2774–2788 Nucleic Acids Research, 2014, Vol. 42, No. 4 Published online 28 November 2013
doi:10.1093/nar/gkt1215

Structural mechanism of transcription regulation of the *Staphylococcus aureus* multidrug efflux operon *mepRA* by the MarR family repressor MepR

Ivan Birukou[1], Susan M. Seo[2], Bryan D. Schindler[2], Glenn W. Kaatz[2,3] and Richard G. Brennan[1,*]

[1]Department of Biochemistry, Duke University School of Medicine, 307 Research Drive, Durham, NC 27710, USA, [2]The John D. Dingell Department of Veterans Affairs Medical Center, B4333 JD Dingel VA Medical Center, 4646 John R, Detroit, MI 48201, USA and [3]Department of Internal Medicine, Division of Infectious Diseases, Wayne State University School of Medicine, 5 Hudson, Harper University Hospital, 3990 John R, Detroit, MI 48201, USA

Received August 29, 2013; Revised October 17, 2013; Accepted November 4, 2013

ABSTRACT

The multidrug efflux pump MepA is a major contributor to multidrug resistance in *Staphylococcus aureus.* MepR, a member of the multiple antibiotic resistance regulator (MarR) family, represses *mepA* and its own gene. Here, we report the structure of a MepR–*mepR* operator complex. Structural comparison of DNA-bound MepR with 'induced' apoMepR reveals the large conformational changes needed to allow the DNA-binding winged helix-turn-helix motifs to interact with the consecutive major and minor grooves of the GTTAG signature sequence. Intriguingly, MepR makes no hydrogen bonds to major groove nucleobases. Rather, recognition-helix residues Thr60, Gly61, Pro62 and Thr63 make sequence-specifying van der Waals contacts with the TTAG bases. Removing these contacts dramatically affects MepR–DNA binding activity. The wings insert into the flanking minor grooves, whereby residue Arg87, buttressed by Asp85, interacts with the O2 of T_4 and O4′ ribosyl oxygens of A_{23} and T_4. Mutating Asp85 and Arg87, both conserved throughout the MarR family, markedly affects MepR repressor activity. The His14′:Arg59 and Arg10′:His35:Phe108 interaction networks stabilize the DNA-binding conformation of MepR thereby contributing significantly to its high affinity binding. A structure-guided model of the MepR–*mepA* operator complex suggests that MepR dimers do not interact directly and cooperative binding is likely achieved by DNA-mediated allosteric effects.

INTRODUCTION

The multiple antibiotic resistance regulator (MarR) family is a large protein family of bacterial transcription factors (1–4). It was named after *Escherichia coli* MarR, a transcription repressor of the *marRAB* operon, which produces MarA, the product of which triggers a multiple antibiotic resistance (mar) phenotype that is characterized by insusceptibility of the bacterium to a diverse repertoire of antibiotics, such as tetracycline, chloramphenicol, β-lactams, puromycin, nalidixic acid and flouoroquinolones (3,5–8). Other members of the MarR family have been shown to regulate the bacterial response to oxidative stress (9–12), expression of virulence determinants (2,13–17), multidrug resistance (3,4,18–27), transport of transition metals across the membrane (28) and catabolism of aromatic xenobiotics (29–31). Most MarR transcription regulators are repressors, whereby repression is achieved by binding of one or more protein dimers to the operator of the regulated gene. The binding site of the repressor often overlaps with crucial promoter elements (the –35 box, the –10 hexamer or the transcription start site (TSS)) thus inhibiting productive RNA polymerase binding. Some MarR-type repressors might affect transcription by competing for DNA binding with a transcription activator and conversely some representatives of the MarR family may act as activators by competing for binding with a repressor (2).

To date, only six structures of MarR family members bound to their cognate DNA are available. These include *Bacillus subtilis* OhrR (32), *Salmonella enterica* SlyA (33), *Mycobacterium tuberculosis* MosR (11), *Yersinia pseudotuberculosis* RovA (13), *Streptomyces coelicolor* SCO3205 (34) and *Sulfolobus tokodaii* ST1710 (22). Five assume a similar binding mode, in which the recognition helices of the winged helix-turn-helix (wHTH) domains from

*To whom correspondence should be addressed. Tel: +1 919 684 9471; Fax: +1 919 684 8885; Email: richard.brennan@duke.edu

both subunits slide into consecutive major grooves, while the wings insert into the flanking minor grooves. Binding specificity is achieved, mostly, by direct reading of the major groove bases by amino acid residues of the recognition helices and, to a lesser extent, via reading of the minor groove bases by the wings. Interestingly, the structure of the ST1710–DNA complex reveals a very different organization, in which only the wing motifs interact with DNA specifically via bases in the minor grooves, whereas the recognition helices neither insert into the major grooves nor establish specific, direct interactions with DNA bases (22).

MepR is a member of MarR family of transcriptional regulators. MepR represses the transcription of *mepA*, which encodes a multidrug resistance efflux pump belonging to the family of **m**ultiple **a**ntibiotics and **t**oxin **e**xtrusion (MATE) pumps, as well as the transcription of its own gene (19,21,23). MepR abolishes transcription by binding to operator regions, which harbour inverted repeats containing a GTTAG signature sequence. The *mepR* operator contains one inverted repeat, extending over the –10 hexamer and TSS, whereas the *mepA* operator contains two such MepR binding sites, which encompass the –35 and –10 elements and the TSS (19). The presence of two abutting MepR binding sites in the *mepA* operator region implies that MepR might bind to the *mepA* operator cooperatively. Repression is relieved upon the binding of MepA substrates to MepR followed by dissociation of the MepR–DNA complex. A typical MepA/MepR substrate is a hydrophobic, very often aromatic, positively charged molecule. Interestingly, the repertoire of substrates of MepR and MepA overlap, but are not identical (18–21).

In order to gain a fuller understanding of the repression and DNA recognition mechanisms of MepR, we determined the structures of a MepR–*mepR* operator DNA complex and characterized the amino acid residues directly and indirectly crucial for specificity and affinity of DNA binding using isothermal titration calorimetry (ITC) and electrophoretic mobility shift assays (EMSA) and *in vivo* by β-galactosidase assays. The MepR–DNA structures reveal a heretofore-unseen mechanism of binding site recognition by a MarR protein. Further, the MepR–*mepR* operator complex structure allowed the building of a model of the MepR–*mepA* operator complex in order to understand better the underlying mechanism of its posited cooperative binding.

MATERIALS AND METHODS

Cloning, mutagenesis, protein expression and purification

MepR with an N-terminal cleavable hexa-histidine tag was used for crystallization (see Birukou *et al.* (35) for details of its preparation). In most ITC experiments, EMSA and β-galactosidase assays recombinant MepR containing a non-cleavable C-terminal hexa-histidine tag was used. The plasmid used for its expression has been described previously (19). Protein expression was achieved using strain C41 DE3 (Lucigen) or BL21 DE3 One Shot (Invitrogen). The cell growth protocol was described by Kumaraswami *et al.* (21). Protein purification was conducted according to the protocol described previously in Birukou *et al.* (35). All mutagenesis was carried out using standard protocols and QuikChange^{TM} Site-Directed Mutagenesis (Agilent Technologies). Selenomethionine-substituted (SeMet) MepR was expressed using the methionine inhibitory pathway protocol (36) and purified as described for native MepR. The protein concentration was determined using UV spectroscopy and a calculated extinction coefficient $\varepsilon^{280} = 8.94\,mM^{-1}\ cm^{-1}$ (determined using ProtParam (37)).

Crystallization, structure determination and refinement of the MepR–DNA complex

Crystallization of the complex was carried out using the hanging drop-vapour diffusion approach. The DNA strands were annealed at 95°C for 5 min and cooled slowly to room temperature on the bench. The protein–DNA complex was prepared in advance by mixing concentrated MepR with 1.2-fold molar excess dsDNA to obtain 0.375–0.5 mM MepR–DNA complex solutions. Oligodeoxynucleotides of different length and sequence were tried in crystallization experiments. The DNA, that yielded diffraction-quality crystals of the MepR–DNA complex was the palindromic blunt-ended 24 base pair (bp) oligodeoxynucleotide, 5′-ATT TAG TTA GAT ATC TAA CTA AAT-3′.

