Structure of the *Escherichia coli* Antitoxin MqsA (YgiT/b3021) Bound to Its Gene Promoter Reveals Extensive Domain Rearrangements and the Specificity of Transcriptional Regulation*

Bacterial cultures, especially biofilms, produce a small number of persister cells, a genetically identical subpopulation of wild type cells that are metabolically dormant, exhibit multi-drug tolerance, and are highly enriched in bacterial toxins. The gene most highly up-regulated in *Escherichia coli* persisters is *mqsR*, a ribonuclease toxin that, along with *mqsA*, forms a novel toxin/antitoxin (TA) system. Like all known TA systems, both the Mqsr-MqsA complex and MqsA alone regulate their own transcription. Despite the importance of TA systems in persistence and biofilms, very little is known about how TA modules, and antitoxins in particular, bind and recognize DNA at a molecular level. Here, we report the crystal structure of MqsA bound to a 26-bp fragment from the *mqsRA* promoter. We show that MqsA binds DNA predominantly via its C-terminal helix-turn-helix domain, with direct binding of recognition helix residues Asn97 and Arg101 to the DNA major groove. Unexpectedly, the structure also revealed that the MqsA N-terminal domain interacts with the DNA phosphate backbone. This results in a more than 105° rotation of the N-terminal domains between the free and complexed states, an unprecedented rearrangement for an antitoxin. The structure also shows that MqsA bends the DNA by more than 55° in order to achieve symmetrical binding. Finally, using a combination of biochemical and NMR studies, we show that the DNA sequence specificity of MqsA is mediated by direct readout.

Biofilms, complex communities of bacteria that are encased in an extracellular matrix and adhere to almost any surface, are responsible for more than 65–80% of human infections (1, 2). Moreover, these infections are extremely difficult to treat because biofilms are both highly resistant to host defenses and tolerant to antimicrobials (3). Currently, a detailed understanding of how biofilms assemble, how they are regulated at a molecular level, and how they achieve antibiotic resistance is only rudimentarily understood. One significant mechanism by which biofilms achieve antibiotic tolerance is through the formation of bacterial persister cells (4–6). Persisters are phenotypic variants of the bulk culture that are slow growing and non-dividing (7). They are also exceptionally tolerant to antibiotics and able to survive (or “persist”) exposure to antibiotic concentrations that are toxic for regular cells (4, 7, 8). Although the frequency of persisters is only about 1 in 1 million in planktonic cultures (9), the frequency is as high as 1 in 100 in biofilms (10). The increased incidence of persister cells in biofilms and their role in human bacterial infections have stimulated renewed efforts to understand the molecular mechanism(s) that underlies the persister phenotype.

Recent studies have demonstrated that the frequency at which persisters occur in populations of *Escherichia coli* is determined by the expression of chromosomal toxin/antitoxin (TA) genes (4, 8, 11, 12). TA pairs are composed of two genes organized in an operon that encode an unstable, labile antitoxin and a stable toxin, which associate to form a stable complex (13–17). There are currently nine established toxin families (18), and similarities between toxins in fold and function are slowly emerging (19, 20). However, there is much greater diversity in structure and toxin recognition among antitoxins. TA systems are regulated by distinct differences in their lifetimes (21). Thus, the most important hallmark of antitoxins is their ability to be rapidly degraded by proteolysis. Under silent conditions, the activity of the toxin is blocked by the antitoxin. However, under conditions of stress, the antitoxins are degraded by proteases (21–24). This degradation outpaces the formation of the TA complex and leads to rapid growth arrest by the cellular effects of the toxin (25–27). TA complexes also regulate their own transcription, where the antitoxins bind to the promoter DNAs within their own operons and the cognate toxins enhance DNA binding (28–31). Although it is

---

*This work was supported by National Institutes of Health Grant F31 NS062630 (to B. L. B.). This work was also supported by National Science Foundation CAREER Award MCB-0952550 (to R. P.).*  
The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1 and Figs. S1–S7.  
The atomic coordinates and structure factors (code 3O9X) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).  
The sequence-specific backbone assignment of MqsA-C has been deposited with the Biological Magnetic Resonance Data Bank with accession number 17009.  
To whom correspondence should be addressed: Brown University, Box GE-4, Providence, RI 02912. Fax: 401-863-9653; E-mail: rebecca_page@brown.edu.

The abbreviations used are: TA, toxin-antitoxin; HTH, helix-turn-helix; MqsA-F, full-length MqsA; MqsA-N, N-terminal toxin-binding domain of MqsA; MqsA-C, C-terminal XRE-HTH domain of MqsA.
believed that multiple factors play important roles in TA expression, stability, and transcription regulation, a detailed molecular picture of the structural transitions that are required for DNA binding and recognition by TA pairs, and antitoxins in particular, is largely missing.

The gene most highly up-regulated in *E. coli* persister cells is *mqsR* (motility quorum-sensing regulator, *ygiL/h3022*) (12, 32, 33). Recently, we determined the crystal structures of MqsA alone and the MqsR-MqsA complex and showed that the *mqsRA* operon defines a novel family of TA systems in *E. coli*, in which MqsR is the toxin and MqsA is the antitoxin (19). The structure revealed that the MqsR toxin is a ribonuclease belonging to the RelE bacterial toxin family (19, 34, 35). MqsR is unique because it is the first toxin linked to biofilms and quorum sensing (33) and is the first toxin that, when deleted, decreases persister cell formation (31). The antitoxin MqsA is even more unusual because it is the first antitoxin that requires a metal, zinc, for structural stability (19); it is the only known antitoxin/TA pair that bind not only their own promoter but also the promoters of other genes that play important roles in *E. coli* physiology, including *mcbR* and *spy* (19) as well as *cspD* (31). Moreover, although 14 antitoxin structures have been solved to date, their primary sequences are highly divergent when compared with MqsA (supplemental Table S1) because the sequence identity between MqsA and the most closely related antitoxin, the *E. coli* protein HigA, is only 13%.

