Published in final edited form as:

J Mol Biol. 2007 September 21; 372(3): 583–593.

Crystal Structure of the Transcriptional Regulator CmeR from *Campylobacter jejuni*

Ruoyu Gu^{1,†}, Chih-Chia Su^{2,†}, Feng Shi³, Ming Li¹, Gerry McDermott⁴, Qijing Zhang³, and Edward W. Yu^{1,2,*}

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

²Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA 50011, USA

³Department of Veterinary Microbiology, College of Veterinary Medicine, Iowa State University, Ames, IA 50011, USA

⁴Department of Anatomy, School of Medicine, University of California, San Francisco, CA 94143, USA

Abstract

The CmeABC multidrug efflux pump, which belongs to the resistance-nodulation-division (RND) family, recognizes and extrudes a broad range of antimicrobial agents and is essential for *Campylobacter jejuni* colonization of the animal intestinal tract by mediating the efflux of bile acids. The expression of CmeABC is controlled by the transcriptional regulator CmeR, whose open reading frame is located immediately upstream of the *cmeABC* operon. To understand the structural basis of CmeR regulation, we have determined the crystal structure of CmeR to 2.2 Å resolution, revealing a dimeric two-domain molecule with an entirely helical architecture similar to members of the TetR family of transcriptional regulators. Unlike the rest of the TetR regulators, CmeR has a large center-to-center distance (54 Å) between two N termini of the dimer, and a large flexible ligand-binding pocket in the C-terminal domain. Each monomer forms a 20 Å long tunnel-like cavity in the ligand-binding domain of CmeR and is occupied by a fortuitous ligand that is identified as glycerol. The binding of glycerol to CmeR induces a conformational state that is incompatible with target DNA. As glycerol has a chemical structure similar to that of potential ligands of CmeR, the structure obtained mimics the induced form of CmeR. These findings reveal novel structural features of a TetR family regulator, and provide new insight into the mechanisms of ligand binding and CmeR regulation.

Keywords

TetR family; *Campylobacter jejuni*; transcriptional regulator; X-ray crystallography; multidrug resistance

Introduction

Campylobacter jejuni is the leading bacterial cause of food-borne diarrhea in the USA and other developed countries.¹ It is also a significant enteric pathogen for young children in developing countries. This Gram-negative enteric organism colonizes the intestinal tracts of animals and has become increasingly resistant to antimicrobials due to the possession of

*Corresponding author E-mail address of the corresponding author: ewyu@iastate.edu.

[†]R.G. and C.S. contributed equally to this work.

multidrug efflux transporters and acquisition of various resistance mechanisms, compromising the effectiveness of antibiotic treatment. According to the genomic sequence of NCTC 11168, *C. jejuni* harbors 13 putative antibiotic efflux transporters of the ATP-binding cassette (ABC), resistance-nodulation-division (RND), multidrug and toxic compound extrusion (MATE), major facilitator (MF), and small multidrug resistance (SMR) families.^{2,3} At present, CmeABC and CmeDEF, which belong to the RND family, are the only two antibiotic efflux transporters that have been functionally characterized in *Campylobacter*.⁴⁻⁶

The CmeABC efflux system is the main efflux pump in *C. jejuni* and consists of three members, including an outer membrane channel (CmeC), an inner membrane drug transporter (CmeB), and a periplasmic membrane fusion protein (CmeA). These three proteins are encoded by a three-gene operon (*cmeABC*) and form an efflux system that extrudes a variety of toxic compounds directly out of *C. jejuni*.⁵ The substrates extruded by CmeABC include commonly used antibiotics (e.g. fluoroquinolones, macrolides, ampicillin, tetracycline, chloramphenicol, cefotaxime, rifampin), metal ions (e.g. Co^{2+} and Cu^{2+}), and lipophilic compounds (e.g. SDS and various bile salts). Thus, CmeABC contributes significantly to the intrinsic and acquired resistance of *Campylobacter* to structurally diverse antimicrobials.⁵⁻⁸ In addition, this efflux system is essential for *Campylobacter* colonization in the animal intestinal tract by conferring resistance to the bile acids that are normally present in the animal intestinal tract and have bactericidal effect.⁹

The expression of *cmeABC* is controlled by the transcriptional regulator CmeR.¹⁰ The *cmeR* gene is located immediately upstream of the *cmeABC* operon and encodes a 210 amino acid residue protein that shares N-terminal sequence and structural similarities with members of the TetR family of transcriptional repressors.^{11,12} Like other members of the TetR family, the N-terminal domain of CmeR contains a predicted DNA-binding helix-turn-helix (HTH) motif, while its C-terminal region has unique sequences and is expected to be involved in the binding of inducing ligands.^{10,11} *cmeR* is transcribed in the same direction as *cmeABC*, and the intergenic region between *cmeR* and *cmeA* contains the 16 bp inverted repeat (IR) operator site for *cmeABC*. As a transcriptional regulator, CmeR binds directly to the IR operator and represses the transcription of *cmeABC*.¹⁰ This regulating process is similar to those of the other TetR family members, such as AcrR of *Escherichia coli*,¹³ MexR of *Pseudomonas aeruginosa*,¹⁴ MtrR of *Neisseria gonorrhoeae*,¹⁵ and QacR of *Staphylococcus aureus*,¹⁶ in which they bind specifically to the promoter sequences of *acrAB*, *mexAB*, *mtrCDE*, and *qacA*, respectively, thus inhibiting the expression of the corresponding efflux pump(s). Deletion of *cmeR* or mutations in the IR operator releases the repression, resulting in the over-expression of CmeABC, which, in turn, leads to the enhanced resistance to multiple antibiotics.¹⁰