One MepR–*mepR* operator complex crystallized in space group P1 and was obtained by mixing the MepR–DNA complex with mother liquor containing 0.14 M $CaCl_2$, 0.7 M Na acetate, pH 4.6, 14% 2-propanol and 30% glycerol in a 1:1 (v:v) ratio. The crystal diffracted to 2.84 Å resolution. X-ray diffraction intensities were collected at ALS beamline B.L. 5.0.2, wavelength = 1.000 Å. Data were processed using HKL3000 (38). Relevant parameters of data collection, integration and scaling are provided in Table 1. The structure of the MepR–*mepR* operator complex was solved via molecular replacement (Phaser-MR in Phenix, (39)) using the structure of the *B. subtilis* OhrR–DNA complex (PDB accession code **1z9c**, (32)) as a search model. Three MepR–DNA complexes were located in the asymmetric unit. After correcting the protein and DNA sequences and initial manual model building in Coot (40), the structure was refined in Phenix applying multiple rounds of positional, B-factor and Translation-Libration-Screw (TLS) refinement, interspersed with manual model rebuilding in Coot. Non-crystallographic symmetry (NCS) restraints were not imposed during refinement. The final model was refined to an $R_{work} = 20.6\%$ and $R_{free} = 24.1\%$.

The second crystal form took space group $P3_1$ and was obtained from solutions of 0.04 M Mg acetate, 0.05 M Na cacodylate, pH 6.0 and 30% (±)-2-methyl-2,4-pentanediol. The crystal diffracted to 3.0 Å resolution. X-ray diffraction intensities were also collected at ALS beamline B.L. 5.0.2 (wavelength = 1.000 Å). Data were integrated and scaled using HKL3000. Relevant parameters of data collection and processing are provided in Table 1. The initial model of the $P3_1$ MepR–DNA structure was obtained by molecular replacement using the

Table 1. Selected X-ray intensity data collection, structure solution and refinement statistics

	MepR-DNA	MepR-DNA	MepR-DNA (SeMet)
PDB accession code	4LLN	4LLL	
Data collection			
Space group	P1	$P3_1$	$P3_1$
Cell dimensions (Å, °)	76.3 76.2 109.4 90.6 104.7 106.2	130.2 130.2 124.7 90.0 90.0 120.0	128.9 128.9 125.7 90.0 90.0 120.0
Wavelength (Å)	1	1	0.979
Resolution (Å)	50-2.84 (2.90-2.84)[a]	50-3.0 (3.05-3.00)	50-3.65 (3.71-3.65)
R_{merge} (%)[b]	9.6 (69.2)	8.1 (>100)	10.8 (67.8)
I/σI	31.7 (2.0)	42.9 (2.0)	33.2 (3.6)
Completeness (%)	91.0 (65.4)[c]	100 (100)	100 (99.9)
Multiplicity	7.3 (6.2)	11.5 (11.1)	11.9 (10.6)
SAD phasing			
Resolution (Å)			50-3.65
Selenium sites			40
Figure of merit[d]			0.58
Refinement			
Resolution (Å)	44.84-2.84	42.61-3.04	
Reflections	47 769	45 454	
R_{work}/R_{free} (%)[e]	20.6/24.1	19.4/23.4	
Protein atoms	6352	8103	
DNA atoms	2934	2934	
Solvent atoms	0	0	
Average B ($Å^2$): protein/DNA	74.2/80.0	77.2/71.4	
RMSD bond lengths (Å)	0.003	0.009	
RMSD bond angles (°)	0.915	1.533	

[a]Values for the highest resolution shell are shown in parentheses.
[b]$R_{merge} = \sum\sum|I_{hkl} - I_{hkl}(j)| / \sum I_{hkl}$, where $I_{hkl}(j)$ is observed intensity and I_{hkl} is the final average value of intensity.
[c]80–100% completeness is observed for all resolution shells 3.0 Å and above.
[d]Figure of merit = $<|\sum P(\alpha)e^{i\alpha}| / \sum P(\alpha)|>$, where α is the phase and $P(\alpha)$ is the phase probability distribution.
[e]$R_{work} = \sum||F_{obs}| - |F_{calc}|| / \sum|F_{obs}|$ and $R_{free} = \sum||F_{obs}| - |F_{calc}|| / \sum|F_{obs}|$, where these reflections belong to a test set of 10% randomly selected data.

structure of the MepR–DNA complex obtained from the P1 crystal. Three MepR–DNA complexes were found in the asymmetric unit. However, after initial refinement the difference (mF_o–DF_c) electron density map contained numerous positive peaks suggesting that the model was incomplete. Thus, in order to obain *de novo* phase information we crystallized a MepR–DNA complex using SeMet-labelled MepR in the $P3_1$ space group using the conditions for the native protein–DNA complex. SeMet-MepR–DNA crystals diffracted only to 3.66 Å resolution. The X-ray diffraction data were collected at ALS beamline B.L. 5.0.2, using a wavelength of 0.979 Å, corresponding to the Se absorbance peak. The structure was solved using SAD phasing in combination with molecular replacement (SAD-MR in AutoSol, Phenix). The previously solved MepR–DNA complex was used as a search model. In total, 40 Se atoms were located in the asymmetric unit, which corresponds to 4 MepR dimers. Density Modification in Resolve (Phenix) resulted in a figure of merit of 0.58 and produced interpretable electron density maps that allowed us to correct the initial molecular replacement model and build a fourth MepR dimer. The complete model was refined further in Phenix using the native, higher resolution (3.0 Å) data. NCS restraints were not imposed on the equivalent protein and DNA chains during refinement of either structure. The final model has an $R_{work} = 19.4\%$ and $R_{free} = 23.4\%$. Selected crystallographic statistics are provided in Table 1. The analysis of DNA conformation was done using Curves+ (41). All structural alignments were performed in Coot (40) and PyMol (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC) and figures of the crystal structures were prepared in PyMol.

Isothermal titration calorimetry

All binding reactions were carried out at 25°C using a VP-ITC calorimeter (Microcal). Data were collected and analysed using the manufacturer supplied software package, Origin 7.0 (OriginLab Corp.). Each ITC run included an initial 4 µl injection followed by thirty 10 µl injections of the MepR solution into the sample cell containing DNA (total volume = 1.45 ml). In some experiments DNA was injected into the cell containing the MepR solution. The experimental details of each experiment are provided in the relevant figure legends. DNA was purchased from IDT and annealed by heating at 95°C for 5 min followed by slow cooling to room temperature on the bench. The buffer used in all ITC experiments was 20 mM Tris HCl pH 7.5 (room temperature), 150 mM NaCl and 5 mM $MgCl_2$. The fitted thermodynamic parameters as well as oligodeoxynucleotide sequences used in these experiments are provided in Supplementary Table S1.

Electrophoretic Mobility Shift Assay

EMSAs employing wild type (WT) MepR or the MepR(R79A) mutant and the WT *mepR* operator site (167 bp) were performed at least twice essentially as described, modified only by elimination of salmon sperm DNA as a non-specific competitor (19).

β-Galactosidase assay

Derivatives of *mepR* were cloned into plasmid pALC2073, placing them under control of the tetracycline-inducible *xyl/tetO* promoter, and then were transduced into the proper host strains (42,43). These included SA-K2908 or SA-K2916, which contain *mepA::lacZ* and *mepR::lacZ* transcriptional fusions, respectively (19,20). The *in vivo* efficiency of MepR derivatives in repression of chromosomal *mepR* or *mepA* expression was quantitated using a fluorescence-based β-galactosidase assay. This procedure has been described previously, and utilizes triplicate cultures of test organisms growing from exponential to stationary phase over a 10-h period. β-Galactosidase activity, which correlates with expression of the gene to which *lacZ* is transcriptionally fused, was quantitated using duplicate aliquots removed at specified time intervals (every 30 min from 1 to 5 h, then at 6, 8 and 10 h) (19). All strains were grown in the absence and presence of 0.05 μg/ml of tetracycline to quantitate the effect of induction of plasmid-based *mepR* expression. A control strain expressing WT MepR also was induced in the same manner and examined concurrently. β-Galactosidase activity was determined using an FLx800 fluorescence microplate reader (BioTek Instruments Inc., Winooski, VT). Cumulative *mepR* or *mepA* expression was determined by integrating the area beneath expression curves with SigmaPlot 12.0 (Systat Software Inc., Chicago, IL). All β-galactosidase assays were replicated a minimum of three times and data were normalized to the activity of WT MepR at the *mepA* or *mepR* operators in SA-K2908 or SA-K2916, respectively, which were considered to be 100%.