We recently showed that MqsA is a dimer, with the MqsR-MqsA complex forming a heterotetramer of composition MqsR-MqsA-MqsA-MqsR. MqsA is composed of an N-terminal zinc-binding domain, MqsA-N, which binds MqsR, and a C-terminal helix-turn-helix (HTH) domain, MqsA-C, which binds DNA and mediates dimerization (19). Electrophoretic mobility shift assays (EMSA) showed that DNA binding is mediated exclusively by MqsA-C because MqsA-N both alone and in complex with MqsR is not capable of binding the *mqsRA* promoter (*PmqsRA*) (19). However, the atomic resolution structure of MqsA suggested that the MqsA-N domains may also play a role in DNA binding. In addition, EMSA studies showed that the MqsR-MqsA complex binds *PmqsRA* DNA with a higher affinity than MqsA alone (35), suggesting that MqsR may also interact with DNA. Finally, the palindromic sequence of *PmqsRA* is separated by a TA tract, which typically leads to changes in DNA curvature (41). This suggests that MqsA recognition of *PmqsRA* DNA may be achieved using a combination of both direct readout (direct discriminatory contacts are made between the protein and the bases in the major groove) and indirect readout (no direct contacts are made between the protein and the bases, and, instead, specificity is achieved by water-mediated hydrogen bonding and/or the induction of DNA distortion) (42, 43).

In order to understand the structural basis of DNA binding, regulation, and specificity by the MqsA antitoxin, the x-ray crystallographic structure of MqsA bound to a 26-base pair DNA fragment derived from the *mqsRA* promoter was determined. These studies reveal that MqsA undergoes extensive domain rearrangements upon DNA binding and explains why both full-length MqsA and the MqsR-MqsA complex bind DNA with higher affinities than MqsA-C alone. A combination of EMSA competition assays and NMR spectroscopy titration studies was then used to show that MqsA recognizes *PmqsRA* DNA using direct readout, which is mediated by MqsA residues Asn97 and Arg101. Finally, our studies reveal that the MqsR-MqsA complex probably uses a mechanism of co-transcriptional regulation in which both MqsA and MqsR bind directly to DNA and provides further evidence that MqsA utilizes steric occlusion to mitigate MqsR-mediated toxicity.

**Experimental Procedures**

**Expression, Purification, and Crystallization**—Full-length MqsA (residues 1–131; MqsA-F), MqsA N-terminal domain (residues 1–76; MqsA-N), and MqsA C-terminal domain (residues 62–131; MqsA-C) were cloned and expressed as described previously (19). Multiple DNA duplexes that correspond to the palindromic region of *PmqsRA* were prepared and screened for crystallization with MqsA as described previously (44). Briefly, the MqsA *PmqsRA* dsDNA complex that formed crystals suitable for structure determination (hereafter referred to as MqsA(*PmqsRA*)) was prepared by incubating 24 µM MqsA with 30 µM *PmqsRA* dsDNA (TGTAATTACCTTTTAGGTTATAACT; palindrome is underlined) in protein stabilization buffer (10 mM Tris, pH 7.0, 50 mM NaCl, 0.5 mM Tris(2-carboxyethyl)phosphine hydrochloride). After incubation at room temperature for 30 min, the complex was concentrated 10-fold using a 3000 Da molecular mass cut-off centrifugal filter (Millipore) and immediately subjected to crystallization trials. A single merohedrally twinned crystal of the MqsA(*PmqsRA*) complex formed in 0.2 M ammonium chloride, 20% PEG 3350 at 4 °C in a drop containing 0.2 µl of protein and 0.4 µl of crystallization condition.

**Data Collection and Structure Determination**—The MqsA(*PmqsRA*) crystal was cryoprotected by a short soak in mother liquor supplemented with 20% glycerol and then frozen by direct transfer to liquid nitrogen. This crystal diffracted to 2.1 Å. X-ray diffraction data were collected at the NSLS X25 microfocus beamline at a wavelength of 0.9788 Å using an ADSC Q315 CCD detector. Data were indexed, integrated, and scaled with DENZO and SCALEPACK from the HKL-2000 program package (45). As reported previously (44), analysis with *phenix.xtriage* (46) confirmed the space group of *P*4_1 or *P*4_3 with a twin fraction of 0.372 and twin operator h,k,l. Phase determination and structure solution were performed with two successive rounds of molecular replacement using Phaser (47). First, a search was performed with two copies of the MqsA N-terminal domain (PDB entry 3GA8) and two copies of the MqsA C-terminal domain (PDB entry 3FMY). This produced a single solution in *P*4_1. A second round of molecular replacement was then used to search for an additional four copies of a 5-mer DNA duplex. A single solution was obtained with a final log likelihood gain of 975.
The \( R_{\text{free}} \) set (9.42%) was assigned using Phenix with the appropriate twin operator. Subsequent model building was performed in Coot (48), and model refinement was performed with Phenix (46) using NCS. The model contains one complex (one MqsA dimer complexed with one DNA duplex) with a total of 264 protein residues and 52 nucleotides. The electron density of the bound DNA is shown in supplemental Fig. S1. Analysis of DNA parameters was carried out using the programs Curves + (49), 3DNA (50), and Enzylate (51).