In addition, bile compounds, including both conjugated (e.g. taurodeoxycholate) and non-conjugated (e.g. cholate), induce the expression of *cmeABC* by inhibiting the binding of CmeR to the promoter of *cmeABC*, suggesting that bile compounds are inducing ligands of CmeR.¹⁷ It is possible that CmeR can be induced by other unidentified ligands. How inducing ligands bind to CmeR and modulate the expression of CmeR-controlled genes is not known. On the basis of the predicted structural features, we hypothesize that binding of inducing ligands to the C-terminal domain of CmeR triggers conformational change in the N-terminal DNA-binding region. This change in conformation results in the release of CmeR from its operator DNA, and thus allows transcription from its cognate promoter. As an initial step to examine this hypothesis and elucidate the mechanisms that CmeR uses to regulate gene expression, we present here the crystal structure at 2.2 Å resolution of the CmeR regulator.

Results

Overall structure of CmeR

We used the multiple-wavelength anomalous dispersion method to solve the selenomethionyl-substituted (SeMet) CmeR crystal structure from *C. jejuni*. Its native crystal structure was then determined to 2.2 Å resolution (Table 1 and Figure 1(a)), revealing that only one CmeR molecule exists in the asymmetric unit. The dimeric arrangement of the protein was found by applying the crystallographic symmetry operators.

The dimeric structure of CmeR, indicating an all-helical protein, is shown in Figure 1(b). As a member of the TetR family of transcriptional regulators, CmeR consists of two functional motifs; an N-terminal DNA-binding domain and a C-terminal ligand-binding domain.^{11,12}

The crystal structure revealed that each subunit of CmeR is composed of ten α helices ($\alpha 1$ – $\alpha 10$ and $\alpha 1'$ – $\alpha 10'$, respectively) and indeed can be divided into two domains. The smaller N-terminal domain shares considerably high levels of sequence and structural similarities with the other TetR family members. For example, residues 12–65 possess 23% amino acid identity with and 43% similarity to that of TetR.¹⁸ This N-terminal region also shows 50% identity with and 71% homology to that of the QacR repressor.¹⁹ Among the TetR family, including TetR,^{18,20} QacR,^{19,21} CprB,²² and EthR,^{23,24} the N-terminal domains of these transcriptional regulators are formed by three α -helix bundles. The structure of CmeR, however, revealed that the third short α -helix, presumably formed by residues 47–53 is missing. Instead, these residues form a random coil with noticeably high *B*-factors, suggesting a mobile nature of this coil. To facilitate the comparison with the structures of other TetR members, helices of CmeR are numbered from the N terminus as $\alpha 1$ (7-29), $\alpha 2$ (36-43), $\alpha 4$ (57-81), $\alpha 5$ (88-104), $\alpha 6$ (106-118), $\alpha 7$ (a(125-136) and b(138-148)), $\alpha 8$ (152-170), $\alpha 9$ (172-180), and $\alpha 10$ (187-203), in which helix $\alpha 3$ has been skipped. In this arrangement, the larger C-terminal domain comprises seven α helices ($\alpha 4$ – $\alpha 10$) and is involved in the dimerization of the repressor. According to the ligand-bound structures of TetR,¹⁸ QacR,¹⁹ and EthR,²⁴ the C-terminal region also forms the drug-binding domain. Like TetR, QacR and EthR, the crystal structure of CmeR suggests that helices 4–9 form the ligand-binding domain of the regulator.

N-terminal domain

The overall structure of the N-terminal DNA-binding domain of CmeR is quite similar to those of the TetR family members. A superposition of C α atoms, between residues 14 and 44, of CmeR with their corresponding residues in QacR gives an overall rmsd of 0.8 Å. The significant sequence conservation of CmeR and QacR spans the entire N-terminal region and extends into the N-terminal end of helix $\alpha 4$. Despite these structural and sequence similarities, the structure of the DNA-binding domain of CmeR presents some noticeable differences from the rest of the TetR family members. One difference comes from helix $\alpha 1$. This first helix consists of 23 residues, which is relatively long compared with all structurally known members of the TetR family. For example, helices $\alpha 1$ in QacR,¹⁹ TetR,¹⁸ and EthR²³ are composed of 16, 13, and 17 residues, respectively.

Perhaps, the most striking difference between structures of CmeR and other TetR members in the DNA-binding domain is the lack of the third N-terminal helix in CmeR. Until now, all known TetR family of regulators possess a helix-turn-helix (HTH) DNA-binding motif, which is formed by helices $\alpha 2$ and $\alpha 3$. Helix $\alpha 3$ is named the “recognition helix,” as it has a key role in binding the target DNA.¹² In the case of CmeR, however, only two N-terminal helices, $\alpha 1$ and $\alpha 2$, are found. According to the sequence alignment, residues 47–53 are supposed to form the recognition helix $\alpha 3$. The crystal structure, however, shows that this region forms a random coil. We predicted the secondary structure of CmeR using the programs GOR V^{†25} and

PROF[‡].²⁶ Both predictions describe the overall crystal structure of CmeR quite accurately, and they both exclude the presence of $\alpha 3$.

