crystallogr.* 57, 1755–1762
33. Fu, D. et al. Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* 290, 481–486 (2000).
34. O. S. Smart, J. M. Goodfellow, B. A. Wallace, The pore dimensions of gramicidin A. *Biophys J* 65, 2455-2460 (1993).

Figures and captions

1a



1b



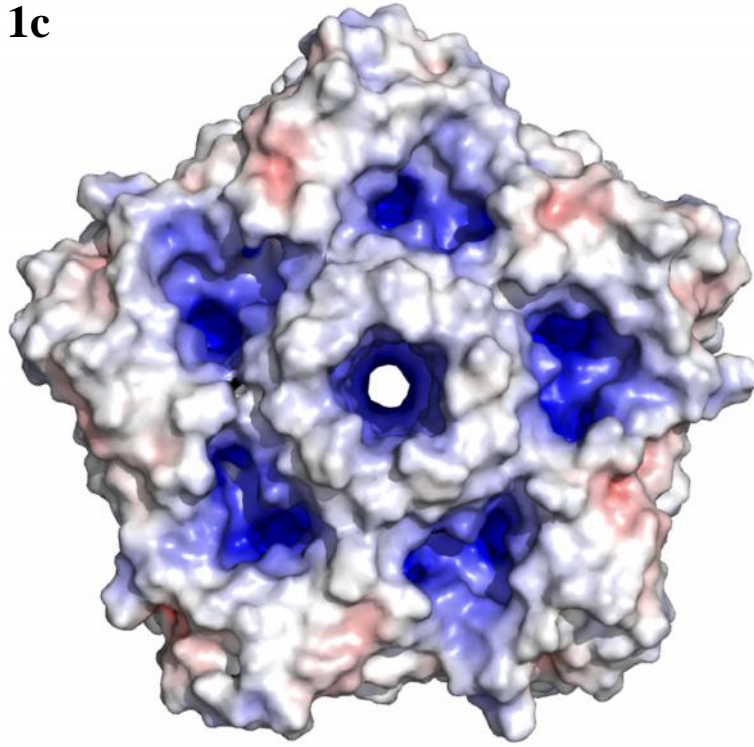
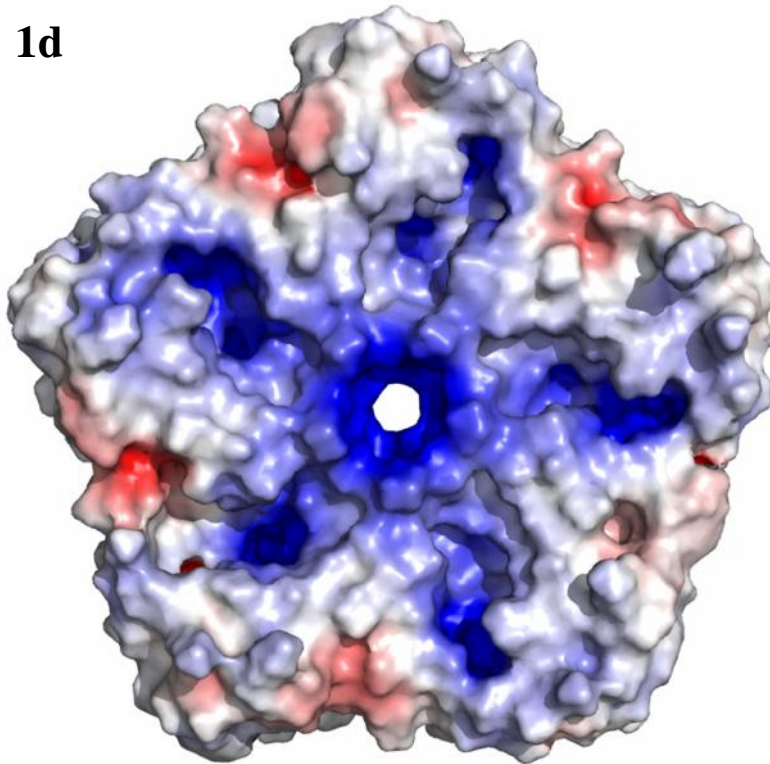
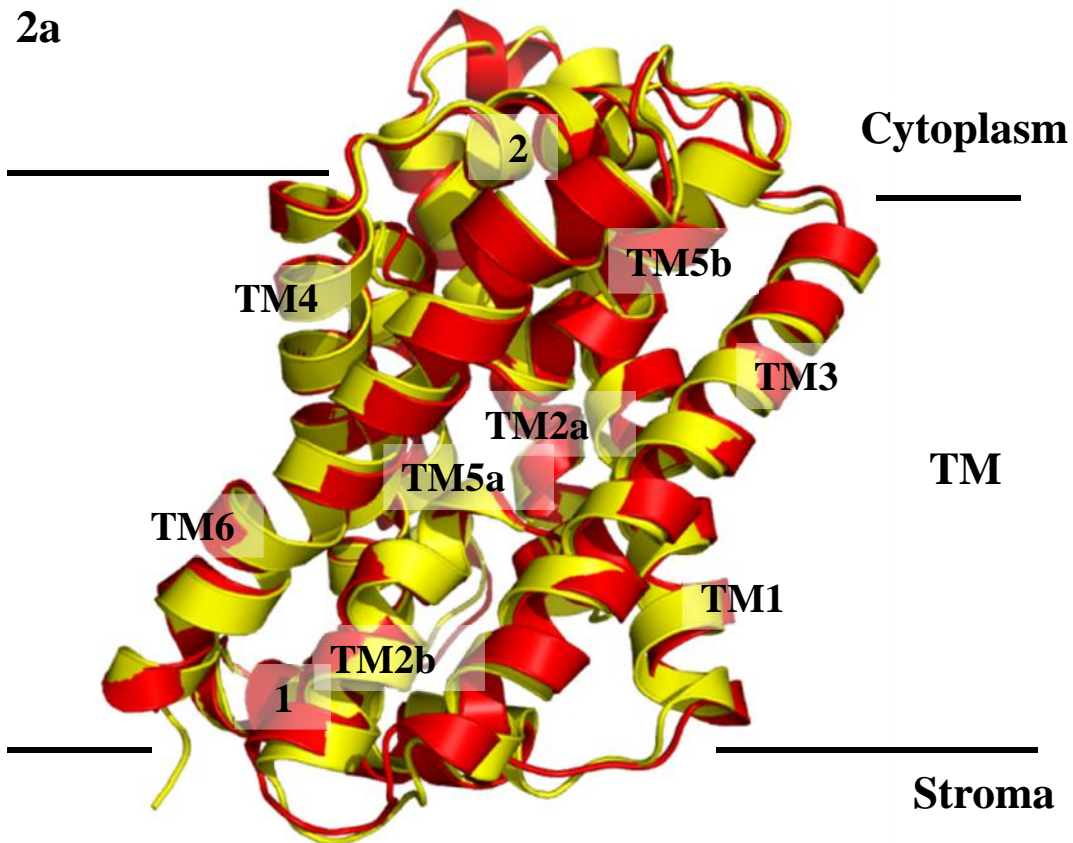
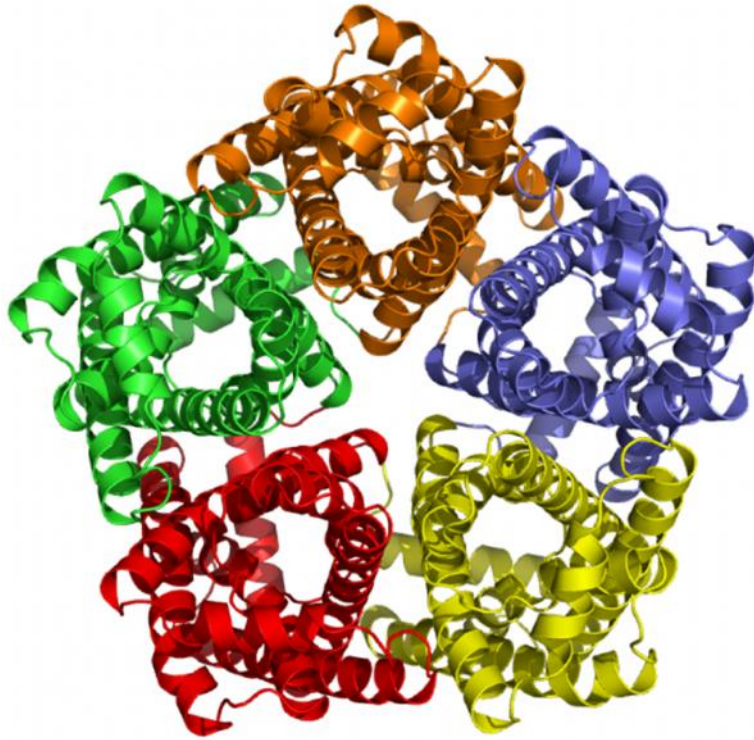
1c**1d**