RESULTS AND DISCUSSION

The structure of MepR in complex with the *mepR* operator DNA

As a prelude to understanding the importance of the *mepR* operator sequence in the overall DNA-binding affinity of MepR and our effort to crystallize a germane MepR–*mepR* operator complex, we carried out several MepR–DNA binding assays using ITC. WT MepR binds to a 27 bp *mepR* operator site containing two 2-fold related GTTAG signature sequences that are separated by a 4 bp linker with a K_d of ~36 nM (Figure 1), which is similar to that previously reported, 24 nM, determined using a fluorescence polarization-based DNA-binding assay (21). Unfavourable positive enthalpy (ΔH = 8.73 kcal/mol) of the reaction along with large favourable entropy (ΔS = 63.3 cal/mol•deg, and corresponding $T\Delta S$ = 18.86 kcal/mol) clearly indicates that the binding is entropy driven (Supplementary Table S1). In protein–DNA complexes, two physical phenomena are responsible for the observed thermodynamic parameters. One is the polyelectrolyte effect, which originates from the displacement of DNA phosphate backbone-associated cations by positively charged amino acid residues of a partner protein and causes a large entropy increase. The second is the hydrophobic effect, which results from the burial of non-polar surfaces upon protein–DNA complex formation and concomitant release of structured water molecules from the hydration shell of a protein and

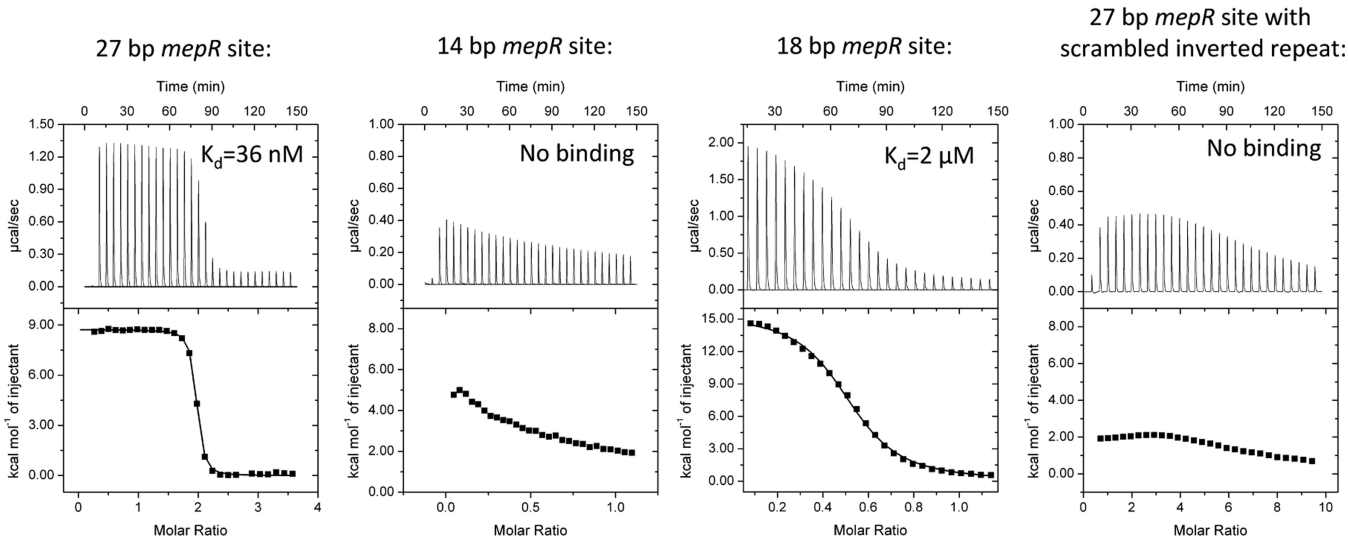

Figure 1. MepR binding to the *mepR* operator. ITC thermograms and binding isotherms for four MepR–DNA complexes. In the first and fourth binding experiment, protein (400 and 660 μM, respectively) was injected into DNA (24 and 15 μM, respectively); a stoichiometry of 2 corresponds to 1 dimer of MepR binding to dsDNA. In the second and third experiments, DNA (250 and 400 μM, respectively) was injected into protein (49 and 75 μM, respectively); a stoichiometry of 0.5 corresponds to a single dsDNA molecule binding to a dimer of MepR. The sequences of the oligodeoxynucleotides and fitted thermodynamic parameters are provided in Supplementary Table S1. The data for the 14 bp sequence and 27 bp DNA site with a scrambled recognition helix-binding region demonstrated no binding and were not fit. The molar concentrations of MepR in all experiments were presented per monomer.

major and minor grooves of a DNA, resulting in increases in the entropy and enthalpy (44). The thermodynamic parameters observed in our ITC experiments on MepR–*mepR* operator binding indicate that the prevalent binding interactions are non-specific electrostatic and non-polar.

Reducing the length of oligodeoxynucleotide to 14 bp to include only the GTTAG signature sequence essentially abolished MepR binding (Figure 1). The addition of a TA dinucleotide to the 5′ end and a GA dinucleotide to the 3′ end of the 14 bp improved the affinity ~30-fold ($K_d \sim 2\,\mu M$, Figure 1), but was still 55-fold lower than the WT value. As demonstrated previously increasing the length further to 22 bp restored high affinity binding (21). To assess the importance of GTTAG signature sequence, we constructed a 27 bp *mepR* operator site, in which the central 14 bp core, which encompasses both signatures sequences, was scrambled. The resulting thermogram demonstrates no specific DNA binding (Figure 1).

Using the binding data from our ITC and fluorescence polarization experiments and in order to crystallize a physiologically relevant MepR–*mepR* operator complex, we tested 10 different oligodeoxynucleotides of varying length and sequence. Multiple crystals were obtained, most of which diffracted to only 5 Å resolution. However, two conditions produced crystals, which diffracted to ~3.0 Å resolution. One took the **triclinic** space group P1 and a second took the **trigonal** space group $P3_1$. The DNA in both these crystals was a blunt-ended 24 bp oligodeoxynucleotide. The sequence of the crystallized oligodeoxynucleotide was modified slightly from the native *mepR* operator sequence to 5′-*T*AT TTA GTT AGA **C**AT CTA AC**G** AAA T*GG*-3′ (19) by removing a thymine from the 5′ end and two guanines from the 3′ end (removed bases are italicized), and by replacing cytosine 12 and guanine 20 (both are shown in bold font) with thymines. The result is the palindromic 24 bp oligodeoxynucleotide, 5′-$A_1T_2T_3T_4A_5G_6T_7T_8A_9G_{10}A_{11}T_{12}$ $A_{13}T_{14}C_{15}T_{16}A_{17}A_{18}C_{19}T_{20}A_{21}A_{22}A_{23}T_{24}$-3′, which allowed us to avoid potential statistical disorder of the DNA bases in the crystal structure.

The unit cell of the triclinic crystal form contains three copies of the MepR–DNA complex. The final model demonstrates excellent stereochemistry with 97.9% of the residues in the most favourable region and no residues in the disallowed region of the Ramachandran plot. The asymmetric unit of the trigonal crystal form contains three copies of the MepR–DNA complex as well as one additional MepR dimer bound to DNA non-specifically. This model also shows excellent stereochemistry with 97.4% of all residues in the most favoured region and the remaining 2.6% in the additionally allowed region of the Ramachandran plot.

