10-2002

Structural Biology of Bacterial Multidrug Resistance Gene Regulators

Michael H. Godsey
Oregon Health & Science University, mgodsey@cu-portland.edu

Ekaterina E. Zheleznova Heldwein
Children's Hospital/Howard Hughes Medical Institute

Richard G. Brennan
Oregon Health & Science University

Follow this and additional works at: http://commons.cu-portland.edu/msfacultyresearch

Part of the Chemistry Commons

Recommended Citation
Godsey, Michael H.; Zheleznova Heldwein, Ekaterina E.; and Brennan, Richard G., "Structural Biology of Bacterial Multidrug Resistance Gene Regulators" (2002). Faculty Research. 36.
http://commons.cu-portland.edu/msfacultyresearch/36

This Article is brought to you for free and open access by the Math & Science Department at CU Commons. It has been accepted for inclusion in Faculty Research by an authorized administrator of CU Commons. For more information, please contact libraryadmin@cu-portland.edu.
Structural Biology of Bacterial Multidrug Resistance Gene Regulators

Multidrug resistance (mdr) can be defined broadly as the ability of a cell to survive ordinarily lethal doses of more than one drug. Clearly, such resistance is a critical problem in the treatment of fungal and bacterial infections and cancer. Four general, but nonexclusive, mechanisms give rise to multidrug resistance: 1) detoxification by enzymatic modification or cleavage of drug; 2) genetic alteration of the intra- or extracellular targets; 3) decreased permeability of the cell membrane; and 4) active drug extrusion by multidrug transporters.

Paramount to our understanding of mdr is the issue of recognition of structurally dissimilar substrates and how drug binding effects function. In bacteria many multidrug transporters are regulated directly (locally) by transcription factors, which also bind the substrates of these transporters, i.e. the drug can act as a transcriptional coactivator or inducer. Multiple mdr transporter genes are also regulated globally by activators such as MarA that do not necessarily bind drugs (1). The regulators are of keen interest because they are more amenable to structural studies than the membrane-bound transporters and thus offer a greater chance to obtain high resolution views of multidrug binding. Moreover, the local gene regulators are equally interesting as their DNA complexes directly reveal the mechanism of multidrug transporter gene regulation. This minireview will summarize the structures of known bacterial mdr regulators. Because our focus is more structural the reader is referred to one of several recent reviews that discuss the more biological aspects of global and local mdr regulation (2–5).

MarA/Rob/MarR

MarA is a member of the AraC family of transcriptional regulators (6) and activates over a dozen genes comprising the Escherichia coli mar (multiple antibiotic resistance) regulon (7, 8). MarA does not bind antibiotics as part of its transcription regulation mechanism. This 129-residue protein binds as a monomer to asymmetric 20-bp operators (1), which contrasts to most prokaryotic transcriptional regulators that are dimers and bind palindromic or pseudopalindromic DNA (9–11). The structure of MarA bound to the marA operator reveals the basis of its DNA binding mechanism (Fig. 1a) (12). The protein consists of seven α helices, six of which comprise two three-helix bundles each containing a helix-turn-helix (HTH) motif. The remaining “linker” helix connects the two HTH motifs, thus creating a monomer with two DNA-reading heads. The independent HTH elements bind adjacent major grooves, and because the two “recognition helices” have different amino acid sequences each makes a distinct set of contacts to the DNA bases, thus explaining the ability of the MarA monomer to bind asymmetric operator sequences. To gain such specificity otherwise would require heterodimerization, such as that displayed by the eukaryotic transcriptional regulators Myc-Max (13, 14), Fos-Jun (15), and AHR-ARNT (16), or a second DNA binding element, e.g. a wing (17, 18).

To bind a cognate DNA site, MarA must significantly bend the DNA. Bending occurs because the recognition helices of MarA are separated by only 27 Å, and the conformational flexibility of the protein appears to be limited by the linker helix. Thus, to accommodate the shorter distance and bind consecutive major grooves, MarA “pulls” the major grooves toward itself, causing two kinks in the DNA at each HTH binding site and the narrowing of the minor groove. These localized kinks result in a global DNA bend of ~35° (12).