**Mutagenesis**—Amino acid mutagenesis was carried out using the QuikChange mutagenesis kit (Stratagene), and all constructs were sequenced (Agencourt). Mutated DNA constructs were purchased as individual complementary oligonucleotides (Integrated DNA Technologies) and then annealed as described previously (44). All complementary oligonucleotides corresponding to biotin-labeled P*\text{mqsRA}-I* were purchased and annealed. To determine dissociation constants, labeled WT P*\text{mqsRA} DNA was incubated with a 10, 20, 40, 70, 100, 150, 200, 400, or 600 nM concentration of either wild-type MqsA-F, MqsA-F N97A, MqsA-F R101A, or MqsA-F N97A/R101A (all proteins were purified immediately prior to EMSA experiments). A similar experiment using 50, 100, 150, 200, 250, 300, 350, 400, and 500 nM MqsA-C was also performed. To determine whether MqsA recognizes the P*\text{mqsRA} DNA via a direct or indirect readout mechanism, 1 fmol of labeled DNA was mixed with a 50-fold excess of MqsA-F and a 200-fold excess of either WT or mutant unlabeled DNA. The binding reactions were incubated at room temperature for 20 min. Samples were then loaded onto a 6% DNA retardation gel (Invitrogen) and subjected to electrophoresis at 4 °C for 75 min and 100 V in TBE (45 mM Tris, pH 8.3, 45 mM boric acid, 1 mM EDTA) buffer. The DNA was transferred to a nylon membrane at 390 mA for 45 min and subsequently UV-cross-linked at 302 nm. Chemiluminescence was performed with the LightShift chemiluminescent EMSA kit (Pierce), and the samples were detected with a CCD imager (Typhoon 9410 imager). Integrated band volumes were calculated using ImageQuant TL software, and dissociation constants were calculated using Sigmaplot 8.0. All experiments were performed in triplicate except for MqsA-C, which was done in duplicate.

**RESULTS**

**Structure of the MqsA-P*\text{mqsRA} DNA Complex**—The structure of full-length MqsA (MqsA-F; residues 1–131; MqsA is a dimer) bound to the 26-mer DNA fragment corresponding to the palindromic of the P*\text{mqsRA} promoter (referred to hereafter as MqsA-P*\text{mqsRA}) was determined to 2.1 Å resolution by molecular replacement (Fig. 1, A–D, and supplemental Fig. S1) using two copies of the MqsA N-terminal domain (MqsA-N), two copies of the MqsA C-terminal domain (MqsA-C), and four copies of a short 5-mer DNA duplex as search models. The space group is P4\(_1\), with one copy of the MqsA-P*\text{mqsRA} complex per asymmetric unit. The final model contains 264 protein residues and 52 nucleotides (Table 1).

The structure of the MqsA-P*\text{mqsRA} complex reveals that a single MqsA dimer binds one P*\text{mqsRA} DNA duplex using a large, positively charged surface on the undersides of the MqsA-C domains (Fig. 1B; the MqsA-C domain mediates dimerization). The MqsA dimerization axis is approximately perpendicular to the helical axis of DNA and passes through the center of the operator (base pair T14-A14). Because the dimerization axis possesses nearly perfect 2-fold symmetry, the protein interactions with the two DNA palindrome half-sites are symmetric. The DNA is further enclosed by the MqsA-N domains, which make direct interactions with the DNA phosphate backbone. The interface between the MqsA protein and P*\text{mqsRA} DNA is extensive, burying a total of...
3263 Å² of solvent-accessible surface area, ~13% of the solvent-accessible surface area of the entire complex.

*MqsA Induces a Large Bend in the DNA*—Complementarity between the DNA and MqsA is achieved by perturbations of the P*mqsRA* base pairs when compared with ideal B-form DNA. These perturbations result in an overall DNA curvature of 55.6°, enabling the MqsA-C recognition helices to bind symmetrically to each palindrome half-site (Fig. 1, D and E).

Protein stabilization of bent DNA is often achieved by binding the minor groove at AT-rich sites (53). This mechanism of stabilization is also observed in the MqsA-P*mqsRA* complex. Namely, the P*mqsRA* palindrome half-sites are separated by a short AT-rich region (Fig. 1C). In the MqsA-P*mqsRA* complex, these nucleotides exhibit a collapsed minor groove and increased helical curvature when compared with ideal B-form DNA. The collapsed minor groove is stabilized by multiple polar interactions between MqsA residues and the phosphate backbone. The central interaction, located at the most narrow region of the minor groove, is mediated by His110 from each MqsA monomer. Specifically, both His110 side chain NE2 atoms hydrogen bond with symmetrical phosphate oxygens from the opposing DNA strands (Fig. 1F). The narrowing of the minor groove is further stabilized by polar interactions from the MqsA-N domains, which serve to both “clamp” and immobilize the DNA.