Each MepR subunit is composed of six α helices and two β strands (Figure 2A). The N terminus of α1 (residues 1–14), the C terminus of α5 (residues 111–119) and α6 (residues 120–139) of both subunits of the MepR dimer form the dimerization domain. The wHTH DNA-binding domain is composed of the helices α2 (residues 31–44), α3 (residues 48–58) and α4 (the recognition helix, residues 60–76), and the wing motif is formed by β strands β1 (residues 77–81) and β2 (residues 89–93) and their connecting loop (residues 82–88). The dimerization domain is connected to the DNA-binding domain via the linker region, which includes the C-terminal half of α1 (residues 15–30) and the N-terminal portion of α5 (residues 95–110). $2F_o$–F_c omit electron density maps for both the triclinic and trigonal crystal forms that encompass the N-terminus of the recognition helix and its major groove-binding site, which contains the GTTAG signature motif, are shown in Supplementary Materials, Supplementary Figure S1.

Structure of *mepR* DNA

For the greater part, MepR binding to the *mepR* operator does not cause any dramatic changes in the DNA-binding site conformation when compared to ideal B-DNA (Figure 2A). We do not observe unusual base pairing or base expulsion. Conformational analysis of the DNA using the Curves+ web server (41) confirms that the DNA is predominantly in B-form. In total, six complexes from both MepR–DNA structures were analysed. Pairwise alignment of the DNA duplexes produced RMS deviations of 0.91–1.49 Å for trigonal crystal form, and 0.81–1.36 Å for triclinic crystal form. The largest deviations, as expected, were localized at the termini of the

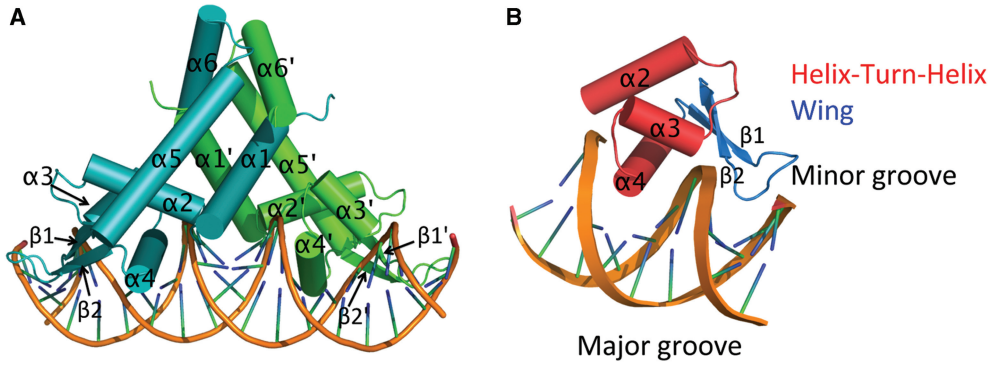

Figure 2. Crystal structure of MepR bound to the *mepR* operator. (**A**) Cartoon representation of MepR bound to a 24 bp oligodeoxynucleotide containing the *mepR* operator site. One subunit of the dimer is coloured blue and the other green. The secondary structure elements of each protomer are labelled with one set of labels primed. (**B**) The DNA-binding domain of MepR interacting with the major and minor grooves of DNA. Each element of the MepR wHTH motif is labelled.

duplexes where the base pairing is the weakest due to the loss of π-stacking stabilization on one of the sides. The duplex DNAs are only slightly bent, typically 5–10°, with only one outlier showing ~23° smooth bend. The average base pair twist is 33.8°, and the rise is 3.26 Å per nucleotide. The minor groove becomes wider and shallower in the centre of the DNA sequence: the width and depth of the central part encompassing bases 9–16 are 8–9 Å and 3–4 Å, respectively. The corresponding parameters for ideal B-DNA are 4.2 and 4.9 Å (41). The width of the major groove in contact with the recognition helices becomes slightly smaller (9–10 Å versus 10.5 Å in ideal B-DNA) but increases in the central part of the 24mer and at the ends (up to 13 Å). Further, the major groove is deeper in the central part of the sequence, 7–8 Å as compare to 5.4 Å in B-DNA, but the depth decreases at the ends (5–7 Å).

Protein–DNA interactions in MepR–DNA complex

Similar to other MarR family members, the *in vivo mepR* operator site is a pseudopalindrome, which we modified slightly to be a true palindrome. Helices α1 and α1′ of the dimerization domain lie atop the central part of the minor groove and make contacts with the sugar-phosphate backbone. Helices α2, α3 and α4 of the wHTH domains of both subunits interact with the major grooves encompassing the signature sequence GTTAG, while the wings insert into the flanking AT-rich minor grooves (Figure 2A and B). The majority of the protein–DNA contacts are non-specific and occur between the wHTH domains and sugar-phosphate backbone of the DNA (Figure 3). Most of these interactions are hydrogen bonds or salt bridges between the phosphate groups and charged or polar amino acid side chains, including residues of the recognition helix and its N-terminal turn (residues Arg59, Thr60, Thr63, Ser65, Arg69 and Asn70), wing residues (Arg79, Gln84, Thr86, Arg88 and Lys89), the N-terminal residues of α3 (Thr49, Gln50 and Asn51) and α2 (residues Thr30 and Gln33), and residue His14 of helix α1. The positively charged dipoles of the N-termini of helices α2 and α3 in both subunits additionally stabilize the binding via electrostatic interactions with the negative phosphate backbone, which is common for most MarR–DNA structures (11,13,30,32–34). There is also a multitude of van der Waals contacts from residues Gln18, Glu32, Arg59, Thr60, Thr63, Asn70, Asp85 and Arg88 to the deoxyribose rings and phosphates.

Interactions between the wings and minor grooves

The MepR wings interact with the minor grooves via residues Arg87 and its dyadic mate Arg87′ (Figure 4A). The guanidine group of Arg87 donates hydrogen bonds to the O2 atom of T_4 and O4′ atoms of the deoxyriboses of A_{23} and T_4 (Figures 3 and 4A). In some cases there is an additional weak hydrogen bond between Arg87 and the O2 of T_3. This residue also makes van der Waals contacts with the T_4 and A_{22} bases (Figures 3 and 4A). At the same time the Arg87 side chain is buttressed by a salt bridge to the carboxylate group of residue Asp85, which in turn is hydrogen bonded to the guanidinium group of residue Arg88 (Figure 4A, inset). To determine the importance

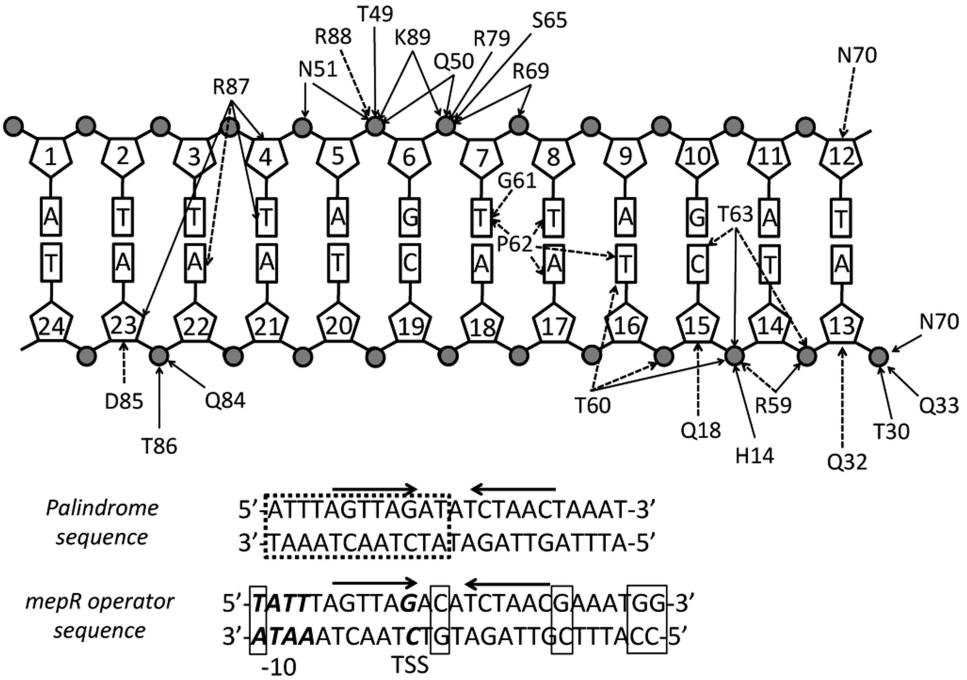

Figure 3. Schematic representation of protein–DNA interactions in the MepR–*mepR* operator complex. Due to the symmetric nature of the complex only half of the 24 bp DNA is shown. Solid arrows indicate electrostatic interactions (hydrogen bonds or salt bridges) and dashed arrows indicate van der Waals contacts. The lower part of the figure compares the sequence of the palindromic oligodeoxynucleotide used in crystallization to the actual sequence of the *mepR* operator. The dashed box indicates the part of the DNA shown in the interaction diagram. The thin line boxes indicate the base pairs, which were changed or removed to obtain the palindromic sequence. The promoter elements (the –10 box and TSS) are shown in bold italics. The arrows above the sequence indicate the inverted repeats GTTAGNNNNCTAAC.