The structure of a second AraC family member, E. coli Rob, has also been solved bound to DNA (19). Rob, like MarA, can activate the mar operon when overexpressed (20). However, unlike MarA, Rob contains an additional ~200-residue C-terminal domain of unknown function but which is structurally similar to the E. coli galactose-1-phosphate uridyltransferase (21). Recent NMR studies on the activation of the rob regulon by 2,2′-dipyridyl have shown that the C-terminal domain of Rob binds this activator directly and that binding is necessary for up-regulation (22). As expected, MarA and the N-terminal, DNA binding domain of Rob, the sequences of which are 51% identical, are structurally similar. An overlay of all the major chain atoms of their conserved domains results in an r.m.s. deviation of 0.9 Å (19), yet their modes of DNA binding are quite different (Fig. 1a). Only the N-terminal HTH motif of Rob is inserted into the major groove. The C-terminal HTH sits on the surface of the double helix and makes a lone DNA contact from residue Arg-90 to the phosphate backbone. This binding mode allows Rob to contact unmeth DNA. Interestingly, Rob can also bind bent DNA (23) but to do so in a MarA-like fashion would require the relocation of a β hairpin (19). Understanding the mechanistic implications of the plasticity of the DNA binding mode(s) of this subclass of the AraC family will require additional structures of Rob-DNA and perhaps MarA-DNA complexes.

Whereas MarA, and perhaps Rob, responds to multidrug intrusion, the former constitutive activator does not bind drugs. Rather, its gene expression is negatively regulated by a dimer of MarR (7), a multidrug binding repressor that acts on the marRAB locus. MarR binds a variety of anionic lipophilic compounds, such as salicylate, which inhibit its repression of the marRAB operon (24). The structure of the 144-residue MarR protein bound to salicylate has been determined and reveals an αβ protein composed of 6 α helices and 3 β strands (25). Overall, the dimer appears to have the shape of a highway safety triangle (Fig. 1b). The dimer interface, which buries
3,570 Å² of accessible surface area, is formed mainly by the N and C termini of the monomers. A winged helix DNA binding motif is found in the middle of the protein sequence. In this conformation, MarR cannot bind B-DNA, which would explain the ability of salicylate to relieve repression. Interestingly, the structure of MarR was solved with two salicylate molecules bound per monomer, though EmR, another family member, appears to bind only one drug molecule per dimer (26). The “drug” binding pockets are readily accessible, and both use an arginine to neutralize the negative charge of the bound salicylate. Although salicylate is known to inhibit MarR (27), it remains unclear if one or both of these binding sites are biologically significant. A ligand-free structure, as well as those of MarR bound to DNA and bound to other “drugs,” is necessary to understand fully the action of this repressor.

**BmrR**

BmrR (Bacterial multidrug resistance Regulator) is a MerR family member (28) from *Bacillus subtilis* that, upon binding one of its structurally unrelated lipophilic cationic ligands, activates transcription of the multidrug transporter gene *bmr* by reconfiguring the 19-bp spacer between the −10 and −35 promoter elements (29, 30). Many BmrR ligands (coactivators) are also substrates of the multidrug transporter Bmr. Like most MerR family members, BmrR contains three domains: an N-terminal DNA binding domain, a linker/dimerization region, and a C-terminal coactivator binding domain. The C termini of MerR family members are quite variable, which is consistent with their functions to respond to a wide variety of cellular stresses.

To facilitate structural studies on BmrR, a 159-residue ligand binding domain, designated Brc (for BmrR C terminus) was crystallized (31), and the structures of its drug-free and tetraphenylphosphonium (TPP⁺) bound forms were determined (32). The BRC-TPP⁺ structure unveiled an internal multidrug binding pocket lined with multiple aromatic amino acids, which participate in van der Waals and stacking interactions with the four phenyl rings of the TPP⁺ ligand. Importantly, the bottom of the pocket features a glutamate residue, Glu-253, buried in the hydrophobic core of the protein. Glu-253 makes the key electrostatic interaction with TPP⁺, which is enhanced by the low dielectric constant of the protein interior (33). Interestingly, the binding pocket cannot be detected in the drug-free structure as residue Glu-253 is buried within the hydrophobic core where its carboxylate group is “neutralized” by hydrogen bonds to the hydroxyl groups of three tyrosine residues. To allow ligand access to this binding site, a short α-helical shield undergoes a helix-to-coil transition and moves away from the protein. The interactions between BRC and TPP⁺ suggest a similar binding mode for other hydrophobic cationic ligands of BRC/BmrR, such as rhodamine 6G (R6G) (32). Clearly, because all BmrR ligands are lipophilic cations, the electrostatic interaction between positively charged ligands and the negatively charged glutamate is key to the cation selectivity of BmrR. Indeed, the replacement of Glu-253 with the isosteric but electroneutral amino acid glutamine diminishes drug binding significantly but does not perturb the global structural integrity of the protein (32).²