**DNA Sequence Recognition Is Mediated by the MqsA Recognition Helix Residues Asn97 and Arg101**—The MqsA-P*mqsRA* interface is composed of 40 protein residues (20 from each monomer) and 24 nucleotides (12 from each half-site) (Fig. 2A, protein-nucleotide interactions identified using Entangle).
TABLE 1
Crystalllographic data collection and refinement statistics

<table>
<thead>
<tr>
<th>MqsA-PmqsRA</th>
<th>Data collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P4</td>
</tr>
<tr>
<td>Unit cell (Å)</td>
<td>a = b = 60.99, c = 148.60</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9788</td>
</tr>
<tr>
<td>Resolution (Å)*</td>
<td>50.0-2.1 (2.14-2.10)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>0.075 (0.515)</td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>29.13 (2.39)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.5 (93.5)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>6.7 (3.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range (Å)</td>
</tr>
<tr>
<td>Rwork (%)</td>
</tr>
<tr>
<td>Rfree (%)</td>
</tr>
<tr>
<td>Root mean square deviation</td>
</tr>
<tr>
<td>Bond length (Å)</td>
</tr>
<tr>
<td>Bond angle (degrees)</td>
</tr>
<tr>
<td>Average B factor (Å²)</td>
</tr>
<tr>
<td>Protein</td>
</tr>
<tr>
<td>Nucleic acid</td>
</tr>
<tr>
<td>Water</td>
</tr>
<tr>
<td>Ions</td>
</tr>
<tr>
<td>Ramachandran plot (%)</td>
</tr>
<tr>
<td>Favored</td>
</tr>
<tr>
<td>Allowed</td>
</tr>
<tr>
<td>Disallowed</td>
</tr>
</tbody>
</table>

*Highest resolution shell data are shown in parentheses.

b Rmerge = ΣhklΣi(|I(hkl)| - <|I(hkl)|>)/ΣhklΣi|I(hkl)| where I(hkl) is the ith observation of a symmetry-equivalent reflection hkl.

(51). Thirty-two of the 40 protein residues responsible for binding the PmqsRA promoter are located in the MqsA-C domain, consistent with the MqsA-C domain being both necessary and sufficient to bind DNA (19). Previously, the MqsA-C domain was classified as a member of the XRE family of helix-turn-helix (XRE-HTH) proteins (19). Indeed, as predicted based on homology to other XRE-HTH proteins, the conserved XRE recognition helices (MqsA-C residues 97–104, colored blue in Fig. 1A) bind into the major groove of DNA. Both MqsA-C recognition helices make equivalent interactions with each palindromic half-site.

Sequence recognition is mediated by residues Asn⁹⁷ and Arg¹⁰¹ from the conserved XRE helix α4, which form discriminatory contacts with the DNA bases in the major groove (Fig. 2B). Asn⁹⁷ and Arg¹⁰¹ make base-specific interactions with a total of eight nucleotides, four from one strand (16AGT19) and four from the antiparallel strand (18CAAT21). In addition to hydrophobic and van der Waals interactions, Asn⁹⁷ and Arg¹⁰¹ form key polar and hydrogen bonding interactions with the PmqsRA half-site nucleotide bases (Fig. 2B). Specifically, the side chain ND2 atom of Asn⁹⁷ forms hydrogen bonds with thymine 19 O4, whereas atom OD1 of Asn⁹⁷ forms a hydrogen bond with adenine 20’ N6. The guanidinium group nitrogens of Arg¹⁰¹ interact with the DNA bases in the major groove (16AGT19) and form hydrogen bonding interactions with the phosphate backbone. Additional nonspecific hydrogen bonds and salt bridge interactions from the MqsA-C domain are mediated by Arg⁸⁸ and Lys⁷⁹ in helix α2; Thr⁸¹, Gln⁸³, and Lys⁸⁶ in helix α3; Gly⁹⁴ and Gly⁹⁵ in the α3-α4 loop; Tyr¹⁰² and Lys¹⁰⁴ in the conserved XRE helix α4; and Ala¹⁰⁷ and Gln¹⁰⁸ in the α4-α5 loop.

To elucidate the importance of Asn⁹⁷ and Arg¹⁰¹ residues in DNA binding and transcription regulation, single mutants of both residues (MqsA-F N97A and MqsA-F R101A) as well as the double mutant (MqsA-F N97A/R101A) were generated and tested for their ability to bind and shift the 26-mer PmqsRA DNA fragment using EMSA (all experiments repeated in triplicate; size exclusion chromatography and circular dichroism polarimetry were used to verify that the mutations did not induce global changes in the protein structure prior to EMSA analysis; supplemental Fig. S3). In the presence of 20 nm WT MqsA-F (2-fold excess versus DNA), all of the labeled PmqsRA shifted, indicating complete binding (Fig. 2C). We also show that this shift is specific because the addition of excess unlabeled DNA abolished the shift of labeled DNA (supplemental Fig. S4). In contrast, at the same concentration, only a partial shift is observed for both single mutants, MqsA-F N97A and MqsA-F R101A, clearly indicating that they bind PmqsRA DNA less effectively than WT. The EMSA data were also used to quantify the interaction of the MqsA proteins (WT and mutants) with PmqsRA DNA. WT MqsA has the strongest affinity for PmqsRA DNA, with a Kd of 0.8 ± 0.4 nm. The interaction of PmqsRA with MqsA-R101A is weaker, with a Kd of 7.5 ± 4.5 nm. Finally, as expected from our structural analysis, the interaction of PmqsRA with MqsA-F N97A is even weaker, with a Kd of only 40.1 ± 7.3 nm. Thus, these data clearly show that Asn⁹⁷ is much more important for DNA binding than Arg¹⁰¹ because the MqsA-F R101A mutant binds DNA nearly 6-fold more strongly than the MqsA-F N97A mutant. Finally, the simultaneous mutation of both residues (double mutant MqsA-F N97A/R101A) completely abolished DNA binding (Fig. 2C), even at concentrations as high as 600 nm (60-fold excess). These data demonstrate that, as predicted by the MqsA-PmqsRA structure, both Asn⁹⁷ and Arg¹⁰¹ are essential for DNA binding and sequence recognition.