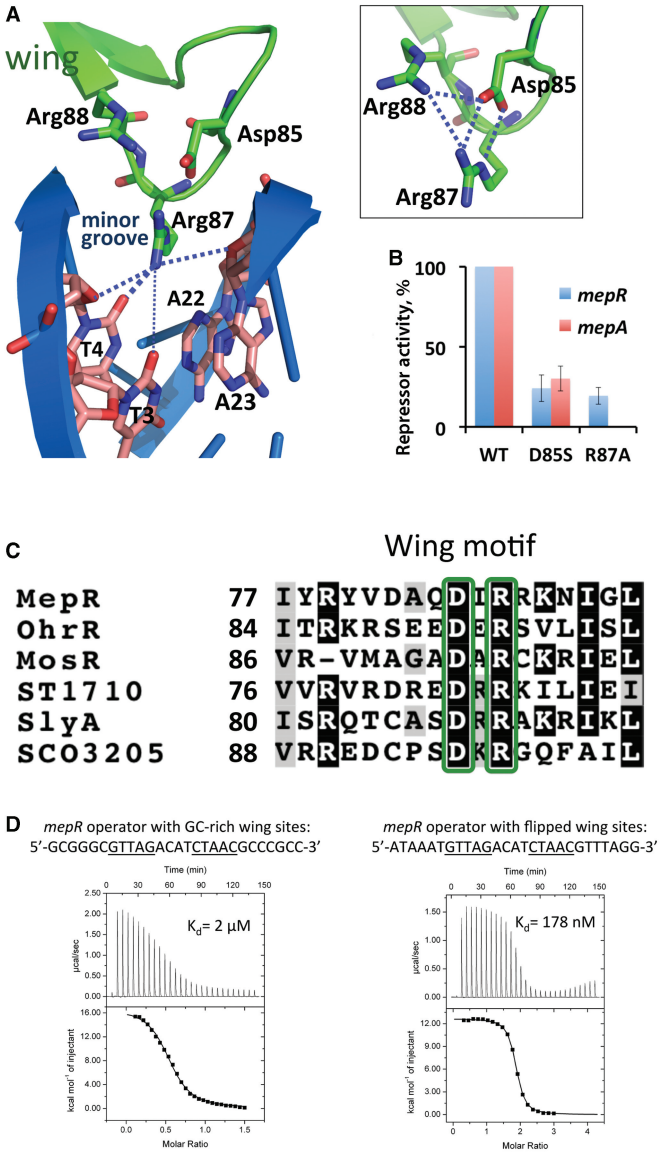

Figure 4. Interaction of the wing motif with the minor groove. (**A**) Residue Arg87 interaction with the O2 atoms of T_4 and T_3, and the ribosyl oxygens of T_4 and A_{23}. The boxed inset shows the electrostatic network between residues Arg87, Asp85 and Arg88. (**B**) Repression of the chromosomal *mepR* or *mepA* operators by MepR mutants D85S and R87A as determined by the ability of mutant MepR to affect the expression of the *mepR::lacZ* and *mepA::lacZ* transcriptional fusions. Data are presented as the mean repressor activity ± the standard deviation. The data were normalized to WT MepR expression and are provided in Supplementary Table S2. (**C**) Sequence alignment of the wing motif region of MarR proteins the structures of which have been solved in complex with cognate DNA. The conserved residues Asp85 and Arg87 are shown in green boxes. The alignment was created using ClustalW (45) and shaded using BoxShade 3.2. (**D**) MepR binding to the *mepR* operator with its wing-binding nucleotides modified. The sequences of oligodeoxynucleotides are indicated above the thermograms. In the left panel, DNA (400 µM) was injected into protein (55 µM); a stoichiometry of 0.5 corresponds to a single dsDNA molecule binding to a dimer of MepR. In the right panel, protein (400 µM) was injected into DNA (20 µM); a stoichiometry of 2 corresponds to 1 dimer of MepR binding to one dsDNA. The fitted thermodynamic parameters are provided in Supplementary Table S1.

of the Asp85–Arg87 pair for DNA binding, we created the MepR single mutants R87A and D85S and tested their ability to repress the expression of β-galactosidase from chromosomal *mepR::lacZ* and *mepA::lacZ* fusions (Figure 4B). Both mutations significantly impair the repressor function of MepR reducing its operator binding activity 5–10-fold in comparison to the WT protein, with the R87A change having a greater effect than the D85S substitution, especially at the *mepA* operator site which is not repressed at all by the MepR(R87A) protein. Similar defects in repressor activity were observed in MexR, SarA and MarR that contained substitutions of the corresponding arginine to alanine, histidine or cysteine (46–48).

The structural configuration of the tip of the wing motif is conserved among MarR family members (Figure 4C). In most MarR family member–DNA complex structures residue Arg87 is found to favour a thymine at the ninth position from the dyad axis of the palindrome (11,13,30,32–34). We tested whether Arg87 indeed discriminates in favour of thymine in MepR–DNA complex using ITC. The replacement of thymines at the 3rd and 4th position by guanines dramatically increases the dissociation constant of protein–DNA complex, from ~36 nM to ~2 µM (Figure 4D). Such a marked change is associated with the unfavourable steric effect of the exocyclic 2-amino group of guanine, which likely causes the Arg87 side chain to relocate to a position outside the minor groove. In contrast, changing the thymines to adenines at the same positions is less detrimental as the binding affinity drops by only ~5-fold (Figure 4D). The absence of an exocyclic group at position 2 of adenine still allows the insertion of the Arg87 side chain into the minor groove. Additionally, the O2 atom of T_{22}, which is complementary to A_3, may accept a hydrogen bond from the amino group of Arg87 side chain. Replacing T_3 and T_4 with cytosines should not disrupt MepR binding since both bases possess an O2 atom and can accept a hydrogen bond from Arg87. Indeed, such a replacement is present in the first MepR binding site of the *mepA* operator sequence (Figure 8), which exhibits an affinity similar to the *mepR* operator site (21).

Interactions between the recognition helices and major grooves

The N-termini of the recognition helices of MepR insert deeply into consecutive major grooves and interact with a number of bases. Intriguingly, these protein–DNA contacts are exclusively van der Waals interactions. Specifically, the side chain of residue Pro62 makes van der Waals contact with the exocyclic methyl group (distance, d = 3.6 – 4.0 Å) and the O4 atom (d = 3.8 Å) of T_7, with the amino group of A_{17} (d = 3.5 Å), with the O4 atom of T_8 (d = 3.2 Å), and with the O4 atom and C7 methyl group of T_{16} (d = 3.5 and 4.0 Å, respectively) (Figures 3 and 5A). Additionally, the carbonyl group of residue Gly61 interacts with the methyl group of T_7 (d = 3.7 Å), the hydroxyl of the Thr60 makes contact with T_{16} methyl group (d = 3.4 Å) and Thr63 hydroxyl group interacts with the C5 atom of C_{15} (d = 3.6 Å) (Figures 3 and 5A). Thus, the major groove is devoid of

base-specifying protein–DNA hydrogen bonds. Such hydrogen bonds are typical for protein–DNA complexes involving transcription factors that contain helix-turn-helix motifs and are observed in the structures of most MarR–DNA complexes (Figure 5C). For example, in the *S. enterica* SlyA–DNA complex, residue Arg65 of the recognition helix directly reads the guanine of its AA TCG signature sequence (33). In *B. subtilis* OhrR bound to cognate DNA many more electrostatic contacts are found: Ser68 makes a direct hydrogen bond to N7 atom of an adenine in the signature sequence TAACAT; in addition, the carbonyl group of residue Gly69 hydrogen bonds to the exocyclic amino group of cytosine and residues Asp67 and Arg77 make water-mediated contacts with the bases of the major groove (32). Similarly, *M. tuberculosis* MosR interacts with its binding site via hydrogen bonds between residues Arg70 and the first guanine, Thr71 and the cytosine complimentary to the second guanine, and Arg75 and the second guanine and the fourth thymine of the palindrome sequence GTGTA (11). More recently, *S. coelicolor* SCO3205 was shown to utilize hydrogen bonds between recognition-helix residues Arg72, Ser73 and His77 and the first, third and fourth bases of the ATTGA signature motif for specific binding (34). An interesting exception to these examples is observed in the crystal structure of the *Sulfolobus tokodaii* ST1710–DNA complex in which a dramatically different binding mode is observed whereby the recognition helix of one subunit lies above the sugar-phosphate backbone and does not insert into the major groove (22).