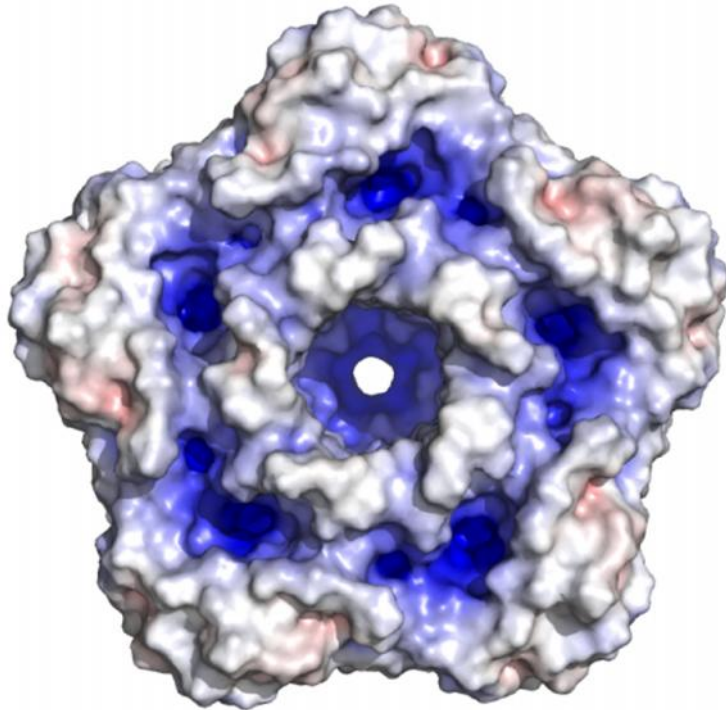
Fig. 1. Structure of the *Chlamydomonas reinhardtii* NAR1.5 (a) Ribbon diagram of a pentamer of NAR1.5 viewed in the membrane plane. The five protomers are colored red, blue, orange, purple and green respectively (b) Ribbon diagram of a pentamer of NAR1.5 as viewed from the top; perpendicular to the plane of the membrane; cytoplasmic view. (c) Surface electrostatic potential of the NAR1.5 pentamer as viewed from the cytoplasmic side (top view) (d) Surface electrostatic potential of the NAR1.5 pentamer as viewed from the stromal side (bottom view). The blue color in (c) and (d) corresponds to the positively charged patches of the pentamer. The Figure was prepared using PyMOL (<http://www.pymol.sourceforge.net>).



2b



2c



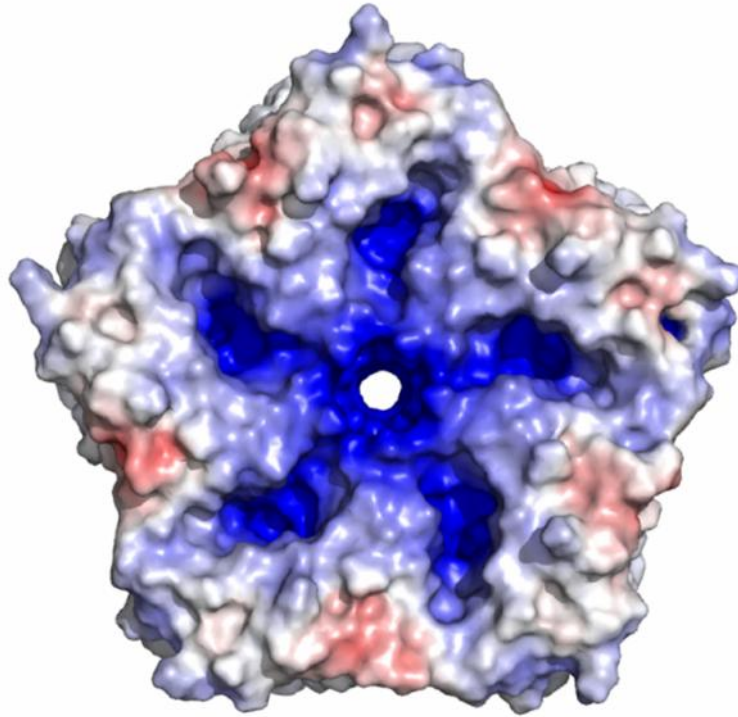
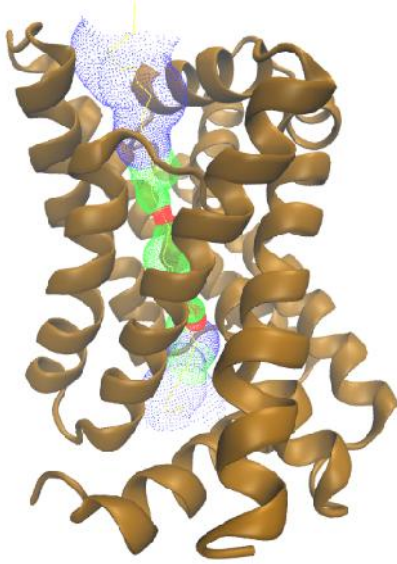
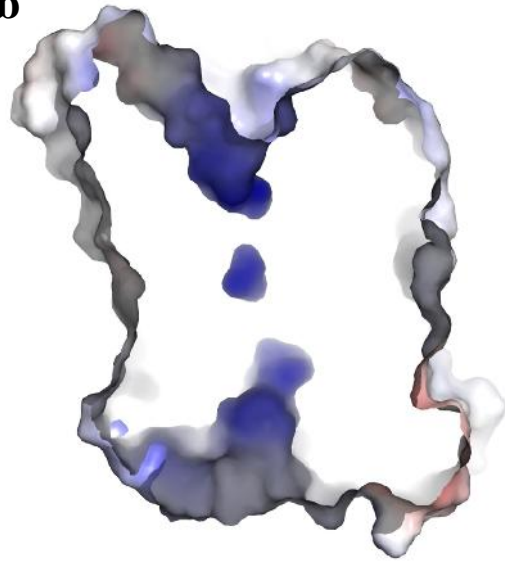
2d

Fig. 2. Structure of the *Chlamydomonas reinhardtii* NAR1.1 (a) Ribbon representation of the superimposition of the protomers of the NAR1.5 (red) and NAR1.1 (yellow) as viewed in the membrane plane. The transmembrane segments (TMs) and α -helices (α s) of the protomers are labeled (b) Ribbon diagram of a pentamer of NAR1.1 as viewed from the top; perpendicular to the plane of the membrane; cytoplasmic view. (c) Surface electrostatic potential of the NAR1.1 pentamer as viewed from the cytoplasmic side (top view) (d) Surface electrostatic potential of the NAR1.1 pentamer as viewed from the stromal side (bottom view). The blue color in (c) and (d) corresponds to the positively charged patches of the pentamer.