One of the unique features of the CmeR structure is its large center-to-center distance between two N termini of the dimer. As CmeR does not have helix $\alpha 3$, we measured this center-to-center distance according to the separation between C $^{\alpha}$ atoms of Tyr51 and Tyr51', which was measured 54 Å. The corresponding distances are 39 Å and 35 Å in the apo forms of QacR,¹⁹ and TetR.²⁷ These center-to-center distances increase upon ligand binding. For the ligand-bound dimers of QacR,¹⁹ TetR,¹⁸ EthR,²⁴ and YfiR,²⁸ these distances become 41 Å, 38 Å, 52 Å and 54 Å, respectively. Thus, the relatively large center-to-center distance observed with CmeR suggested that CmeR was liganded.

C-terminal domain

The C-terminal domain of CmeR consists of eight helices ($\alpha 4$ – $\alpha 10$). Except helices 6 and 9, these helices form an anti-parallel, five-helix bundle. In view of the crystal structure, helices 6, 8, 9 and 10 are involved in the formation of the dimer. Dimerization occurs mainly by couplings between pairs of helices ($\alpha 6$ and $\alpha 9'$, $\alpha 8$ and $\alpha 10'$, and their identical counter pairs). A surface area of 1950 Å² per monomer (probe radius of 1.4 Å) is buried in the contact region of the dimer. The interaction surface is mostly hydrophobic in character. Within 3.5 Å, the close contact pairs in the helices involve I130 and I180', K160 and H193', I114 and Y120', Y153 and I205', L161 and F196', Y116 and Y116', and T167 and P172'. Additional contact interfaces are provided by the loop connecting helices 6 and 7, and the loop region right after helix 10. At the flexible loop between helices $\alpha 6$ and $\alpha 7$, Y120 makes contact with I114' in helix $\alpha 6'$, while residue I205 close to the end of the C-terminus forms hydrophobic contacts with M154' and L202' from the other subunit. The dimer interface is further strengthened by two cross-interface hydrogen bonds formed between backbone atoms (between L202 and I205', and between K113 and V119'), and a water-mediated hydrogen bond (between Y120 and A110'). The overall structure of the C-terminal domain of CmeR is closest to that of QacR among the TetR family members. Superposition of the C-terminal domains of CmeR and QacR suggests that helices 4–6, 7a-b, 8, and 10 of CmeR correspond to helices 4–6, 7, 8, and 9, respectively, of QacR. In QacR, helix $\alpha 8$ transits directly to the last helix through a nearly 180° turn.¹⁹ Similar anti-parallel arrangements of the last two C-terminal helices are found in other TetR family members, such as EthR,²³ YfiR,²⁸ and CprB,²² and these last two helices are contribute mainly to dimerization. The C-terminal domain of CmeR is distinct, in that helix, $\alpha 9$, which is between the two anti-parallel helices $\alpha 8$ and $\alpha 10$, deviates from the direction of $\alpha 8$ by 40°. Thus, helix $\alpha 9$ bends toward the next subunit of the dimer, interacting with $\alpha 6'$ and $\alpha 7'a$ to secure interaction between the dimer. The C-terminal domain of CmeR forms a large tunnel-like cavity (Figure 2). This tunnel, formed predominantly by helices 4–9, opens horizontally from the front to the back of each protomer. The length of this tunnel is about 20 Å. Helices $\alpha 7$ and $\alpha 8$ from one subunit, and $\alpha 9'$ from the other subunit of the regulator make the entrance of the tunnel. Residues I130, Q134, F137 and Y139 of $\alpha 7b$; E159, V163 and T167 of $\alpha 8$; and P172', Y173', L176' and I180' of $\alpha 9'$ are involved in the formation of this entrance.

Surrounding the inner wall of this tunnel are I68, C69, F99, F103, A108, F137 and S138, in which many of these residues are hydrophobic in nature. Helices 4–6, with the side-chains of I68, H72, I102, E107 and F111, contribute to form the end of this hydrophobic tunnel. Similar hydrophobic tunnels have been found in the EthR^{23,24} and YfiR²⁸ repressors. In the case of EthR, the tunnel opens vertically to the bottom of the molecule.^{23,24} For YfiR, however, the

[‡]<http://gor.bb.iastate.edu>

[§]<http://cubic.bioc.columbia.edu/20services/NORSp/submit.html>

long tunnel opens on one end at the subunit interface, and this end of the tunnel is nearly blocked by the second subunit.²⁸

CmeR was liganded

The initial solvent-flattened multiple-wavelength anomalous dispersion map showed a positive density, presumably from an unidentified ligand that was purified and crystallized with the protein, inside the hydrophobic tunnel of CmeR. We used the program “putative active sites with spheres” (PASS)²⁹ to search for potential ligand-binding sites in the CmeR structure, in which the top two predicted binding sites were located inside the hydrophobic tunnel. One of these two predicted sites overlaps with the unidentified positive electron density. The simulated annealing omit map shows that the ligand density appears to come from a small molecule.

Gas chromatography coupled with mass spectrometry (GC-MS) suggests that the bound ligand is glycerol (1,2,3-propanetriol), as it was detected as the major component present in the extraction solvent (see Supplementary Data Figure S1). We used solutions containing glycerol to purify and crystallize the protein and it was not surprising that the identified ligand is glycerol, although the finding was unexpected. The chemical structure of glycerol is compatible with the positive density in the simulated annealing omit map, and it fits unambiguously into the CmeR structure (Figure 3).