The ligand selectivity of BRC depends not only on the presence of the negatively charged glutamate but also on the architecture and chemistry of the binding site. Although the structural flexibility of the binding site is unknown, the hydrophobic side chains likely rotate to accommodate different drugs. Mutational analysis of the BRC residues directly participating in TPP⁺ binding showed that each mutation affected the binding affinity of a variety of ligands differently (34). This suggests that although drugs bind in the same general location, each forms a distinct set of stacking and van der Waals contacts with the residues in the binding site. The structures of additional BRC-drug complexes should clarify this issue.

The structures of full-length BmrR in complex with DNA and a drug (35) and of MtaN (multidrug transporter activation, N terminus), also from *B. subtilis*, have been solved (36) (Fig. 2). BmrR is the first intact MerR family member to have its structure determined, and the structure of MtaN, a constitutive activator of transcription, is the first of a MerR family member solved in the absence of DNA or coactivator. Drug binding by BmrR is similar to that of the BRC in that residue Glu-253 is still involved and aromatic residues line the pocket and stack with TPP⁺. However, the BmrR structure also reveals the possible involvement of residue Asp-47, which is located in the DNA binding domain of the other subunit, and Glu-266. Site-directed mutagenesis of each of these acidic residues and drug binding studies are needed to elucidate their contribution to binding. Regardless, Glu-253 appears to be the electrostatic key to TPP⁺ binding because the BRC, which is missing Asp-47, binds this activator.

² J. L. Huffman, M. C. Miller, and R. G. Brennan, unpublished data.
DNA shown in **b**. QacR molecules bound to RG6, Et, and is displayed as respective binding sites, the R6G molecule displayed as ribbons. Shown in their respective. The backbones were overlaid from residue 55 to the C terminus and are displayed as ribbons. Shown in their respective binding sites, the R6G molecule is displayed as red balls and sticks and the Et as blue balls and sticks. The backbones of the two drug-bound forms are nearly identical with an r.m.s. deviation of 0.66 Å. However, the DNA-bound form differs significantly as described in the text. **c**, close-up view of the drug binding sites reveals that the two sites are distinct, yet partially overlapping. The "neutralizing" glutamates are shown as pink balls and sticks (Glu-90-R6G) and cyan balls and sticks (Glu-120-Et).

only ~5-fold less well than the full-length BmrR. BmrR and MtaN contain a similar winged helix DNA binding motif, which is contained within the structurally conserved N-terminal four-helix domain. However, a crucial difference between the two proteins is evident in their dimer conformations. The major groove binding helices (the "recognition" helices) are separated by 33.3 Å in the MtaN structure but only 30.6 Å in the DNA- and coactivator-bound form of BmrR, reflecting the shortening of the effective length of the consecutive major grooves of BmrR-bound bmr promoter. In addition, there is a 15° twist of the dimer partners, which is mediated through the antiparallel coiled-coil dimerization interface (36). This twist causes a 7.5 Å offset in the relative positions of the major groove binding helices. These differences suggest a mechanism of DNA activation by MerR family members in which DNA and coactivator binding cause a shift in the dimer conformation of the protein that leads to disruption of a centrally located Watson-Crick base pair and shortening and undertwisting of the DNA (28, 35, 37, 38). The unusual mechanism of QacR DNA binding is revealed by the crystal structure of a QacR-IR1 complex (42). QacR is an all helical protein, which contains a TetR-like HTH DNA binding motif embedded within the first three helices (43). Helices 4 and 5 comprise the drug binding and dimerization domains. Perhaps the most intriguing aspect of QacR-IR1 binding is that a tetramer does indeed bind as a pair of dimers but without the dimers contacting each other (Fig. 3a). Rather, cooperativity is effected by the DNA, the structure of which differs significantly from canonical B-DNA. The IR1 site is underwound (32.1° twist per bp and 11.2 bp per turn) and displays a widened major groove. These features are necessary to allow the QacR recognition helices of each dimer, which display a center-to-center distance of 37 Å, to bind consecutive major grooves. Thus, binding of the first QacR dimer and remodeling of the IR1 site allows the facile binding of the second dimer.