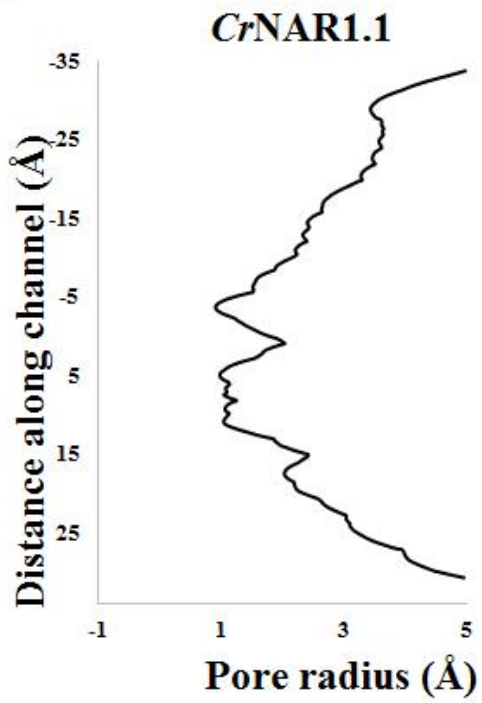
3a



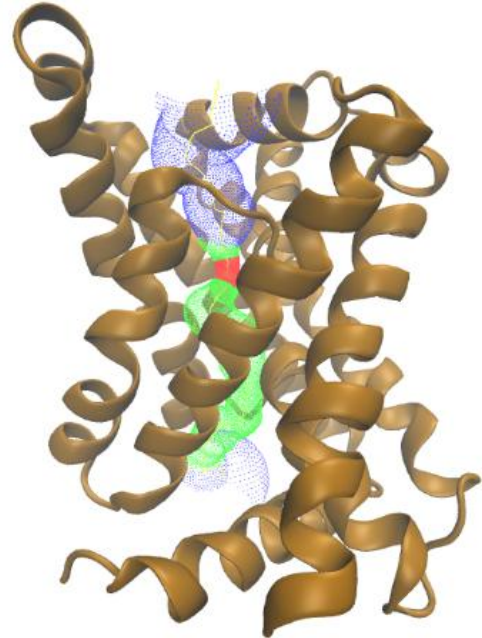
3b



3c



3d



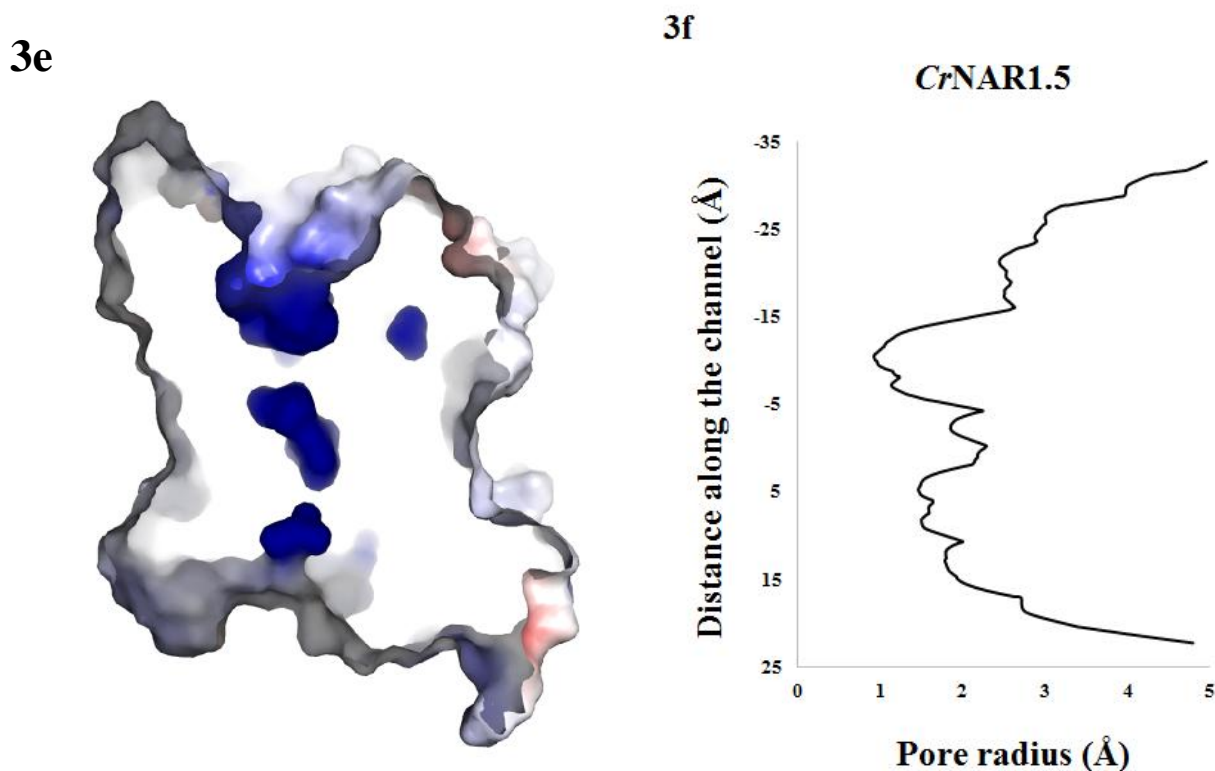
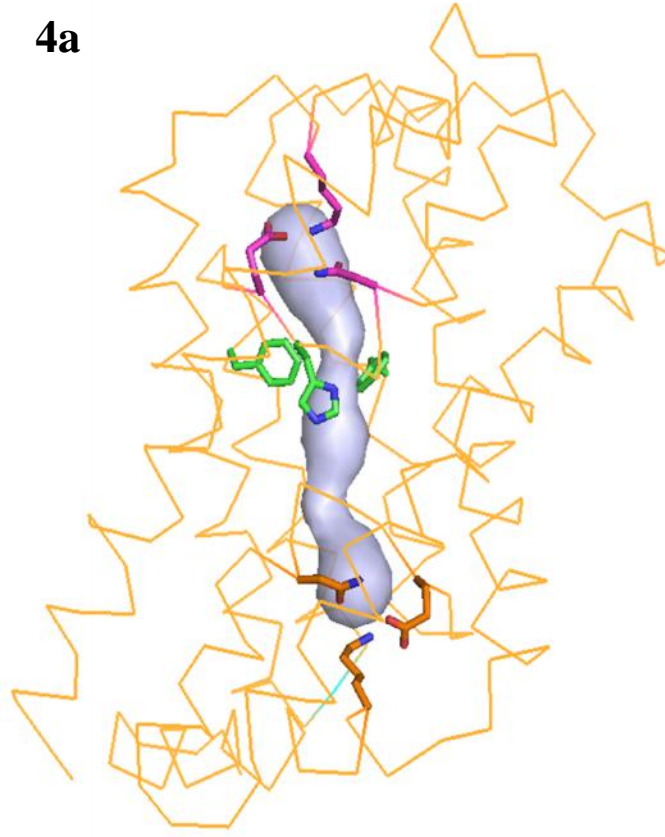


Fig. 3. Channel of NAR1.1 and NAR1.5. (a) Each subunit of NAR1.1 forms a channel (colored blue, green and red) spanning the entire transmembrane. This channel was calculated using the program HOLE (36). (b) Electrostatic surface potential properties of the cytoplasmic and stromal openings of the NAR1.1 channel facilitate the nitrite transport. (c) Structure profile of the NAR1.1 channel as calculated with HOLE (34). (d) Each subunit of NAR1.5 forms a channel (colored blue, green and red) spanning the entire transmembrane. This channel was calculated using the program HOLE (34). (e) Electrostatic surface potential properties of the cytoplasmic and stromal openings of the NAR1.5 channel facilitate the nitrite transport. (f) Structure profile of the NAR1.5 channel as calculated with HOLE (34).