In addition, it is very interesting to note that no N¹H⁻N correlation peak in the two-dimensional ¹H−¹⁵N HSQC spectrum of MqsA-C was detected for Asn⁹⁷ (supplemental Fig. S2). This indicates that Asn⁹⁷ must either undergo a μs-ms conformational exchange or exchange with the solvent without adopting a single major conformation in the unbound state. This might prime Asn⁹⁷ for its role in DNA binding because it will allow for a rapid, enthalpic adaptation of the DNA bound conformation. The only other unassigned MqsA residue, Ser¹¹², does not participate in DNA binding but rather is located in helix α5 at the MqsA-C dimer interface.

DNA Binding Is Accompanied by Extensive Domain Rearrangements in MqsA—In the absence of DNA, MqsA (PDB entry 3GN5) is a highly elongated dimer, composed of a central MqsA-C domain dimerization core, with the MqsA-N domains of each monomer projecting outward from the dimer core (Fig. 3A, light blue) (19). Upon binding DNA, the MqsA-C domain dimer stays essentially unchanged, superimposing with that from apo-MqsA with a root mean square deviation of 0.934 Å. The differences between the apo- and DNA-bound MqsA-C dimers are due to a small collapsing of...
the XRE recognition helices toward the dimerization interface, which facilitates their interaction with DNA. This change is accommodated by the \(93\text{GGG}95\) loop that immediately precedes the XRE helices.

In contrast, both MqsA-N domains undergo extensive rearrangements upon binding DNA. Both MqsA-N domains rotate and twist by more than 106° (106°, monomer A; 120°, monomer B) (Fig. 3, A and B). This causes the MqsA-N domains to collapse about the DNA, transforming MqsA from a highly extended conformation into a narrow, elongated DNA “clamp.” Unexpectedly, these rotations transform the MqsA dimerization axis from a pseudo- into a nearly perfect 2-fold axis between both monomers, allowing both MqsA-N domains to have symmetrical interactions with the bound DNA (Fig. 3B).

These domain rearrangements are mediated by a short, two-residue flexible segment linking the MqsA-N and MqsA-C domains, composed of residues Thr68 and Val69. Flexibility between these domains has been shown previously by proteolytic digestion experiments (19). Whereas the MqsA-N and MqsA-C domains behave as rigid bodies, the \(\Phi/\Psi\) angles of the residues in the linker undergo large changes, by \(-45\) and \(165°\), respectively, for Thr68. These DNA binding-induced rotations result in the formation of a novel interface between the MqsA-N and MqsA-C domains, which buries 627 \(\AA^2\) of solvent-accessible surface area (supplemental Fig. S5). The central residue of this new interface is MqsA-N residue Asn65. In the apo-MqsA structure, Asn65 is entirely accessible to solvent. In contrast, in the DNA-bound structure, Asn65, whose position changes by 7.8 Å as a conse-

FIGURE 2. MqsA specificity for \(P\text{mqsRA}\) is mediated by Asn97 and Arg101 and stabilized by multiple nonspecific interactions. A, schematic diagram of all MqsA-\(P\text{mqsRA}\) interactions, including electrostatic, hydrogen bond, hydrophobic, and van der Waals interactions (identified using the program Entangle) (S1). Interactions for only one palindrome half-site are shown because there are equivalent interactions with the second half-site. All residues that interact with DNA are labeled. The residues that mediate base-specific electrostatic contacts, Asn97 and Arg101, are labeled in red. Residues from the MqsA-N domain that stabilize DNA binding are labeled in blue. B, primary base-specific hydrogen bonding interactions (dashed lines) mediated by Asn97 and Arg101 (shown as sticks). Amino acid labels are shown in black, and nucleic acid labels are shown in magenta. C, EMSA experiments using biotin-labeled \(P\text{mqsRA}\) (the same construct used in crystallization) and increasing amounts of either WT MqsA, the MqsA single mutants (MqsA-F N97A or MqsA-F R101A), or the MqsA double mutant (MqsA-F N97A/R101A), WT MqsA (top left), MqsA-F N97A (top right), and MqsA-F R101A (bottom left) bind the \(P\text{mqsRA}\) construct as shown by a change in DNA migration as compared with the DNA alone (lane 1 in each panel). However, the N97A and R101A mutants do so with a lower affinity than WT. MqsA-F N97A/R101A is unable to bind DNA (bottom right; gels have two control lanes that were removed for clarity, and full gels are shown in supplemental Fig. S6). All binding reactions contain the nonspecific poly(dI-dC) probe.
 Influence of the MqsA-N domain rotation, is almost completely buried from solvent. In this conformation, the Asn65-ND2 atom hydrogen-bonds to the backbone carbonyl of Gly105 from MqsA-C, whereas the Asn65-OD1 atom forms a hydrogen bond with Gln108-NE2. A second consequence of the MqsA-N rotation is that both residues Arg23 and Arg61 move from a position in which they face outwards, away from the dimer interface, into a position in which they face inwards, toward the bottom of the MqsA-C dimer (Fig. 3C).