The crystal structures of QacR bound to six different drugs reveal its multidrug binding and drug induction mechanisms (44). The most striking finding is the presence of multiple but overlapping drug binding sites, which is consistent with data on multidrug transporters (4, 45). The elongated binding site, made from all C-terminal helices but α9, has dimensions of 10 × 9 × 23 Å and essentially contains two connected pockets. The R6G and the Et (ethidium) pockets of the site were named for the drugs which bind in them (Fig. 3c). The monovalent compounds crystal violet and malachite green bind in the R6G-Et overlap regions of the pockets, whereas berberine is located in the R6G pocket. The long, bivalent drug desquinilnum spans the site and fills both the R6G and Et pockets. Containing an elongated, multifaceted drug binding site, QacR displays a now obvious yet elegant mechanism for multiple drug recognition. Interestingly, QacR binds only one drug molecule per homodimer, unlike TetR, which binds two tetracycline molecules per dimer (44).
The QacR and BmrR drug binding sites contain a number of similar features. Both are lined by aromatic residues, which are found to stack against the ring structures of bound drugs. Moreover each protein uses buried or partially buried carboxylates to recognize their cationic ligands. In BmrR, Glu-253 appears to be key for the recognition of cationic drugs. By contrast, QacR contains four partially buried glutamates (residues 57, 58, 90, and 120) that line the extended drug binding pocket. In the binding of each cationic drug, the positive charge is neutralized by one or more of these glutamates. The QacR drug binding site also has several polar residues and solvent molecules, which interact with the hydrogen bond donors and acceptors of each drug, likely in a “nondiscriminatory” manner as each can act as a donor or acceptor.

QacR is induced by each drug, mono- or bivalent, by an identical mechanism. Upon binding drug, helix α5 of the drug-bound subunit undergoes a coil-to-helix transition and expels residues Tyr-92 and Tyr-93 from the hydrophobic core of the protein. This coil-to-helix transition extends the α5 helix by a turn, which now impinges upon helix α6 moving the latter helix toward the DNA binding domain. To maintain contacts between α6 and α1, the DNA binding domain translates by 9.1 Å and rotates by 36.7° (Fig. 3b). The change in one subunit is transmitted to the other (non-drug-bound) subunit through interactions between the α6 helices. However, the change in the position of the DNA binding domain of the drug-free monomer is smaller (3.9 Å translation and 18.3° rotation). The end result of these drug binding-induced changes is an increased distance between the recognition helices from 37 to 48 Å, thus eliminating specific DNA binding.

**Perspective**

We are now at the initial stages of elucidating the structural underpinnings of multidrug recognition and how these toxic molecules can regulate bacterial gene expression. Clearly many major questions remain. For instance, are multidrug binding pockets limited in their design or have multiple scaffolds arisen? The QacR and BmrR drug binding pockets share similarities such as the interaction of one or more acidic residues with their cationic ligands. In an analogous manner MarR uses an arginine to contact one of its anionic ligands. QacR and BmrR have multiple aromatic side chains stacking interactions with their ligands but MarR does not. The multidrug binding sites of QacR and BmrR are solvent-inaccessible and require a conformational change for drug access. By contrast, both salicylate binding pockets of MarR are solvent-exposed. Thus, the answer is still unclear. However, the coordinates of MexR, a MarR family member, have been deposited in the Protein Data Bank (46), and its structure should shed some light on the multidrug binding mechanism of this family. Can bacterial mdr gene regulators, like their transporter counterparts, bind more than one drug simultaneously? The structure of the QacR-dequalinium complex and the 1.100 Å volume of the multidrug binding pocket of QacR suggest such binding is possible. A similarly large volume is calculated for the drug binding pocket of PXR, the human nuclear xenobiotic response regulator (47). Formal proof will require the structure of a protein–two drug complex. Is there a structural relationship between the drug binding pockets of the cytosolic mdr gene regulators and membrane-bound transporters? Although not an mdr transporter per se, a 4.5 Å resolution crystal structure of E. coli MsbA, an ABC transporter, has been reported (48).