4a



4b

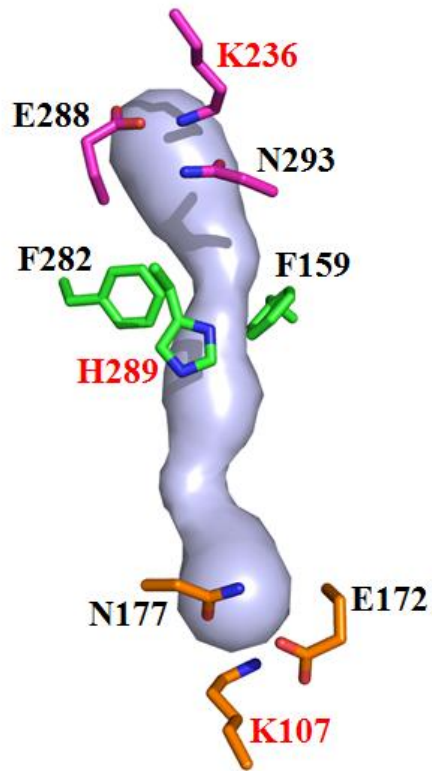
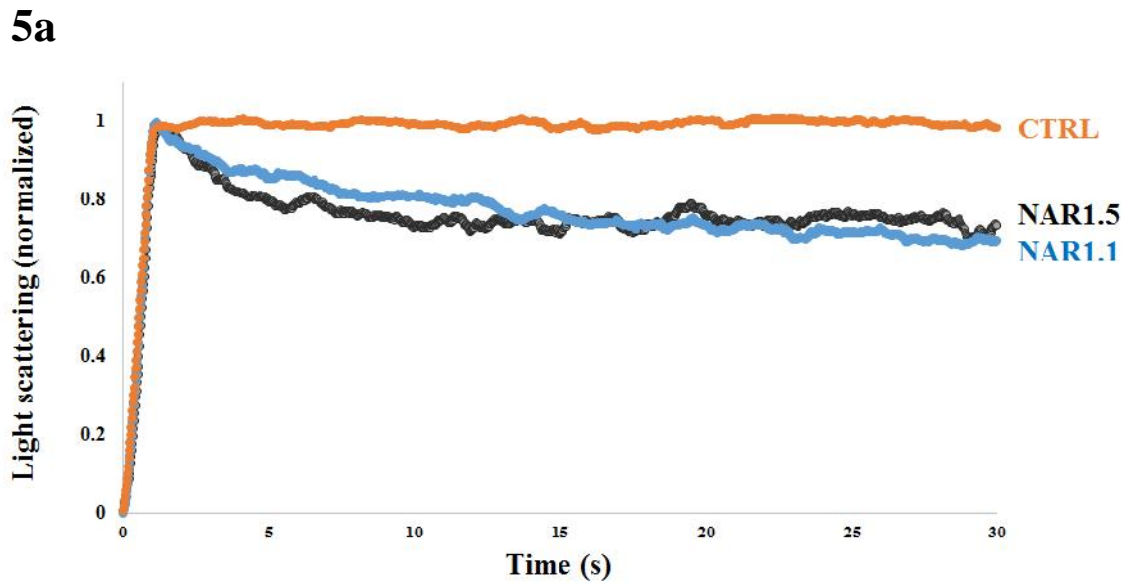


Fig. 4. Gating regions of NAR1.1 (a) The locations of the gating regions of a protomer of NAR1.1. The two salt triads; one at the cytoplasmic side and the other at the stromal side are colored in purple respectively. The residues in green form the narrowest region of the channel CS1. This channel was calculated using the program CAVER (<http://loschmidt.chemi.muni.cz/caver>). (b) K236, E288 and N293 form the cytoplasmic salt triad. K107, E172 and N177 form the stromal salt triad. F159, F282 and H289 form CS1. The residues colored red are positively charged.



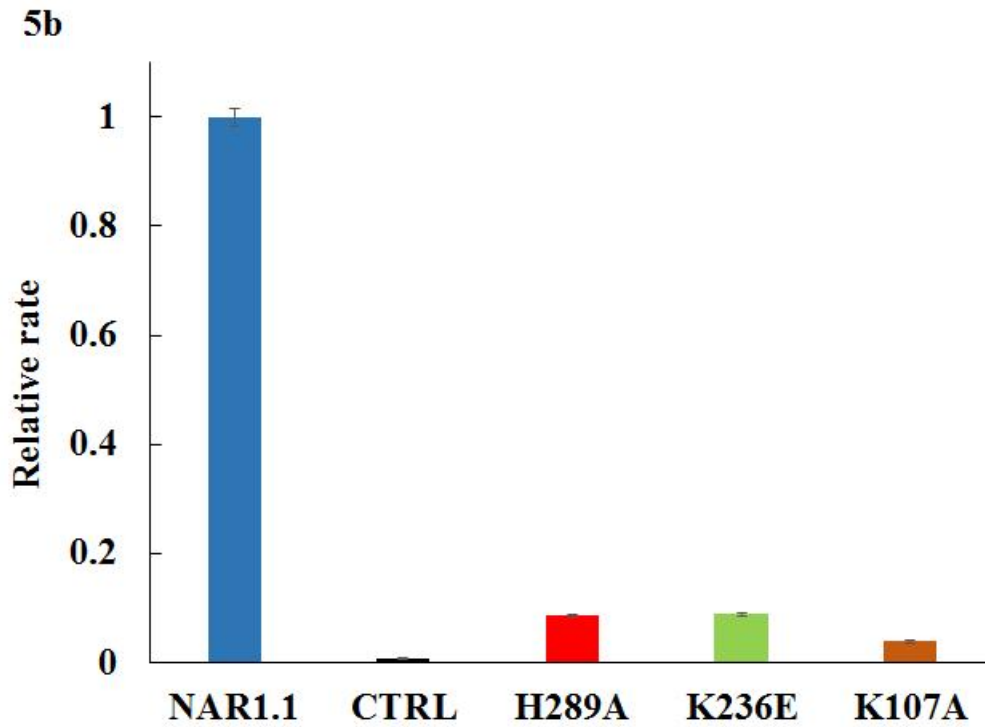
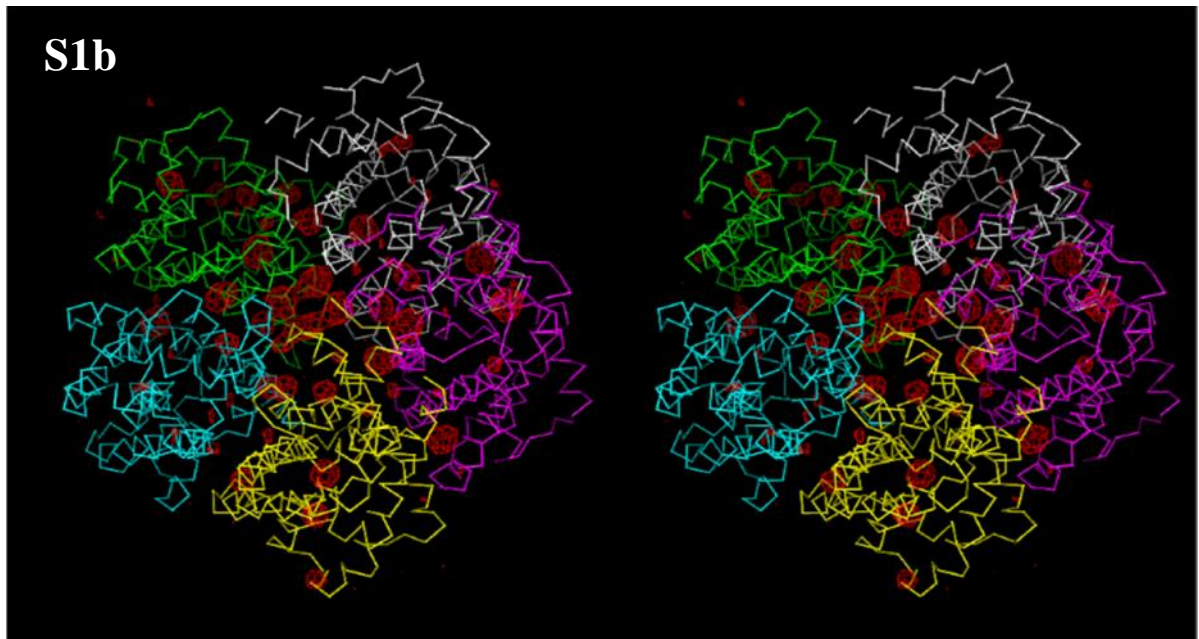
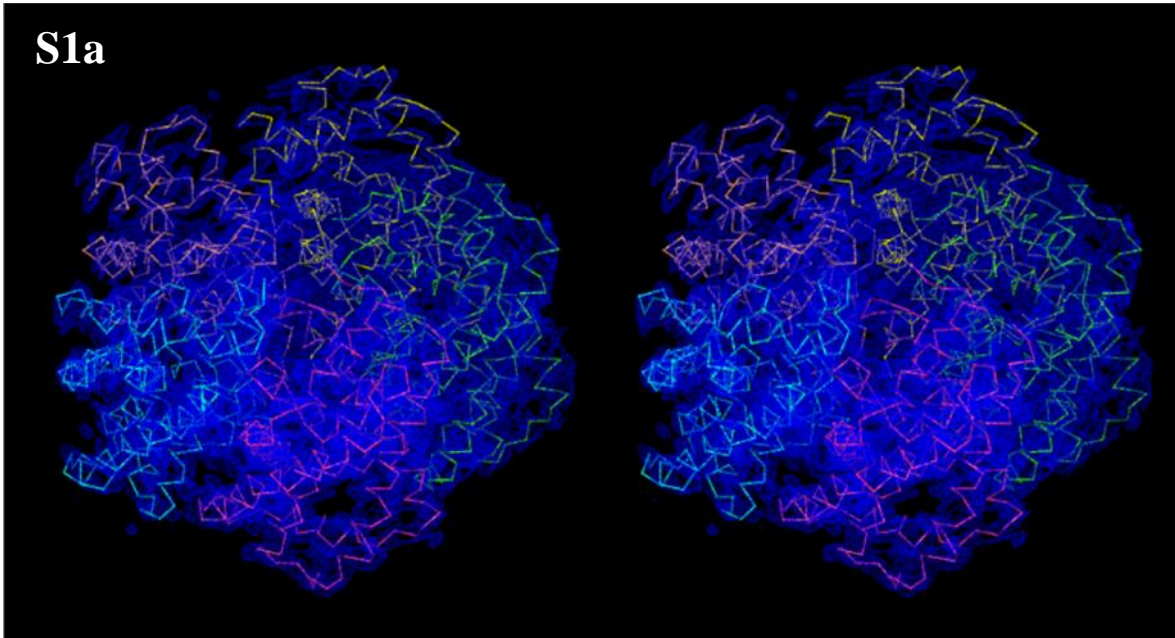