The absence of hydrogen bonds and other electrostatic interactions between the residues of recognition helix and the bases of the major groove in the MepR–DNA complex is unusual and unprecedented for the MarR family. Often, electrostatic interactions can be water-mediated as was observed in the OhrR–DNA complex structure (32). We did not observe any electron density, which could be attributed to solvent molecules in the vicinity of the recognition helices. To exclude the possibility, that the relatively low resolution of our structures prevented us from observing protein–DNA contact-mediating solvent molecules, we created a series of MepR mutants, in which residues potentially capable of establishing water-mediated contacts with the bases of the major groove, including Thr63, Asn66, Arg69 and Asn70, were replaced by alanine. We tested the DNA-binding activity of these MepR mutants using β-galactosidase assay. None of these mutations affected the repressor function of MepR (Figure 5B, Supplementary Table S2), with the exception of T63A, which demonstrated an ~2-fold loss of repressor activity for the *mepR* operator in comparison to WT MepR, but no loss of activity at the *mepA* operator. In contrast, ITC results showed that the MepR(T63A) protein binds the *mepR* operator site with WT affinity (Supplementary Figure S2). The reasons behind the observed 2-fold loss of the *in vivo* repressor function of MepR(T63A) are not clear but might be due to a lower expression level or stability in the SA-K2916 strain.

Thus, the crystal structure of the MepR–DNA complex and *in vivo* β-galactosidase assays with the selected MepR recognition-helix mutants support the idea that the specificity of MepR binding to cognate DNA is achieved solely via van der Waals interactions. Residues Thr60, Gly61, Pro62 and Thr63 appear to be the only structural

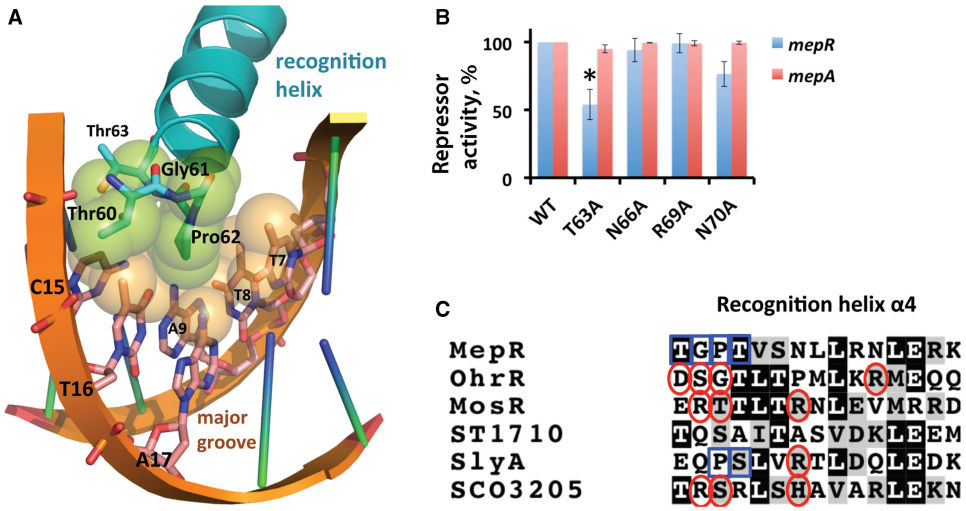

Figure 5. Interaction of the recognition helix (α4) with the major groove. (**A**) Residues Thr60, Gly61, Pro62 and Thr63, shown as atom-coloured sticks, establish van der Waals contacts with the TTAG bases of the GTTAG signature sequence, shown as sticks and atom-coloured. The atoms involved in van der Waals interactions are shown as semi-transparent spheres with the carbon atoms of the DNA bases coloured beige and those of the amino acid residues coloured green. Key residues and bases are labelled. (**B**) Repression of the chromosomal *mepR* or *mepA* operator by MepR mutants T63A, N66A, R69A and N70A as determined by the ability of the mutant protein to affect the expression of *mepR::lacZ* and *mepA::lacZ* transcriptional fusions. The data are presented as the mean repressor activity +/– the standard deviation. The data were normalized to MepR WT and are provided in Supplementary Table S2. (**C**) Sequence alignment of the recognition helices of selected MarR family proteins. Red circles indicate the residues interacting electrostatically with the major groove bases in their corresponding protein–DNA complexes and the blue boxes represent van der Waals contacts. The alignment was created using ClustalW (45) and shaded using BoxShade 3.2.

factors that discriminate in favour of the TTAG bases of the GTTAG signature sequence. To test further whether the observed van der Waals interactions indeed determine specificity of binding, we monitored WT MepR binding to DNA containing altered GTTAG half-site sequences. Therefore, we designed three mutated sequences, GTTT G, GATAG and GTAAG, which altered only one of the signature sequence bases and would eliminate their observed van der Waals interactions between the recognition helices and the major groove. These mutated sequences bound WT MepR with only marginal decreases, at most ~3-fold, in DNA-binding affinity (Figure 6). However, when all three mutations are combined to give the quintet, GAATG, the dissociation constant increases to 10 micromolar (Figure 6). Modelling of the substitution of the last guanine:cytosine with an adenine:thymine pair (GTTAG to GTTAA) reveal the introduction of severe steric clash between the methyl group of the thymine