Residues from the MqsA-N Domain Enhance DNA Binding—The structure of the MqsA-PmqsRA complex reveals that residues from the MqsA-N domain also contribute to DNA binding. The large rearrangement of the MqsA-N domains that accompanies DNA binding (Fig. 3, A and B) results in the formation of a DNA-binding pocket, with the C-terminal domain dimer forming the top of the pocket and the N-terminal domain β1-β2 turn and helix α1 forming the sides of the pocket. This structural change optimally positions multiple MqsA-N domain residues (Phe22, Arg23, Lys58, and Arg61, shown as sticks with black labels) to interact with the phosphate backbone of nucleotides G11', T14, T15, A19', and A20' (labeled in magenta). Polar interactions are indicated by black dashes; interactions for only one MqsA-N monomer are shown. E, the MqsA-N domain enhances DNA binding. Increasing amounts of MqsA-C were incubated with labeled DNA and monitored for protein binding. Compare with the upper left panel in Fig. 2C (MqsA-F). All binding reactions contain the nonspecific poly(dI-dC) probe.
Structure of the MqsA-PmqsRA DNA Complex
gens from each chain both directly, via salt bridge interac-
tions, and indirectly, via water-mediated hydrogen bonds. 
Additional electrostatic interactions with the DNA sugar and 
phosphate backbone are mediated by MqsA-N residues Lys58 
and Arg91, whereas hydrophobic interactions are mediated by 
Phe22.
To demonstrate that the MqsA-N domain plays a role in 
DNA binding, we performed titration EMSA experiments 
with the 26-mer PmqsRA promoter DNA and the MqsA-C 
domain (Fig. 3E). As can be seen, the MqsA-C domain alone 
binds DNA more weakly than MqsA-F (compare Fig. 2C (up-
per left) with Fig. 3E) because 40–50 nm protein (4–5-fold 
excess) results in only a partial shift for MqsA-C (Fig. 3E) but 
a complete shift for MqsA-F (Fig. 2C, upper left). The EMSA 
data were also used to determine the dissociation constant for 
MqsA-C and PmqsRA (Kd = 41.5 ± 4.3 nM), which shows 
that MqsA-C binds PmqsRA 50-fold more weakly than 
MqsA-F (Kd = 0.8 ± 0.4 nm). Thus, the EMSA data confirm 
that the conformational changes between apo-MqsA and 
DNA-bound MqsA are biologically relevant.
MqsA Recognition of DNA Is Mediated by Direct Versus 
Indirect Readout—DNA sequence recognition by DNA-bind-
ing proteins is often mediated by direct readout, in which dis-
criminatory interactions are made between the protein and 
the DNA bases. However, in some recognition events, no di-
rect contacts are made between the protein and bases. In-
stead, in these cases, DNA recognition is mediated by indirect 
readout, in which the proteins bind distinct conformations of 
distorted, or bent, DNA; these DNA distortions are often ac-
commodated by AT-rich sequences (42, 43, 53). Many mem-
bers of the XRE family of DNA-binding proteins, including 
MqsA, make direct contacts with the DNA bases in the major 
groove via the XRE recognition helix (27, 54). However, many 
XRE DNA-binding proteins, including MqsA, bind bent 
DNA, with the DNA distortions accommodated near AT-rich 
segments, and thus these proteins may also use indirect read-
out for sequence recognition (27). The PmqsRA palindrome is 
separated by a short AT-rich segment (\textsuperscript{13}TTT\textsuperscript{15}) that, in the 
MqsA-PmqsRA complex, mediates the 55.6° DNA bend via a 
very narrow minor groove (Fig. 1, C and E).
In order to determine the importance of direct versus indi-
rect readout in sequence recognition by MqsA, we performed 
EMSA experiments coupled with NMR spectroscopy titration 
studies between MqsA and various WT and mutant PmqsRA 
DNA duplexes (Fig. 4, A–D). The NMR studies were per-
formed with only the C-terminal domain because it is both 
necessary and sufficient for DNA binding (19), and it provides 
of all of the base-specific DNA interactions. Furthermore, the 
size of the MqsA-C-PmqsRA complex (~32 kDa) is reasonable 
for NMR measurements, whereas the size of the full-length 
complex (~45.9 kDa) would make measurements exceedingly 
more difficult. For these interaction studies, the sequence-
specific backbone assignment of the MqsA-C domain has 
been determined (supplemental Fig. S2).
To investigate the importance of direct readout in sequence 
recognition, two DNA mutants were generated in which the 
nucleotides involved in base-specific protein recognition were 
either mutated to the opposite purine or pyrimidine nucleo-
tides (e.g. A to G or T to C; PmqsRA-D1) or to the typical 
paired base of the opposite purine or pyrimidine (e.g. A to C 
or T to G; PmqsRA-D2; Table 2). EMSA competition experi-
ments were performed to determine how effectively unlabeled 
mutant DNA duplexes prohibit the binding of labeled WT 
DNA. Specifically, a 50-fold molar excess of MqsA was incu-
bated with labeled WT PmqsRA in the presence of an excess 
of unlabeled DNA (either WT or mutant) for 20 min and sub-
sequently analyzed by EMSA to detect binding. As expected, 
the excess unlabeled WT PmqsRA DNA effectively out-com-
peted the binding of the labeled WT DNA to MqsA, causing 
the intensity of the shifted band to disappear completely (Fig. 
4A, lane 3). However, the addition of excess unlabeled mutant 
PmqsRA-D1 or PmqsRA-D2 DNAs resulted in no decrease in 
the intensities of the shifted bands (Fig. 4A, compare lane 3 
with lanes 4 and 5), demonstrating that the two mutant DNA 
duplexes are not able to out-compete WT DNA for MqsA.
We then used NMR spectroscopy to determine how the 
changes in binding between WT and mutant DNAs are mediat-
ed at a molecular level. Specifically, two-dimensional \textsuperscript{1}H,\textsuperscript{15}N 
HSQC spectra were measured of the MqsA-C domain dimer 
alone and titrated with either WT or mutant DNA. Chemical 
shift differences between the spectra show changes of the lo-
cal environment. As expected, in the presence of WT 
PmqsRA, many significant changes in chemical shifts between 
apo- and DNA-bound MqsA-C are observed, indicative of 
binding (Fig. 4B). However, upon the addition of either 
PmqsRA-D1 (Fig. 4C) or PmqsRA-D2 (data not shown), addi-
tional significant differences in the chemical shifts are ob-
erved when directly compared with WT PmqsRA, demon-
strating that the mutant DNAs behave very differently from 
WT.
To determine if DNA geometry or curvature plays a role in 
protein binding, the three nucleotides in the intervening AT-
rich region were mutated to GC base pairs (e.g. A to G or T to 
C; PmqsRA-I; Table 2). EMSA competition experiments 
showed that DNA mutant PmqsRA-I behaves virtually identi-
cally to WT PmqsRA, with unlabeled PmqsRA-I out-compet-
ing labeled PmqsRA for MqsA-C binding nearly as effectively 
as unlabeled WT PmqsRA (Fig. 4A, compare lanes 3 and 6). 
The dissociation constant of PmqsRA-I for MqsA-F (2.8 ± 0.6 
nm) is also nearly identical to that of PmqsRA for MqsA-F 
(0.8 ± 0.4 nm; supplemental Fig. S6F). This demonstrates that 
the identity of the nucleotides separating the palindrome half-
sites is not important because the three intervening nucleo-
tides can be changed without significantly affecting the bind-
ing affinity of MqsA for PmqsRA. Furthermore, the NMR 
titration experiments confirm the EMSA competition results. 
Namely, the chemical shift changes in the two-dimensional 
\textsuperscript{1}H,\textsuperscript{15}N HSQC spectrum of MqsA-C titrated with PmqsRA-I 
were nearly identical to those observed with WT PmqsRA 
(Fig. 4D), demonstrating that MqsA-C binds PmqsRA and 
PmqsRA-I via identical interactions. These results, in con-
junction with the MqsA Asn\textsuperscript{97} and Arg\textsuperscript{101} mutant EMSA 
Studies (Fig. 2C), show that MqsA mediates DNA sequence 
recognition using direct readout.
DISCUSSION