Fig. 5. Stopped flow light scattering experiment. (a) Protein-free liposomes and NAR1.1 and NAR1.5 loaded liposomes are mixed with 100mM NaNO₂ in a stopped flow apparatus. In response to high osmotic pressure water molecules diffused through the lipid, causing vesicle to deflate. The rapid changes in vesicle size are reflected by change in scattering. Both NAR1.1 and NAR1.5 are permeable to nitrite. (b) Site generated mutagenesis revealed that H289A, K236E and K107A mutants are not permeable to nitrite.

Supplemental Materials



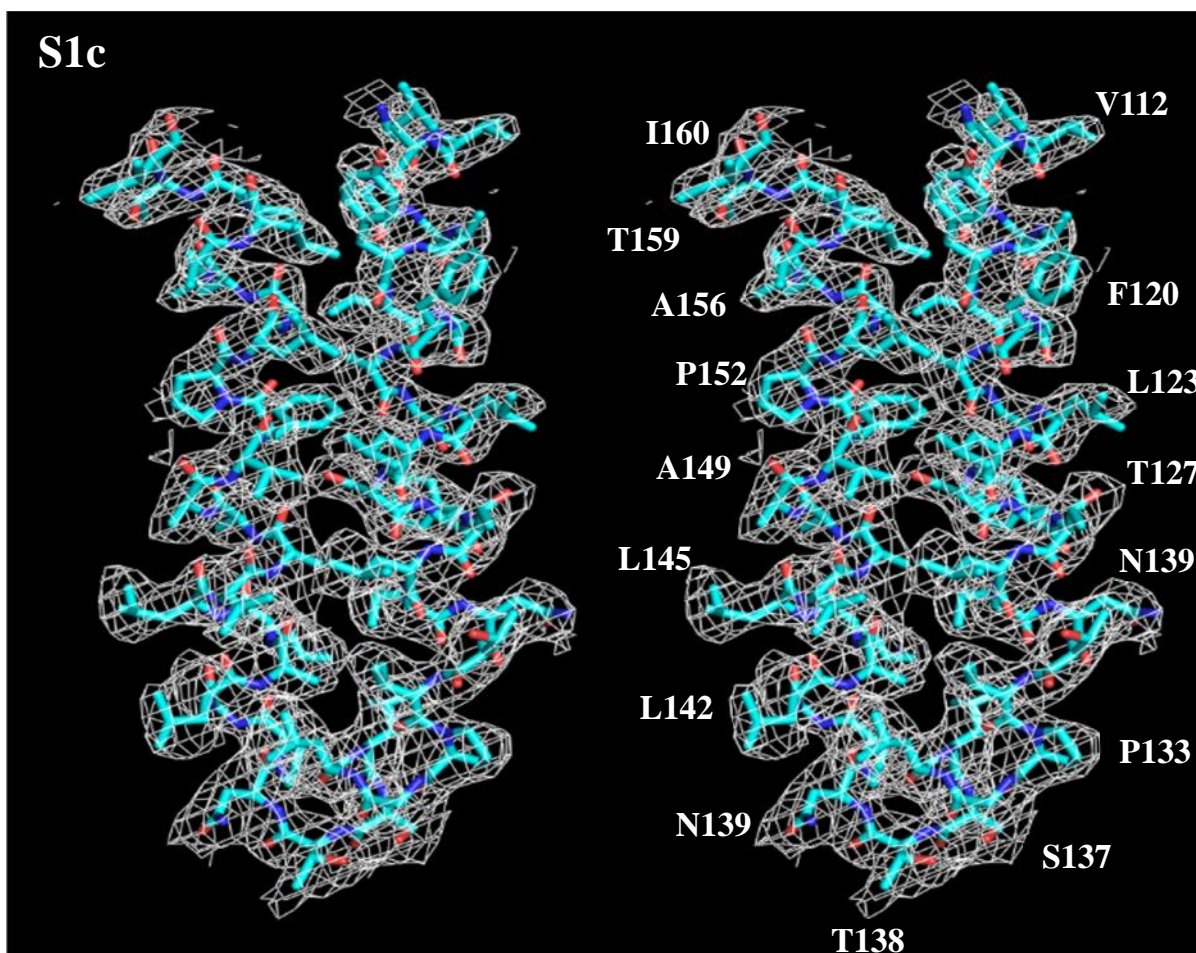


Fig. S1. Stereo view of the electron density maps of NAR1.5 at a resolution of 3.29 Å.

(a) The electron density maps are contoured at 1.2σ . The $C\alpha$ traces of the NAR1.5 pentamer in the asymmetric unit are included. (b) Anomalous maps of the 37 selenium sites (contoured at 3σ). Five protomers forming a pentamer of NAR1.5 are found in the asymmetric unit. Each protomer contributes 8 selenium sites corresponding to the 8 methionines (red). The $C\alpha$ traces of the five NAR1.5 monomers are colored green, white, purple, yellow and blue. (c) Representative section of the electron density in the vicinity of TMs 1 and 2a of NAR1.5. The electron density (colored white) is contoured at the 1.2σ level and superimposed with the final refined model (light blue, carbon; red, oxygen; dark blue, nitrogen).