Glycerol has not been proved to be a ligand of CmeR or an inducer of the *cmeABC* operon, the structure obtained probably mimics the ligand-bound form of CmeR. For the glycerol binding, the hydroxyl atom O3 forms hydrogen bonds with S138 and the backbone N atom of Y139 at distances of 3.1 Å and 2.7 Å, respectively. The repressor protein further anchors the bound glycerol through two water-mediated hydrogen bonds, between T167 and hydroxyl atom O2 (through OW1) and between S138 and hydroxyl atom O1 (through OW2) of the bound glycerol. The SG of C166 and NZ of K170 are less than 4 Å away from O2 and O1, respectively, interacting with the bound glycerol and securing the binding.

Predicting the structure of the DNA-bound form of CmeR

Although we expect that the DNA-binding mode of CmeR is similar to that of TetR, the overall crystal structure and sequence alignment suggest that the CmeR protein is more similar to the QacR repressor. Thus, a speculative model of DNA-bound CmeR was generated by the alignment of its individual domains with those of the DNA-bound QacR (Figure 4). This model reveals an extensive movement in CmeR that might allow a shift from a ligand-bound form to a DNA-bound form of the repressor, although we cannot exclude the fact that the target DNA itself may also undergo conformational changes allowing binding to CmeR as seen for TetR,²⁰ QacR²¹ and BmrR.³⁰ On the basis of this DNA-bound model, it is speculated that helix $\alpha 6$ moves toward $\alpha 9$ during DNA binding, resulting in a decrease in volume of the ligand-binding site.

Docking of ligands into the hydrophobic tunnel

To elucidate different binding modes of CmeR to a variety of ligands, we used the program MEDock³¹ to identify potential binding pockets for two bile acids, taurocholate and cholate. We first predicted a glycerol-binding site in CmeR (Figure 5(a)). This predicted site overlaps with the glycerol-binding site identified from the crystal structure, suggesting MEDock is sufficiently precise for identification of potential protein-binding pockets. When MEDock was used to search for a taurocholate-binding site in CmeR, it was found that the taurocholate molecule was bound inside the tunnel, spanning the two PASS predicted ligand-binding sites. Taurocholate is a 19 Å long ligand, and spans almost the entire length of the ligand-binding tunnel (Figure 5(b)). For the predicted cholate binding in CmeR, cholate binds inside the tunnel, very similar to taurocholate binding (Figure 5(c)).

Discussion

The structural similarity of the N-terminal domains of members of the TetR family suggests a similar mode of interaction with target DNAs. CmeR represses the transcription of *cmeABC* by binding directly to the inverting 16 bp IR sequence in the promoter region of the efflux operon. This IR sequence is similar in length to that of the 15 bp *tetO* bound by TetR, but is different from the long 28 bp IR1 recognized by QacR. TetR binds as a single dimer to the *tetO* operator;²⁰ however, two dimers of QacR interact with one IR1.²¹ On the basis of the IR sequence, we reasoned that CmeR might bind its operator as a dimer, similar to the TetR DNA binding.

The separation between two successive major grooves of a 16 bp double helix should be less than 40 Å (the distance between two consecutive major grooves of *B*-DNA is 34 Å). The ligand-bound structure of CmeR indicates that the two DNA recognition regions of the dimer are separated by 54 Å, which is incompatible with the binding of the regulator to its 16 bp operator. Thus, it is possible that a drastic change in conformation of the DNA-binding domain of CmeR might take place during the process of binding target DNA. This change should be greater compared with that of QacR and TetR. In the case of QacR, the change in conformation is accompanied by an increase in the center-to-center distance from 37 Å of the DNA-bound form to 48 Å of the ligand-bound form.²² For TetR, however, this change is less obvious. The center-to-center distance of the TetR dimer shifts only from 35 Å in the DNA-bound structure to 38 Å in the ligand-bound structure.²⁰ On the basis of the DNA-bound model of CmeR, it is speculated that this center-to-center distance may decrease to 36 Å upon DNA binding (Figure 4). This change may trigger a coupling movement of helices $\alpha 6$ and $\alpha 9$, resulting in a decrease in size of the ligand-binding tunnel, which in turn hinders the inducer ligand to enter the ligand-binding site. The consequence is that the inducer ligand has to overcome this steric hindrance in order to bind CmeR. Thus, it is likely that CmeR induction may be governed by steric repulsion that takes place during inducer binding. The crystal structures of both ligand-free and DNA-bound CmeR would be necessary to infer the mechanisms of CmeR induction, and to confirm our speculative models based on the glycerol-bound CmeR structure.