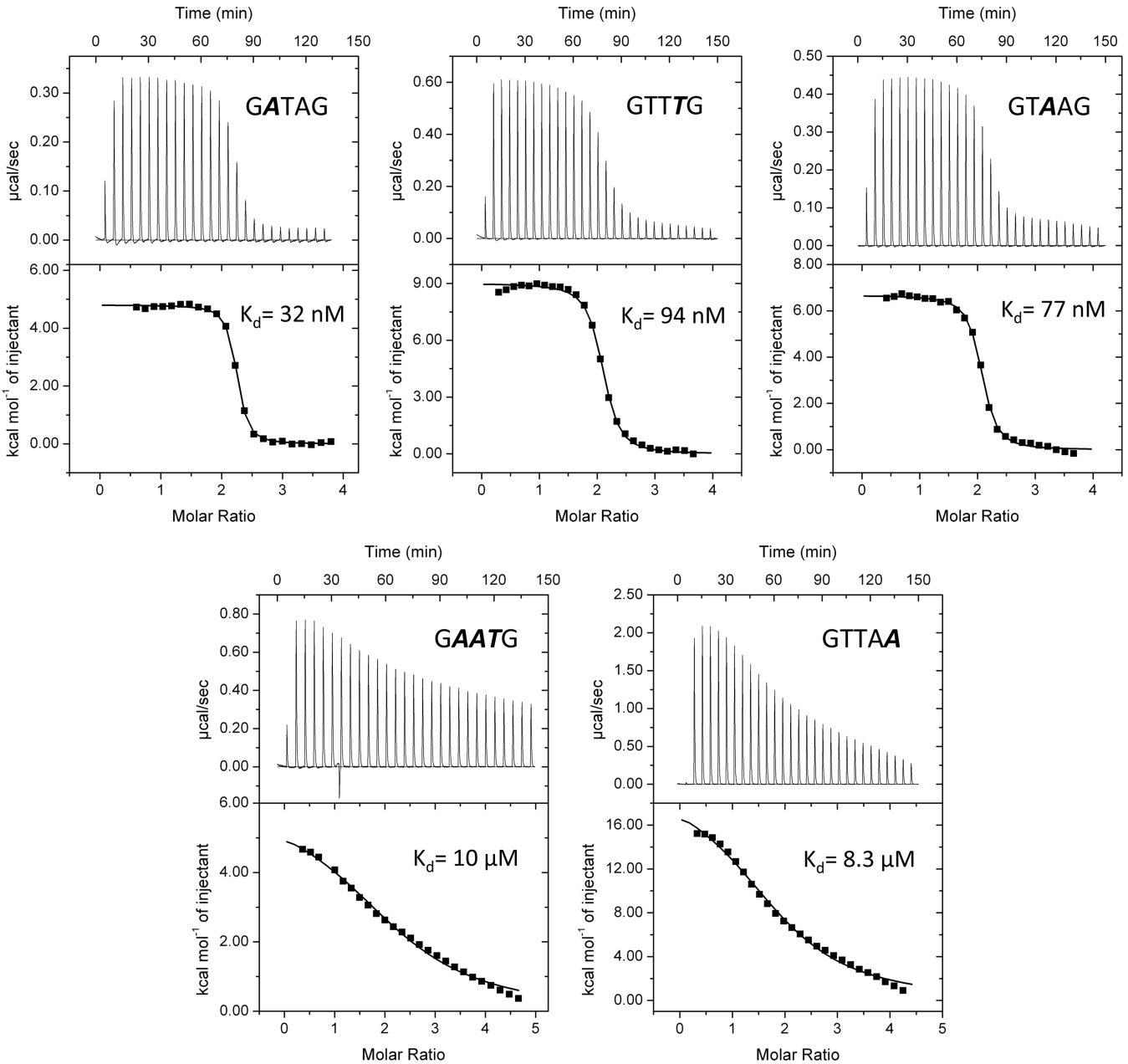

Figure 6. MepR binding to mutated signature sequences of the *mepR* operator. Variants of the WT signature sequence GTTAG are shown with mutated nucleotides shown in bolded italics. The 'palindromized' 26 bp oligodeoxynucleotide 5′-TATTTAGTTAGATATCTAACTAAATA-3′ was used as the template in which to make appropriate base pair changes in the signature sequence for binding studies. In all experiments, protein was injected into DNA. A binding stoichiometry of $n = 2$ corresponds to 1 dimer of MepR binding to one dsDNA. The fitted thermodynamic parameters are provided in Supplementary Table S1. In the first experiment, 200 µM MepR was injected into 10 µM dsDNA; in the second and third—185 µM MepR was injected into 10 µM dsDNA; in the fourth—452 µM MepR was injected into 20 µM dsDNA; and in the fifth—412 µM MepR was injected into 20 µM dsDNA.

base and residue Thr63. As expected, this mutation completely abrogates MepR binding. Thus, the recognition of specific DNA-binding sites by MepR is dictated by van der Waals interactions. Because this type of interaction is less discriminating and energetic than a typical hydrogen bond, more van der Waals contacts are necessary to achieve the required DNA specificity, which explains the small increases in K_d for the single mutations of the TTA triplet of GTTAG motif and the necessity to introduce multiple changes to lower the binding affinity significantly. In accord, neither the single MepR mutation P62A, previously reported (49), nor the single mutation T63A, described in this work, have affected the repressor function of MepR. In addition, MepR does not only discriminate in favour of a particular sequence, but also against some specific sequence alterations, as observed for the GTTAA and GAATG mutations (Figure 6).

Distal protein–protein interactions stabilize the DNA-binding conformation of MepR

A more detailed investigation of the DNA-bound form of MepR revealed two protein–protein interaction networks, which are absent from the apoMepR structure but we posit are important in stabilizing the DNA-binding conformation of MepR. One network involves the side chains of the residue Arg10′ (where the prime indicates a residue from the other subunit of the MepR dimer), His35 and Phe108, which stack upon each other in a manner that is similar to DNA base stacking (Figure 7). In most of the independent copies of the MepR–DNA complex, the planes of the guanidine, imidazole and benzyl side chains are nearly parallel, and the distances between these planes are 3–4 Å. The side chains of residues Arg10′, His35 and Phe108 contain delocalized π-electrons and can participate in π–π or cation–π interactions. In the apoMepR structure, residues Arg10′, His35 and Phe108 do not interact and are located 7–10 Å away from each other (Supplementary Figure S3B). To test whether this triplet interaction contributes to, or is absolutely necessary for specific DNA binding we created the single MepR mutants R10S and H35A and measured their DNA-binding affinities (Supplementary Figure S3A). Both MepR(R10S) and MepR(H35A) displayed ∼10-fold decreased affinities for the *mepR* operator (K_d (R10S) = 300 nM and K_d (H35A) = 420 nM). The second protein–protein interaction network involves residues His14′ of α1 and Arg59 found in the loop between helices α3 and α4 of the neighbouring subunit (Figure 7). The planes of the imidazole side chain of His14′ and the guanidinium group of Arg59 are not necessarily parallel to each other but all within 3–4 Å, which again indicates π–π or cation–π interactions similar to those in Arg10′:His35:Phe108 triad. In apoMepR, residues His14′ and Arg59 do not interact: the closest His14′–Arg59 approach is 13.5 Å, (Supplementary Figure S3B). The effect of substituting His14 with alanine on DNA-binding activity of MepR is identical to the R10A and H35A changes: an ∼10-fold increase in the dissociation constant (K_d = 380 nM) compared to WT MepR (Supplementary Figure S3A).

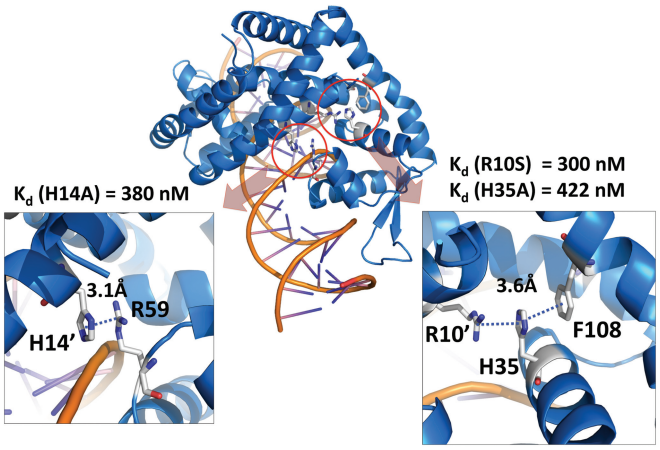

Figure 7. Intersubunit contacts buttress the DNA-binding conformation of MepR. Interacting residues of His14′:Arg59 and Arg10′:His35:Phe108 are shown as sticks; the distances between the residues are indicated. The dissociation constants (K_d) of MepR mutants H14A, R10S, H35A are provided.

Beyond these interactions, residue Arg79, which is located in the middle of the β1 strand of the wing, also plays an important role in the DNA-binding activity of MepR. Residue Arg79 is involved in a network of electrostatic interactions with residues Gln50 of α3 and Glu72 of α4 (Supplementary Figure S3C), which helps to preserve the proper fold of the helix-turn-helix motif, as this network is present in both the apo and DNA-bound forms of MepR. The same Gln–Arg–Glu triad is found in the SlyA–DNA structure, whereas in the OhrR–DNA complex only the Arg–Glu interaction is observed. The MepR mutant in which residue Arg79 is substituted by alanine has very poor DNA-binding activity as evidenced by EMSA (Supplementary Figure S3D).