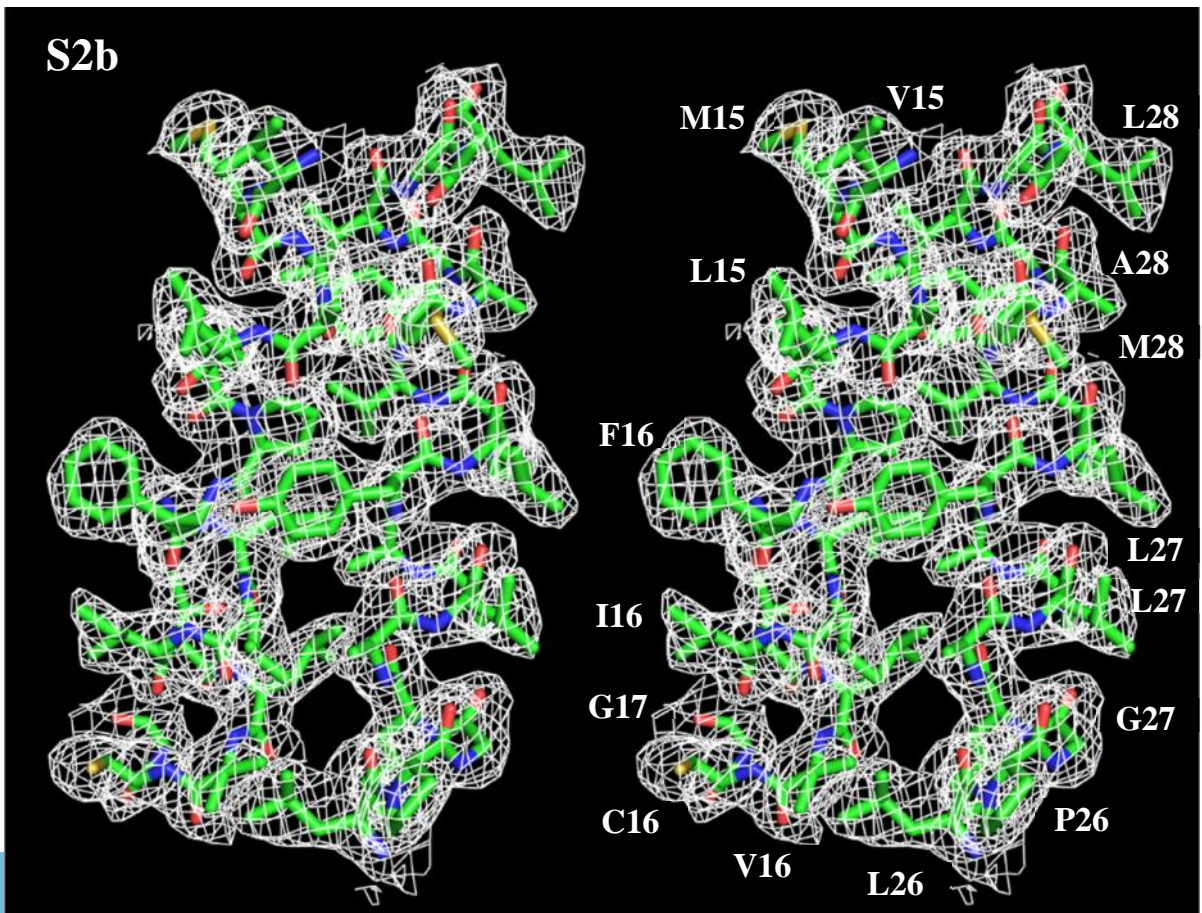
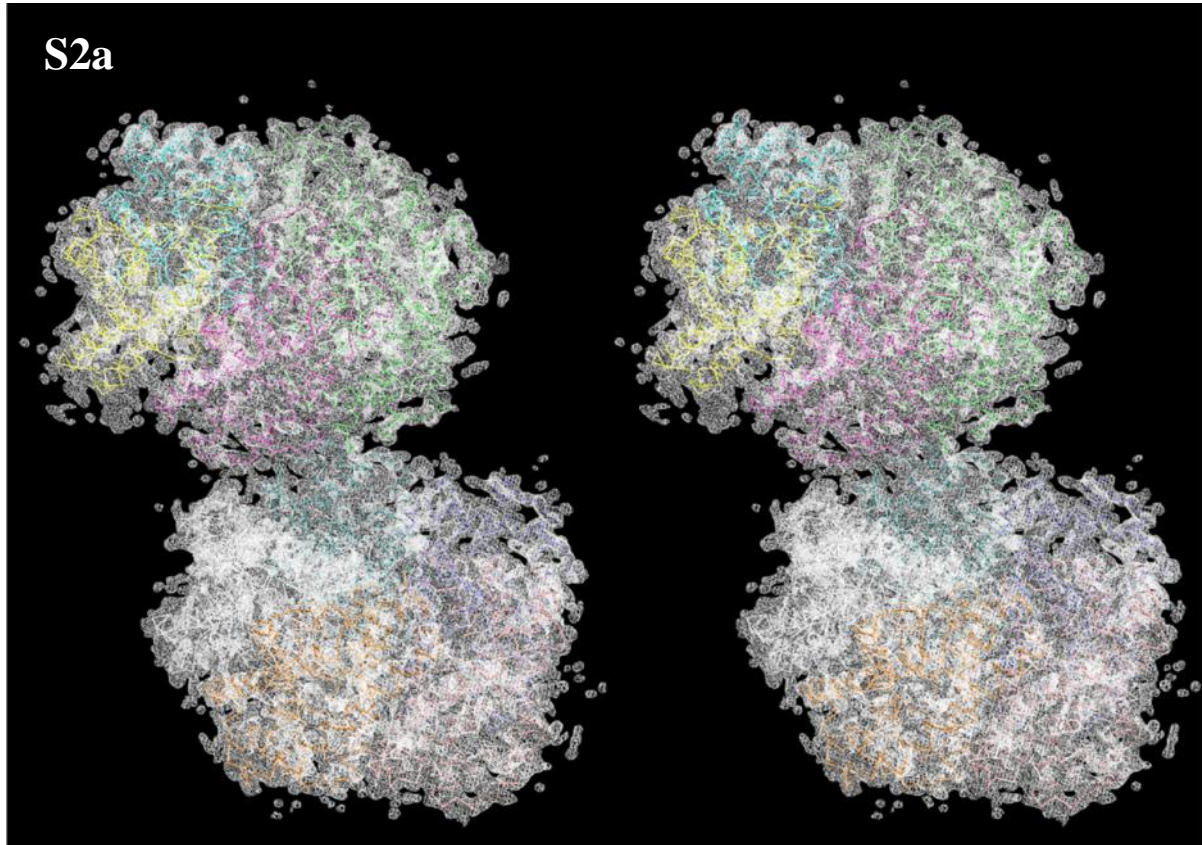


Fig. S2. Stereo view of the electron density maps of NAR1.1 at a resolution of 2.49 Å.

(a) The electron density maps are contoured at 1.2 σ . The C α traces of the two NAR1.1 pentamer in the asymmetric unit are included. (b) Representative section of the electron density in the vicinity of TMs 2a and 5a of NAR1.1. The electron density (colored white) is contoured at the 1.2 σ level and superimposed with the final refined model (green, carbon; red, oxygen; blue, nitrogen).

Table S1. Data collection, phasing and structural refinement statistics of NAR1.1 and NAR1.5.