The lack of the recognition helix, $\alpha 3$, in the DNA-binding domain of the CmeR structure is unique among members of the TetR family. Secondary structure prediction using programs GOR V²⁵ and PROF²⁶ also suggests that this segment (residues 47–53) is likely to form a random coil. An α -helix is stabilized mainly by a favorable enthalpic contribution from the formation of the backbone hydrogen bonds and van der Waals interactions.^{32,33} However, a random coil is mostly favored by conformational entropy that degenerates different conformational states of the coil.³⁴ The entropic cost of fixing the backbone dihedral angles in forming an α -helical structure is within 2 kcal/mol at 25 °C.^{33,35,36} We reasoned that in the case of CmeR, the energy difference between α -helical and randomly coil states of the recognition helix are very close, in the range of 1–2 kcal/mol. If this is the case, it is possible that the segment forming $\alpha 3$ is more favorable for formation of the flexible coil state in the absence of target DNA. This segment will transform into an α -helical conformation upon DNA binding due to the release of energy from the formation of hydrogen bonds and contact interactions between the repressor and target DNA. On the basis of a helix propensity scale,³⁷ we estimated the amount of energy involved in helix formation of the last five residues in the recognition segment. This estimated energy in CmeR is 3.3 kcal/mol, which is about 1 kcal/mol greater when compared with those of TetR (2.5 kcal/mol), QacR (2.5 kcal/mol), EthR (2.7 kcal/mol), and CprB (1.9 kcal/mol). It seems that in CmeR this segment has less tendency to form α -helix. However, this extra 1 kcal/mol is not excessive and can be compensated easily by the release of energy during repressor–operator interaction.

One striking feature of the CmeR structure is the large ligand-binding tunnel in each monomer. This tunnel, nearly 20 Å in length, is surrounded by mostly hydrophobic residues of helices 4–9, and occupies a volume of about 1000 Å³. Each hydrophobic tunnel spans horizontally across the C-terminal domain, and can be seen through from the front to the back of the dimer without obstruction. This unique feature, in the TetR family of regulators, highlights the flexibility of the CmeR regulator. Unexpectedly, the crystal structure of CmeR revealed the presence of a glycerol molecule inside this large ligand-binding tunnel. Glycerol binds identically in each subunit, as indicated by the crystallographic 2-fold symmetry of the CmeR dimer (Figure 1(b)). This ligand-binding mode is different from that of QacR, in which one dimer of QacR binds one drug,¹⁹ but similar to that of TetR, which interacts with tetracycline in a 1:1 monomer/drug molar ratio.¹⁸ The volume of the ligand-binding tunnel of CmeR is large enough to accommodate a few of the ligand molecules. Additional water molecules fill the portion of the large tunnel that is unoccupied by ligand. When PASS²⁹ was used to search for potential ligand-binding sites in the CmeR structure, the top two predicted sites, which are 8.4 Å apart, were found inside this hydrophobic tunnel. One of these predicted sites corresponds to the glycerol-binding site. The second predicted site, however, is still empty. Thus, CmeR might be able to accommodate a much bigger ligand that spans across these two predicted binding sites. There is also a good chance that CmeR, like QacR, could bind two drug molecules at a time.³⁸ In any case, the flexibility of the large ligand-binding tunnel suggests that CmeR is a multiple ligand-binding protein.

Although the ligand-binding tunnel is mainly hydrophobic in nature, the electrostatic surface diagram (Figure 2(a)), somewhat surprising, displays a patch of positive surface potential inside the tunnel. This positive potential indicates that CmeR may be more favorable to bind neutral and negatively charged ligands. In fact, many of the CmeR ligands, such as bile acids, are negative in charge. To examine how CmeR binds a variety of ligands, we used MEdock to predict the binding sites of cholate and taurocholate. The docking study suggested that the bound cholate and taurocholate, respectively, occupied both of the PASS predicted sites in the hydrophobic tunnel. These binding modes are similar to that of the dequalinium binding in QacR, in which the bound dequalinium took both the rhodamine 6G and ethidium-binding sites.¹⁹

CmeR represses the expression of the CmeABC efflux pump that extrudes various bile salts, such as cholate, deoxycholate and taurocholate. It also recognizes commonly used antibiotics, including fluoroquinolones, macrolides, tetracycline and rifampin. Thus, these compounds are the CmeR ligands. It has not been shown that glycerol is a natural ligand of CmeR. The biological effect of CmeR binding by glycerol remains to be determined. Recently, it has been shown using DNA microarray that CmeR may function as a pleiotropic regulator modulating the expression of multiple membrane transporters, including two C₄-dicarboxylate transporters DcuA and DcuB (Q.Z., unpublished results). This finding suggests that C₄-dicarboxylates, such as malate, fumarate, succinate and aspartate, might be ligands for CmeR. The chemical structures of glycerol (which is bound by CmeR) and C₄-dicarboxylates are quite similar, suggesting that CmeR may recognize and respond to C₄-dicarboxylate compounds.

Materials and Methods

Purification, crystallization and data collection

Recombinant CmeR, containing a His₆ tag at the N terminus, was produced in *Escherichia coli* using the pQE30 vector. The cloning, expression, purification and crystallization procedures have been described.^{5,9,10,39} Diffraction data sets of both the native and SeMet-CmeR crystals were taken at the Advanced Light Source (beam-line 8.2.2) at cryogenic temperature (100 K) using an ADSC Quantum 315 CCD-based detector.

Structural determination and refinement

Diffraction data sets were processed with DENZO and scaled with SCALEPACK.⁴⁰ Both native and SeMet crystals took the space group of $P2_12_12$, with the unit cell parameters given in Table 1. Initial phase calculation was carried out at 2.8 Å resolution using the program BnP⁴¹ after finding and refinement of all three selenium sites. The electron density map obtained was applied to density modification (DM) using the program RESOLVE.⁴² The auto-interpretation routine program in RESOLVE led to an initial model containing 78% amino acid residues, 50% of which contained side-chains. The remaining part of the model was constructed manually using the program O.⁴³ The model, comprising residues 6–207, was then refined against native data at 2.2 Å using the programs CNS⁴⁴ and REFMAC5.^{45,46} Solvent atoms were initially built using the program ARP/warp^{45,47} and later added or removed by manual inspection. The final R -factor and R_{free} (calculated with 5% of the reflections omitted from the refinement) were 21.9% and 24.3%, respectively.