DNA Binding by MqsA Is Unique among All Known Bacterial Antitoxins—These studies provide new and additional evidence that MqsA defines a novel family of antitoxins. First, MqsA binds DNA predominantly via its MqsA-C and not MqsA-N domains (Fig. 1), with 32 of the 40 residues contributing to DNA binding coming from the MqsA-C domain. Although MqsA binds bent DNA, whose collapsed minor groove is stabilized by multiple protein interactions, we showed that sequence recognition is mediated exclusively by direct readout (Fig. 4). XRE recognition helix residues Asn97 and Arg101 mediate the direct readout, making base-specific contacts with eight nucleotides from each palindrome half-site (Fig. 2). Moreover, when both residues are deleted, DNA binding is totally abolished. Several other MqsA-C residues also stabilize DNA binding through nonspecific interactions with the phosphate backbone or ribose sugar moieties. Second, in the presence of DNA, the MqsA-N domains twist and collapse by over 105° via a two-residue hinge that allows these domains to rotate freely as rigid bodies (Fig. 3). This rotation “clamps” the bound DNA and confers nearly perfect 2-fold symmetry of the MqsA dimer, previously unseen in the apo-MqsA structure. It also results in significant interactions between MqsA-N residues Phe22, Arg23, Lys58, and Arg61 with the DNA backbone. Indeed, the DNA binding affinity for full-
length MqsA is significantly higher than the binding affinity to the MqsA-C domain alone, further supporting the structural evidence that MqsA-N domains stabilize DNA binding (Fig. 3E). These domain rearrangements are unprecedented for DNA binding by a bacterial antitoxin.

To date, three antitoxin-DNA structures have been reported: the chromosome-encoded hipBA module from E. coli (27), the plasmid-encoded ccdAB module from E. coli (55), and the chromosome-encoded fitAB module from Neisseria gonorrhoeae (56). HipB, like MqsA, is one of the few antitoxins identified to date that recognizes DNA via an XRE-HTH DNA-binding motif (19, 27, 57, 58). Despite sharing a common DNA recognition fold, the sequence identity between the MqsA and HipB DNA-binding domains is only 16%.

As expected, both MqsA and HipB bind the major groove of DNA via their respective XRE recognition helices. However, because the MqsA and HipB dimerization interfaces are completely distinct from one another (supplemental Fig. S7), their bound DNAs curve in different directions. As a consequence, the XRE residues that mediate sequence recognition are not conserved, either in position or in sequence (supplemental Fig. S7). Moreover, HipB binds both DNA and the HipA antitoxin via its structured N-terminal domain. Whether or not conformational changes accompany DNA binding remains to be determined because the apo-structure of HipB has not yet been reported. However, because the HipB protein is composed nearly entirely of the XRE-HTH domain, large conformational changes like those observed in MqsA are not expected upon DNA binding.