Data set	NAR1.1	NAR1.5	Se-Met NAR1.5
Data collection			
Wavelength (Å)	0.989	0.989	0.989
Space group	P3 ₁ 21	C2	P2 ₁ 2 ₁ 2
Cell constants (Å)			
a	102.16	176.33	124.53
b	102.16	101.61	143.94
c	556.60	104.28	101.32
α, β, γ (°)	90,90,120	90,112.93,90	90,90,90
Resolution (Å)	50-2.49 (2.58-2.49)	50-3.29(3.41-3.29)	50-4.20(4.35-4.20)
Completeness (%)	99.8 (99.6)	99.4(99.8)	91.5(68.6)
Total reflections	6,197,948	825,959	3,710,425
Unique reflections	120,617	25,900	13,695
Redundancy	6.6 (6.1)	3.8(3.9)	5.4(3.6)
CC _{1/2} (High resolution shell)	0.714	0.875	0.819
R _{pim} (%)	3.4 (55.4)	3.7(43.0)	4.5(44.1)
$\langle I/\sigma(I) \rangle$	23.1 (1.4)	34.9(3.5)	21.4(1.0)
Phasing			
Number of sites			37
Resolution used (Å)			4.2
Figure of merit			0.401
Refinement			
R _{work} (%)	20.55	26.37	
R _{free} (%)	25.73	30.69	
B-factors			
Overall(Å ²)	25.34	138.84	
Rms deviations			
Bond (Å)	0.009	0.011	
Angles (°)	0.999	1.459	
Ramachandran analysis			
Most favored (%)	98.2	97.0	
Allowed (%)	1.8	3.0	
Disallowed (%)	0.0	0.0	

CHAPTER 5. GENERAL CONCLUSIONS

The emergence of drug resistant strains of bacterial pathogens has contributed to increased human casualties [1, 2]. Emerging strains of *Mycobacterium tuberculosis* and *Campylobacter jejuni* show increased resistance towards existing antibiotics [3, 4]. Bacterial drug efflux pumps are important therapeutic targets for treating pathogens because they help in pumping drug molecules out of cells, thereby preventing the accumulation of antibiotics above a certain level [5]. An important objective of this dissertation research is to elucidate the structural details of drug efflux pump components and the mechanism of their regulation, with the long term goal of identifying new therapeutic pathways. Equally important for the future is the identification and characterization of new and more efficient therapeutic targets to combat infectious pathogens.

In Chapter 2, we explore the structure and mechanism of regulation of the transcriptional regulator Rv0678 of *M. tuberculosis*. Rv0678 is a MarR-like transcriptional regulator that controls the expression of the MmpS5-MmpL5 efflux system [6]. The MmpS5-MmpL5 efflux system plays a key role in the biosynthesis and export of lipids [7]. These lipids are synthesized in the cytoplasm and are transported to the periplasm by the MmpS5-MmpL5 efflux system, where they are incorporated into the cell wall. Cell wall remodeling plays a critical role in bacterial defense against drugs and antibiotics. This makes the MmpS5-MmpL5 system an important therapeutic target. Our crystal structure of Rv0678 revealed the presence of a fortuitous ligand at the binding site, which co-purified with the protein. GCMS studies revealed the ligand to be 2-stearoylglycerol, which is a fatty acid glycerol ester. *In-vitro* studies using electrophoretic mobility shift assay (EMSA) and dye primer-based DNase I footprint assay showed that Rv0678 controls the expression levels of MmpS5-MmpL5, MmpS4-MmpL4, and MmpS2-MmpL2 transport systems. EMSA using 1-

stearoyl-rac-glycerol also showed that the ligand reduces the binding of Rv0678 with DNA. Fluorescence polarization assays using labelled 26-bp DNA enabled us to determine the binding constant and helped us identify important residues for Rv0678 DNA binding. Together with the *in-vitro* studies the crystal structure of Rv0678 suggests a possible mechanism for the binding of DNA to Rv0678.

Chapter 3 presents the crystal structure of the outer membrane channel protein, CmeC, from CmeABC tripartite drug efflux system of *Campylobacter jejuni*. CmeC was expressed in *E. coli* and was purified and crystallized. The crystal diffracted to 2.37 Å resolution. Three molecules of CmeC comes together to form a trimeric channel in a closed conformation state. The charge distribution inside the channel is strikingly electronegative, facilitating the export of antimicrobials such as fluoroquinolones. The Arginine 104 (R104) residues from different protomers were found to interact with each other and form a gate at the upper part of the α -barrel, potentially regulating the movement of substrates. The availability of the crystal structure and identification of important residues of CmeC may allow us to design novel drugs to block the function of this channel and increase the antibiotic sensitivity of pathogenic *C. jejuni*.

Chapter 4 describes the crystal structures and functional characterization of NAR1.1 and NAR1.5 of *Chlamydomonas reinhardtii*. *C. reinhardtii* is a photosynthetic green alga found predominantly in soil, fresh water bodies and sea beds that relies on photosynthesis to harvest energy of which carbon is an important raw material [8, 9]. *C. reinhardtii* possesses a carbon concentrating mechanism (CCM) to survive in carbon limiting conditions. CCM consists of many proteins found in different locations [10]. NAR1.2, also called LCIA, is a component of CCM and is believed to assist in the transport of HCO_3^- across the chloroplast envelope [11]. Chloroplast envelope membranes also possess two other proteins encoded by the *nar1* multigene family, NAR1.1 and NAR1.5, which are widely believed to transport

nitrite ions. All three proteins belong to the Formate-Nitrite family (FNT) of proteins which are widely found in bacteria but are absent in higher order plants [11]. NAR1.1 and NAR1.5 were expressed, purified and crystallized in *E. coli*. The crystals diffracted to 2.1 Å and 3.1 Å resolution respectively. The crystal structure shows a pentameric arrangement for both the proteins, with each monomer possessing a channel that spans the entire protein. Light scattering studies on proteoliposomes provided evidence that the proteins transport nitrite (NO_2^-) ions and identified important residues that could form a possible gating arrangement. Based on the crystal structure and *in vitro* studies, we have suggested a possible mechanism of transport in these channels.

References

1. Walsh C. 2003 Antibiotics: actions, origins, resistance. Washington, DC: ASM Press.
2. Jones DS, Podolsky SH, Greene JA. 2012. The burden of disease and the changing task of medicine. *N Engl J Med* 366, 2333–2338.
3. Maartens G., Wilkinson R. J. 2007. Tuberculosis. *Lancet* 370, 2030–2043
4. Ruiz-Palacios G. M. 2007. The health burden of *Campylobacter* infection and the impact of antimicrobial resistance: Playing chicken. *Clin Infect Dis* 44:701–703.
5. Coleman, K., M. Athalye, A. Clancey, M. Davison, D. J. Payne, C. R. Perry, and I. Chopra. 1994. Bacterial resistance mechanisms as therapeutic targets. *J. Antimicrob. Chemother.* 33:1091-1116
6. Ellison D. W., Miller V. L. 2006. Regulation of virulence by members of the MarR/SlyA family. *Curr. Opin. Microbiol.* 9, 153–159
7. Luo M., Fadeev E. A., Groves J. T. 2005. Mycobactin-mediated iron acquisition within macrophages. *Nat. Chem. Bio.* 1, 149 – 153

8. Spalding M. H. 2008. Microalgal carbon-dioxide-concentrating mechanisms: Chlamydomonas inorganic carbon transporters. *J Exp Bot* **59**, 1463-1473
9. Wang Y., Stessman D. J., Spalding M. H. 2015. The CO₂ concentrating mechanism and photosynthetic carbon assimilation in limiting CO₂ : how Chlamydomonas works against the gradient. *Plant J* **82**, 429-448
10. Giordano M., Beardall J., Raven J. A. 2005. CO₂ concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu Rev Plant Biol* **56**, 99-131
11. Mariscal V., Moulina P., Orselb M., Millerb A. J., Fernández E., Galvana A. 2006. Differential regulation of the Chlamydomonas Nar1 gene family by carbon and nitrogen. *Protist*.**157**, 421–433

ACKNOWLEDGEMENTS

I would like to thank my major advisor, Dr. Edward W. Yu for being a constant source of support and encouragement during my graduate studies at Iowa State University. It would have been close to impossible to finish my Ph. D. with such satisfaction if not for his motivation and guidance, which have contributed immensely towards the development and progress of my scientific career. I would like to thank him for being patient and supportive during my ups and downs in the last five years. I am fortunate to have Ed as my advisor and mentor.