Modeling of DNA-bound form of CmeR

The model of the DNA-bound form of CmeR was generated using O.⁴³ In brief, the N and C-terminal domains of CmeR were separately aligned with the corresponding domains of the DNA-bound QacR (1JT021). The resulting model was then idealized using REFMAC5.^{45,46} The CmeR dimer was generated by applying symmetry operators obtained from the ligand-bound CmeR crystal structure. The recognition helix $\alpha 3$ was also placed accordingly. The final center-to-center distance of the DNA-bound form of CmeR is 36 Å.

Prediction of ligand-binding sites

The MEdock web server³¹ was used for prediction of the taurocholate (PDB tch) and cholate (PDB chd) bindings in CmeR. A global search strategy that exploits the maximum entropy property of the Gaussian distribution was employed. For the docking protocol, the maximum generation in each run was set to 1000. A grid of 0.375 Å spacing was used for the calculation. Five separate docking calculations were performed for each ligand. Each calculation was performed with a population size of 50, and a probability of 0.05 to invoke local search.

Identification of fortuitous ligand

We used GC-MS to identify the nature of the bound ligand in crystals of CmeR. The CmeR crystals were washed extensively with the crystallization buffer and transferred into deionized water. The mixture was incubated at 100 °C for 5 min, and subsequently methanol was added into the mixture to a final concentration of 80% (v/v) to denature the protein and allow for the extraction of ligand. The GC-MS results suggested that the bound ligand is glycerol (1,2,3-propanetriol) (see Supplementary Data).

Protein Data Bank accession code

Coordinates and structural factors for the structure of CmeR have been deposited with the RCSB Protein Data Bank with accession code 2QCO.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Initial crystal screens were performed at the Stanford Synchrotron Radiation Laboratory (SSRL, beamline BL9-1) and Advanced Photon Source (APS, beamline 22ID). The complete X-ray data sets of both native and SeMet CmeR were collected at the Advanced Light Source (ALS, beamline 8.2.2). This work was supported by NIH grants DK063008 (to Q.Z.) and GM074027 (to E.W.Y.).

References

1. Friedman, CR.; Neimann, J.; Wegener, HC.; Tauxe, RV. Epidemiology of *Campylobacter jejuni* infections in the United States and other industrialized nations. In: Nachamkin, I.; Blaser, MJ., editors. *Campylobacter*. ASM Press; Washington, DC: 2000. p. 121-138.
2. Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, et al. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 2000;403:665–668. [PubMed: 10688204]
3. Lin, J.; Akiba, M.; Zhang, Q. Multidrug efflux systems in *Campylobacter*. In: Ketley, JM.; Konkel, ME., editors. *Campylobacter: Molecular and Cell Biology*. Horizon Bioscience; Norfolk, UK: 2005. p. 205-218.
4. Akiba M, Lin J, Barton YW, Zhang QJ. Interaction of CmeABC and CmeDEF in conferring antimicrobial resistance and maintaining cell viability in *Campylobacter jejuni*. *J Antimicrob Chemother* 2006;57:52–60. [PubMed: 16303882]
5. Lin J, Michel LO, Zhang Q. CmeABC functions as a multidrug efflux system in *Campylobacter jejuni*. *Antimicrob Agents Chemother* 2002;46:2124–2131. [PubMed: 12069964]
6. Pumbwe L, Piddock LJ. Identification and molecular characterization of CmeB, a *Campylobacter jejuni* multidrug efflux pump. *FEMS Microbiol Letters* 2002;206:185–189.
7. Cagliero C, Mouline C, Payot S, Cloeckaert A. Involvement of the CmeABC efflux pump in the macrolide resistance of *Campylobacter coli*. *J Antimicrob Chemother* 2005;56:948–950. [PubMed: 16157618]
8. Luo N, Sahin O, Lin J, Michel LO, Zhang Q. *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* 2003;47:390–394. [PubMed: 12499221]
9. Lin J, Sahin O, Michel LO, Zhang Q. Critical role of multidrug efflux pump CmeABC in bile resistance and *in vivo* colonization of *Campylobacter jejuni*. *Infect Immun* 2003;71:4250–4259. [PubMed: 12874300]
10. Lin J, Akiba M, Sahin O, Zhang Q. CmeR functions as a transcriptional repressor for the multidrug efflux pump CmeABC in *Campylobacter jejuni*. *Antimicrob Agents Chemother* 2005;49:1067–1075. [PubMed: 15728904]
11. Grkovic S, Brown MH, Skurray RA. Regulation of bacterial drug export systems. *Microbiol Mol Biol Rev* 2002;66:671–701. [PubMed: 12456787]
12. Ramos JL, Martinez-Bueno M, Molina-Henares AJ, Teran W, Watanabe K, Zhang XD, et al. The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* 2005;69:326–356. [PubMed: 15944459]
13. Ma D, Alberti M, Lynch C, Nikaido H, Hearst JE. The local repressor AcrR plays a moderating role in the regulation of *acrAB* genes of *Escherichia coli* by global stress signals. *Mol Microbiol* 1996;19:101–112. [PubMed: 8821940]
14. Evans K, Adewoye L, Poole K. MexR repressor of the *mexAB-oprM* multidrug efflux operon of *Pseudomonas aeruginosa*: identification of MexR binding sites in the *mexA-mexR* intergenic region. *J Bacteriol* 2001;183:807–812. [PubMed: 11208776]
15. Hagman KE, Shafer WM. Transcriptional control of the *mtr* efflux system of *Neisseria gonorrhoeae*. *J Bacteriol* 1995;177:4162–4165. [PubMed: 7608095]
16. Grkovic S, Brown MH, Roberts NJ, Paulsen IT, Skurray RA. QacR is a repressor protein that regulates expression of the *Staphylococcus aureus* multidrug efflux pump QacA. *J Biol Chem* 1998;273:18665–18673. [PubMed: 9660841]
17. Lin J, Cagliero C, Guo B, Barton YW, Maurel MC, Payot S, Zhang Q. Bile salts modulate expression of the CmeABC multidrug efflux pump in *Campylobacter jejuni*. *J Bacteriol* 2005;187:7417–7424. [PubMed: 16237025]
18. Hinrichs W, Kisker C, Duvel M, Muller A, Tovar K, Hillen W, Saenger W. Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. *Science* 1994;264:418–420. [PubMed: 8153629]

19. Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA, Brennan RG. Structural mechanisms of QacR induction and multidrug recognition. *Science* 2001;294:2158–2163. [PubMed: 11739955]
20. Orth P, Schnappinger D, Hillen W, Saenger W, Hinrichs W. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nature Struct Biol* 2000;7:215–219. [PubMed: 10700280]
21. Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA, Brennan RG. Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J* 2002;21:1210–1218. [PubMed: 11867549]
22. Natsume R, Ohnishi Y, Senda T, Horinouchi S. Crystal structure of a γ -butyrolactone auto-regulator receptor protein in *Streptomyces coelicolor* A3 (2). *J Mol Biol* 2003;336:409–419. [PubMed: 14757054]
23. Dover LG, Corsino PE, Daniels IR, Cocklin SL, Tatituri V, Besra GS, Futterer K. Crystal structure of the TetR/CamR family repressor *Mycobacterium tuberculosis* EthR implicated in ethionamide resistance. *J Mol Biol* 2004;340:1095–1105. [PubMed: 15236969]
24. Frenois F, Engohang-Ndong J, Locht C, Baulard AR, Villeret V. Structure of EthR in a ligand bound conformation reveals therapeutic perspectives against tuberculosis. *Mol Cell* 2004;16:301–307. [PubMed: 15494316]
25. Kloczkowski A, Ting K-L, Jernigan RL, Garnier J. Combining the GOR V algorithm with evolutionary information for protein secondary structure prediction from amino acid sequence. *Proteins: Struct Funct Genet* 2002;49:154–166. [PubMed: 12210997]
26. Rost B. Predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol* 1996;266:525–539. [PubMed: 8743704]
27. Orth P, Cordes F, Schnappinger D, Hillen W, Saenger W, Hinrichs W. Conformational changes of the Tet repressor induced by tetracycline trapping. *J Mol Biol* 1998;279:439–447. [PubMed: 9642048]
28. Rajan SS, Yang X, Shuvalova L, Collart F, Anderson WF. Crystal structure of YfiR, an unusual TetR/CamR-type putative transcriptional regulator from *Bacillus subtilis*. *Proteins: Struct Funct Genet* 2006;65:255–257. [PubMed: 16862575]
29. Brady GP Jr, Stouten PFW. Fast prediction and visualization of protein binding pockets with PASS. *J Comput Aided Mol Des* 2000;14:383–401. [PubMed: 10815774]
30. Heldwein EEZ, Brennan RG. Crystal structure of the transcriptional activator BmrR bound to DNA and a drug. *Nature* 2001;409:378–382. [PubMed: 11201751]
31. Chang DT, Oyang Y-J, Lin J-H. MEdock: a web server for efficient prediction of ligand binding sites based on a novel optimization algorithm. *Nucl Acids Res* 2005;33:W233–W238. [PubMed: 15991337]
32. Scholtz JM, Marqusee S, Baldwin RL, York EJ, Stewart JM, Santoro M, Bolen DW. Calorimetric determination of the entropy change for the σ -helix to coil transition of an alanine peptide in water. *Proc Natl Acad Sci USA* 1991;88:2854–2858. [PubMed: 2011594]
33. Yang A-S, Honig B. Free energy determinants of secondary structure formation. I. α -Helices. *J Mol Biol* 1995;252:351–365. [PubMed: 7563056]
34. Aurora R, Creamer TP, Srinivasan R, Rose GD. Local interactions in protein folding: lessons from the α -helix. *J Biol Chem* 1997;272:1413–1416. [PubMed: 9019474]
35. Nemethy G, Leach SJ, Scheraga HA. The influence of amino acid side-chains on the free energy of helix-coil transitions. *J Phys Chem* 1966;70:998–1004.
36. Wang J, Purisma EO. Analysis of thermodynamic determinants in helix propensities of nonpolar amino acids through a novel free energy calculation. *J Am Chem Soc* 1996;118:995–1001.
37. Pace CN, Scholtz JM. A helix propensity scale based on experimental studies of peptides and proteins. *Biophys J* 1998;75:422–427.
38. Schumacher MA, Miller MC, Brennan RG. Structural mechanism of the simultaneous binding of two drugs to a multidrug-binding protein. *EMBO J* 2004;23:2923–2930. [PubMed: 15257299]
39. Su C-C, Shi F, Gu R, Li M, McDermott G, Yu EW, Zhang Q. Preliminary structural studies of the transcriptional regulator CmeR from *Campylobacter jejuni*. *Acta Crystallogr Sect F* 2007;63:34–36.