In contrast, both the antitoxins CcdA and FitA recognize their operon promoters via a ribbon-helix-helix DNA-binding motif, in which the β-strands of the ribbon-helix-helix domain insert into the DNA major groove. In these proteins, the CcdA and FitA C-terminal toxin-binding domains do not interact with the bound DNA, and instead their DNA-binding and toxin-binding domains are entirely distinct. This differs significantly from MqsA, in which both the MqsA-C DNA-binding and the MqsA-N toxin-binding domains interact directly with the bound DNA. Aside from MqsA, whose structure is reported herein, the only antitoxin whose structure has been determined in both its apo- and DNA-bound state is CcdA (the structures of FitA and HipB have only been determined in the DNA-bound or toxin-bound state). In both the apo- and DNA-bound states, the CcdA toxin-binding domains are unstructured and do not contribute to DNA binding. In contrast, the MqsA-N domain, which binds the MqsR toxin, makes direct interactions with DNA and enhances DNA binding even in the absence of MqsR. Only MqsA, which is unique among antitoxins because it is structured throughout its entire sequence in both the apo-, DNA-bound,
and toxin-bound forms, exhibits the extensive conformational changes of the MqsA-N domains between the apo- and DNA-bound states that stabilize the MqsA-PmqsRA complex and enhance DNA binding (Fig. 3).

**MqsR-MqsA-PmqsRA Model Predicts a Role for the MqsR Toxin in Transcriptional Co-regulation**—Neither MqsR (35) nor the MqsR-MqsA-N complex (19) is capable of binding DNA. However, it is well established across many TA systems, including MqsR-MqsA (35), that TA complexes bind the operon promoter with higher affinity than the antitoxin alone. Two toxin-antitoxin-DNA structures are currently available (FitAB-FitA-DNA (56) and HipA-HipB-DNA (27)), and each displays a different mechanism of toxin co-regulation. In the case of the FitAB module, two FitB toxin monomers bind to two FitA antitoxin dimers to form a heterotetramer. This serves to “tether” the FitA antitoxin dimers on the DNA. Although the toxins do not interact with DNA, they facilitate binding by stabilizing the FitAB heterotetramer. In contrast, in the case of the HipBA module from *E. coli*, the HipA toxin interacts directly with the DNA backbone. In this case, DNA binding is enhanced through direct interactions between the DNA and the toxin.

To determine the potential mechanism utilized by the mqsRA TA module, we generated a model of the MqsR-MqsA-PmqsRA transcriptional regulator complex by superimposing the structure of the MqsR-MqsA-N complex (19) with the MqsA-PmqsRA complex (Fig. 5). Our model of the MqsR-MqsA-PmqsRA complex shows that the large rearrangements of the MqsA-N domains that occur upon DNA binding position the bound toxins directly below the XRE recognition helices. This would enable the toxins to interact with the bound DNA, most likely via nonspecific interactions with the phosphate backbone because MqsR itself does not bind DNA (19, 35). Thus, this model predicts that the mqsRA TA module most likely utilizes a mechanism of transcriptional co-regulation in which the MqsR toxin interacts directly with the PmqsRA DNA, a mechanism similar to that seen with the hipBA TA module.

**MqsR Toxicity Is Mitigated by Steric Occlusion**—Our model also provides new insights into the mechanism used by MqsA to mitigate MqsR-mediated toxicity. Previously identified residues that comprise the MqsR active site are Lys<sup>56</sup>, Gln<sup>68</sup>, Tyr<sup>81</sup>, and Lys<sup>96</sup>, all of which are accessible when bound to the MqsA-N domain (19). Unexpectedly, in the MqsR-MqsA-PmqsRA complex model, these residues are now facing inward, toward the pseudo-2-fold rotation axis of the MqsA dimer and underneath the bound DNA (Fig. 5). In this DNA-bound orientation, the active sites of the two MqsR toxins face one another and are separated by only ~13–15 Å. This orientation probably severely limits the accessibility of the active sites for mRNA and/or the ribosome, supporting a mechanism of toxin neutralization through steric occlusion.

**CONCLUSIONS**

The structure of the MqsA-PmqsRA transcription-regulatory complex provides new evidence that MqsA is highly unique among all known antitoxins. Unlike other antitoxins, MqsA requires metal for structural stability; toxin binding is mediated by the N-terminal domain, whereas DNA binding is mediated by the C-terminal HTH domain; and, most significantly, MqsA is structured throughout its entire sequence in the free, the toxin-bound, and now the DNA-bound forms (19). The structure of the MqsA-PmqsRA complex now reveals that MqsA is also unique because, unlike other antitoxins, MqsA undergoes unprecedented domain rearrangements upon binding DNA. These rotations result in orientations of the MqsA-N domains that enable the MqsR toxins to bind directly to the DNA, providing a molecular basis for the observation that MqsR enhances DNA binding and stability. Finally, these conformational changes provide new evidence that MqsA neutralizes MqsR toxicity by steric occlusion because, as modeled when bound to DNA, the MqsR functional sites face one another and define a narrow 13–15 Å channel immediately below the bound DNA, making them nearly inaccessible to other biomacromolecules, such as mRNA.

Therefore, this work has provided the first insights into understanding how this unique antitoxin regulates transcription at a molecular level and provides a basis for developing novel antibacterial therapies that target TA pairs.

**Acknowledgment**—Data for this study were measured at beamline X25 of the National Synchrotron Light Source (supported principally by the Offices of Biological and Environmental Research and of Basic Energy Sciences of the United States Department of Energy and by the National Center for Research Resources of the National Institutes of Health).

**REFERENCES**

Structure of the MqsA-PmqsRA DNA Complex