I would like to thank my committee members: Dr. Robert Sam Houk, Dr. Drena L. Dobbs, Dr. Young-Jin Lee and Dr. Wenyu Huang for their valuable suggestions and advices.

I would like to thank my colleagues Tsung-Han Chou and Nitin Kumar for being extremely supportive on both academic and personal fronts. Collaborations with them have taught me to be both brave and persistent. I would like to thank my lab mates Dr. Chi-Chia Su, Dr. Feng Long, Dr. Hsien-Ting Lei, Dr. Jani Reddy Bolla, Sylvia Do and Jared Delmar. Interactions with them were extremely helpful during the dark periods of research and looking back, these moments have become sweet memories.

I am deeply indebted to my roommates over the years, Dipak Mainali, Raghu Vamsy Maligal Ganesh and Jagadeesh Kolattoor for being such fun and interactions with them ensured I never carried the burden of research home. I would also like to thank my friends in Ames, numerous to mention, who have made my stay in ISU a memorable experience.

I would like to thank Dr. Kanagalaghatta R. Rajashankar and all the staff members of NECAT in Advanced Photon Source for their help and assistance during X-ray data collection, processing and structure determination. I would also like to thank our collaborators Dr. Martin H. Spalding, Dr. Richard Martin, Dr. Katherine Osteryoung, Dr.

Georgiana Purdy and their lab members for the helpful interactions and suggestions in various projects. I want to thank Dr. William Colonna and Dr. John Strohl of the ISU fermentation facility for their generous support.

Last but not the least, I would like to thank my parents, C. Radhakrishnan Nair and P. S. Santhakumari for being a constant source of motivation and encouragement throughout my life. In many ways, the privileged and protected life I lead now is a result of the hard work they have put in, the dreams they have seen. I am indebted to my brothers Kalyan Kumar R. P. S. and Partha Sankar R. P. S. and my fiancée R. Joshya Shyamala for being my closest friends whom I could fight without ever being apologetic and still be a source of encouragement and support in times of despair and struggle; the constants of my life.

APPENDIX. CURRICULUM VITAE**Author:** Abhijith Kumar Radhakrishnan Pattathil Santha**EDUCATION**

2011-2016 PhD Chemistry
Iowa State University, Ames, IA
Advisor: Prof. Edward W. Yu

2005-2010 Integrated M. Sc. Chemistry
Pondicherry University, India
Advisor: Prof. M. M. Balakrishnarajan

RESEARCH EXPERIENCE

2011-2016 Research Assistant
Department of Chemistry
Iowa State University, Ames, IA

TEACHING EXPERIENCE

2011-2015 Teaching Assistant
Department of Chemistry
Iowa State University, Ames, IA

Supervised and graded undergraduate classes and labs of 20 students for a total of 9 semesters, which included:

- General Chemistry Recitation
- General Chemistry Laboratory
- Chemical Thermodynamics

PUBLICATIONS

1) Kumar N*, Su CC*, Yin L*, **Radhakrishnan A**, Chou TH, Dai L, Delmar JA, Zhang Q, Shin YK and Yu EW (2016). Structures and transport dynamics of a multidrug efflux pump. (manuscript submitted)

2) Kumar N*, Su CC*, Chou TH, **Radhakrishnan A**, Delmar JA, Rajashankar KR, and Yu EW (2016). Crystal structure of a hopanoid transporter of *Burkholderia*. (manuscript submitted)

3) Kumar N*, **Radhakrishnan A**, Su CC, Osteryoung KW, Yu EW (2015). Crystal structure of a conserved domain in the intermembrane space region of the plastid division protein ARC6. *Protein Sci.* doi:10.1002/pro.2825. PMID: [26452626](#)

4) Delmar JA, Chou TH, Wright CC, Licon MH, Doh JK, **Radhakrishnan A**, Kumar N, Lei HT, Bolla JR, Rajashankar KR, Su CC, Purdy GE, Yu EW. Structural Basis for the Regulation of the MmpL Transporters of *Mycobacterium tuberculosis*. 2015 Sep 22. *J Biol Chem.* pii: jbc.M115.683797. PMID: [26362239](#)

5) Chou TH, Delmar JA, Wright CC, Kumar N, **Radhakrishnan A**, Doh JK, Licon MH, Bolla JR, Lei HT, Rajashankar KR, Su CC, Purdy GE, Yu EW. Crystal structure of the *Mycobacterium tuberculosis* transcriptional regulator Rv0302. 2015 Sep 12. *Protein Sci.* doi: 10.1002/pro.2802. PMID: [26362239](#)

- 6) Bolla JR, Su CC, Delmar JA, **Radhakrishnan A**, Kumar N, Chou TH, Long F, Rajashankar KR, Yu EW. Crystal structure of the *Alcanivorax borkumensis* YdaH transporter reveals an unusual topology. *Nat Commun.* 2015 Apr 20 doi: 10.1038/ncomms7874. PMID: [25892120](#)
- 7) Su CC, Bolla JR, Kumar N, **Radhakrishnan A**, Long F, Delmar JA, Chou TH, Rajashankar KR, Shafer WM, Yu EW. Structure and function of *Neisseria gonorrhoeae* MtrF illuminates a class of antimetabolite efflux pumps. *Cell Rep.* 2015 Mar 26 doi: 10.1016/j.celrep.2015.03.003. PMID: [25818299](#)
- 8) Bolla JR, Su CC, Do SV, **Radhakrishnan A**, Kumar N, Long F, Chou TH, Delmar JA, Lei HT, Rajashankar KR, Shafer WM, Yu EW. Crystal structure of the *Neisseria gonorrhoeae* MtrD inner membrane multidrug efflux pump. *PLoS One.* 2014 Jun 5 doi:10.1371/journal.pone.0097903.PMID: [24901477](#)
- 9) Lei HT, Chou TH, Su CC, Bolla JR, Kumar N, **Radhakrishnan A**, Long F, Delmar JA, Do SV, Rajashankar KR, Shafer WM, Yu EW. Crystal structure of the open state of the *Neisseria gonorrhoeae* MtrE outer membrane channel. *PLoS One.* 2014 Jun 5 doi: 10.1371/journal.pone.0097475. PMID: [24901251](#)
- 10) **Radhakrishnan A***, Kumar N*, Su CC*, Long F, Bolla JR, Lei HT, Delmar JA, Do SV, Chou TH, Rajashankar KR, Zhang Q, Yu EW. Crystal structure of the *Campylobacter jejuni* CmeC outer membrane channel. *Protein Sci.* 2014 July 6 doi:10.1002/pro.2478. PMID: [24753291](#)
- 11) **Radhakrishnan A***, Kumar N*, Wright CC, Chou TH, Tringides ML, Bolla JR, Lei HT, Rajashankar KR, Su CC, Purdy GE, Yu EW. Crystal structure of the

transcriptional regulator Rv0678 of *Mycobacterium tuberculosis*. *J Biol Chem*. 2014 Jun 6 doi: 10.1074/jbc.M113.538959. PMID: [24737322](#)

12) **Radhakrishnan A***, Kumar N*, Wright CC, Chou TH, Lei HT, Bolla JR, Tringides ML, Rajashankar KR, Su CC, Purdy GE, Yu EW. Crystal Structure of the transcriptional regulator Rv1219c of *Mycobacterium tuberculosis*. *Protein Sci*. 2014 Jan 14 doi: 10.1002/pro.2424. PMID: [24424575](#)

* *Co-first author*

PRESENTATIONS

“Crystal structure of the transcriptional regulator Rv0678 of *Mycobacterium tuberculosis*” at the 49th Midwest Regional Meeting of the American Chemical Society - MWRM 2014, University of Missouri – Columbia, MO (Oral Presentation).