40. Otwinowski Z, Minor M. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997;276:307–326.
41. Weeks CM, Shah N, Green ML, Miller R, Furey H. Automated web- and grid-based protein phasing with *BnP*. *Acta Crystallog sect A* 2005;61:C152.
42. Terwilliger TC. Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallog sect D* 2001;57:1755–1762.
43. Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallog sect A* 1991;47:110–120.
44. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, et al. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog sect D* 1998;54:905–921.
45. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallog sect D* 1994;50:764–767.
46. Murshudov A, Vagin A, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog sect D* 1997;53:240–255.
47. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. *Nature Struct Biol* 1999;6:458–463. [PubMed: 10331874]

Figure 1.

Stereo view of the experimental density map and ribbon diagram of CmeR. (a) Representative section of electron density at the subunit interface. The solvent-flattened electron density (50–2.3 Å) is contoured at the 1σ level and superimposed with the final refined model (yellow, carbon; red, oxygen; blue nitrogen; green, sulfur). (b) Ribbon diagram of the ligand-bound CmeR homodimer generated by crystallographic symmetry. The Figure was prepared using PyMOL [<http://www.pymol.sourceforge.net>].

Figure 2.

Views of the tunnel-like cavity in the ligand-binding domain of CmeR. (a) Electrostatic surface potential of one subunit of CmeR. This view shows the long tunnel spanning through the C-terminal domain of CmeR. Blue ($+15 k_B T$) and red ($-15 k_B T$) indicate the positively and negatively charged areas of the protein, respectively, (b) View of the hydrophobic cavity with residues forming the tunnel. The Figure was prepared using PyMOL [<http://www.pymol.sourceforge.net>].

Figure 3.

Simulated annealing omit map of the glycerol binding pocket. A stereo view of the composed electron density omit map (contoured at the 1.5σ level) calculated by excluding glycerol from the model. Carbon atoms are colored grey for bound glycerol and light orange for protein residues. Nitrogen, oxygen, and sulfur atoms are colored blue, red, and orange, respectively. The two top binding-site centers predicted from PASS are depicted as pink dotted-spheres. Two water molecules (OW1 and OW2) located at the glycerol-binding site are shown as red balls.

Figure 4. Speculative model of CmeR in its DNA-bound form. The N and C-terminal domains of the ligand-bound dimeric CmeR were individually aligned with those of the DNA-bound QacR (1JT0) to generate the DNA-bound form of CmeR. The two DNA recognition $\alpha 3$ helices in the dimer of CmeR are included in the model.

Figure 5.

Binding site prediction for CmeR. (a) CmeR complexed with glycerol. The bound glycerol molecules from the crystal structure and from prediction are colored pink and yellow, respectively. (b) CmeR complexed with taurocholate. The bound taurocholate from prediction is colored orange. (c) CmeR complexed with cholate. The bound cholate from prediction is colored yellow. All predictions were done using MEDock. The two top binding-site centers predicted from PASS are depicted as green dotted-spheres.

Table 1

Data collection, phasing and structural refinement statistics

	Native	SeMet infect. point	SeMet peak	SeMet remote
<i>A. Data collection</i>				
Wavelength (Å)	0.9795	0.9798	0.9795	0.9662
Space group	<i>P</i> 21212	<i>P</i> 21212	<i>P</i> 21212	<i>P</i> 21212
Cell constants				
<i>a</i> (Å)	93.3	93.0	93.0	93.0
<i>b</i> (Å)	37.4	37.3	37.3	37.3
<i>c</i> (Å)	57.6	57.5	57.5	57.5
Resolution (Å)	2.24 (2.33–2.24)	2.10 (2.18–2.10)	2.07 (2.28–2.07)	2.07 (2.14–2.07)
Completeness (%)	99.6 (97.7)	97.6 (85.1)	93.4 (84.1)	97.2 (80.3)
No. reflections	297,224	168,174	179,712	179,712
No. unique reflections	10,172	12,493	14,087	12,897
<i>R</i> _{sym} (%)	5.5(26.2)	5.8(25.0)	5.2(26.4)	5.1(27.3)
Average <i>I</i> /σ	16.6(4.7)	23.3(3.8)	22.3(5.3)	18.2(3.2)
<i>B. Phasing</i>				
Selenium atom sites			3	
Resolution range of data used (Å)			50–2.80	
Overall figure of merit			0.59	
<i>C. Refinement</i>				
<i>R</i> _{work} (%)	21.9			
<i>R</i> _{free} (%)	24.3			
rms deviation from ideal				
Bond angles (deg.)	0.8			
Bond lengths (Å)	0.006			
Ramachandran analysis				
Most favored regions (%)	91.4			
Allowed regions (%)	7.5			
Generously allowed regions (%)	1.1			
Disallowed regions (%)	0.0			