The structure of a non-specific MepR–DNA complex

In addition to the three independent specific MepR–DNA complexes that are seen in the asymmetric unit of the trigonal crystals, an additional MepR dimer is bound to DNA. However, unlike the specific complexes, this fourth dimer interacts with two different DNA molecules acting as a non-specific tether (Supplementary Figure S4A). The dimerization domain of the non-specifically bound MepR dimer is significantly disordered and due to poor electron density, residues 115–124 could not be built. Regardless, the overall configuration of the dimer is discernable and is distinct from the specific DNA-binding conformation of MepR. Indeed, the distance between the centres of the recognition helices is ∼42 Å, and the orientations of recognition helices are substantially different from those of MepR in complex with DNA. As a result, the alignment of the dimer from the MepR–DNA complex with the non-specifically bound MepR dimer produces a striking RMSD of 6.5 Å. Further, the conformation of the non-specific MepR dimer is also distinct from the apoMepR dimer reported by Kumaraswami *et al.* (21) (RMSD = 4.9 Å), which again demonstrates the great flexibility and adaptability of the MarR family members. In general

terms, almost all protein–DNA interactions observed in the non-specific complex are confined to the DNA backbone and involve many, but far fewer of the residues that also make DNA-backbone interactions in the specific complexes (compare Figure 3 with Supplementary Figure S4B). In contrast, the wing is still engaged in hydrogen bonds with minor groove bases but not those contacted in the specific complex (Figure 3, Supplementary Figure S4B). We propose that this MepR–DNA complex provides fortuitous insight into one of the non-specific DNA-binding modes of MepR and potentially other MarR family members. A more detailed description of the non-specific MepR–DNA complex is provided in the Supplementary Materials.

Comparison of DNA-bound MepR with a DNA-free MepR

Comparison of the structure of DNA-free MepR reported previously (21) to DNA-bound MepR demonstrates that significant conformational changes are required for specific DNA binding. Particularly, the wHTH domain of each MepR subunit undergoes ~45° rotation towards the α1 helix of the neighbouring subunit during complex formation (Supplementary Figure S5A). Such movements of the DNA-binding domains place the recognition helices ~32 Å apart (the distance from the Cα atoms of residues Leu68 and Leu68′), which is an ideal distance for the interaction of the recognition helices with consecutive major grooves and the wings with the flanking minor grooves of B-DNA. Key to the observed conformational change is the distortion of the linker region such that a kink is created at position 103 of helix α5 and helix α1 is bent at position 16. Such kinking and bending of helices α5 an α1 in the specifically DNA-bound MepR corroborates our recent finding, based on a comparison of the apoMepR structure with the structure of a second ligand-free MepR dimer, which assumed a conformation more compatible with DNA binding and underscores the importance of flexibility of the MepR linker region in adopting a quaternary structure that is optimal for specific DNA binding (35).

Comparison of the MepR–DNA complex to other MarR–DNA complexes

The two crystal structures described herein provide six independent copies of the MepR–DNA complex. These align well within experimental errors with RMSDs between 0.3 and 0.6 Å (here and below only the protein dimers of each complex were included in the alignment). The overall shape of the MepR–DNA complex is triangular, which is similar to other structures of MarR family members bound to DNA (Supplementary Figure S5B, (11,13,30,32–34)), with the one exception, ST1710 (22). The structures of DNA-bound MepR dimers superimpose very well on the dimeric structures of DNA-bound OhrR (PDB accession code **1z9c**) (RMSD = 1.6 Å) and DNA-bound SCO3205 (PDB accession code **3zpl**) (RMSD = 1.7 Å) with the largest deviations in both alignments found in the dimerization domains. Also in SCO3205, the C-terminal tail is 12 residues longer than in MepR and extends into the DNA major groove, where its terminal residue, Arg163, interacts with residue Arg72 of the recognition helix, which in turn makes a sequence-specific contact to A_5 or a backbone contact with the phosphate group of G_4 (34). A significantly worse alignment is observed with DNA-bound SlyA dimer (PDB accession code **3q5f**) (RMSD = 2.6 Å), in great part because the DNA-binding domains are farther apart in the SlyA–DNA complex. However, when individual MepR subunits are aligned with subunits of SlyA from the SlyA–DNA complex, the RMSD diminishes to 1.0 Å. Structural alignment of the DNA-bound forms of dimeric MepR and MosR (PDB accession code **3gfi**) also reveals a fairly large RMSD (2.3 Å). Similar to OhrR, the largest discrepancy is concentrated in the dimerization domain, while the wHTH motifs are structurally conserved. Finally, the structure of ST1710–DNA complex (PDB accession code **4fx4**) stands out due to its unusual mode of interaction with DNA: the sequence-specific interactions are found only between the wing motifs and the minor grooves, whereas at the same time, the recognition helices provide only non-specific contacts to the sugar-phosphate backbone of the DNA and do not interact with the major grooves. The resulting superposition of a DNA-bound MepR dimer and a DNA-bound ST1710 dimer results in an RMSD = 4.7 Å. Thus with the exception of ST1710, the MarR family members bind their cognate DNA sites using similar, but not identical conformations with their major structural differences confined primarily to the dimerization interfaces.

A model of MepR bound to *mepA* operator

The *mepR* operator contains a single inverted repeat and, thus, binds only one MepR dimer (Figure 3, (21)). In contrast, the *mepA* operator sequence encompasses two inverted repeats, which allows binding of two MepR dimers (Figure 8, (21)). Inspection of the *mepA* operator sequence reveals that the individual MepR binding sites are essentially adjoining as only a single T–A base pair separates the 3′ end of the 1st site from the 5′ end of the 2nd site (Figure 8). Such close positioning of the binding sites raises the possibility that MepR dimers interact co-operatively to bind to the *mepA* operator. To investigate this possibility, we used the structure of MepR–*mepR* operator site to model the MepR–*mepA* complex (Figure 8). The model demonstrates that the MepR dimers are positioned on nearly opposite faces of the DNA: the 2-fold axes of the two dimers create a dihedral angle of ~130° with the DNA axis. Interestingly, no interactions between MepR dimers could be inferred from the model, which at first glance would imply that MepR binds the *mepA* operator sites independently, in a non-cooperative fashion.

To investigate the thermodynamics of MepR binding to the *mepA* operator and to ascertain whether or not such binding is cooperative, we carried out a series of DNA-binding experiments. MepR binding to the WT *mepA* operator is characterized by a single sharp transition (Figure 9). The stoichiometry of binding ($n = 4$) indicates that two dimers of MepR interact with one DNA duplex,

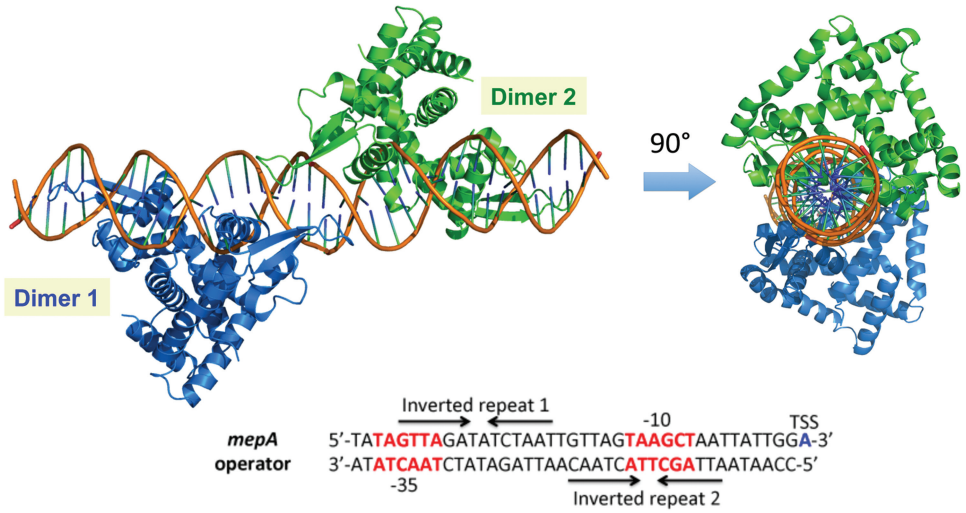

Figure 8. A structure-guided model of MepR bound to the *mepA* operator. The sequence of the *mepA* operator is shown with key promoter elements highlighted (the –35 and –10 boxes are shown in red font and the TSS in blue). Two MepR binding sites (Inverted repeats 1 and 2 corresponding to the 1st and the 2nd sites, respectively) are indicated by arrows.

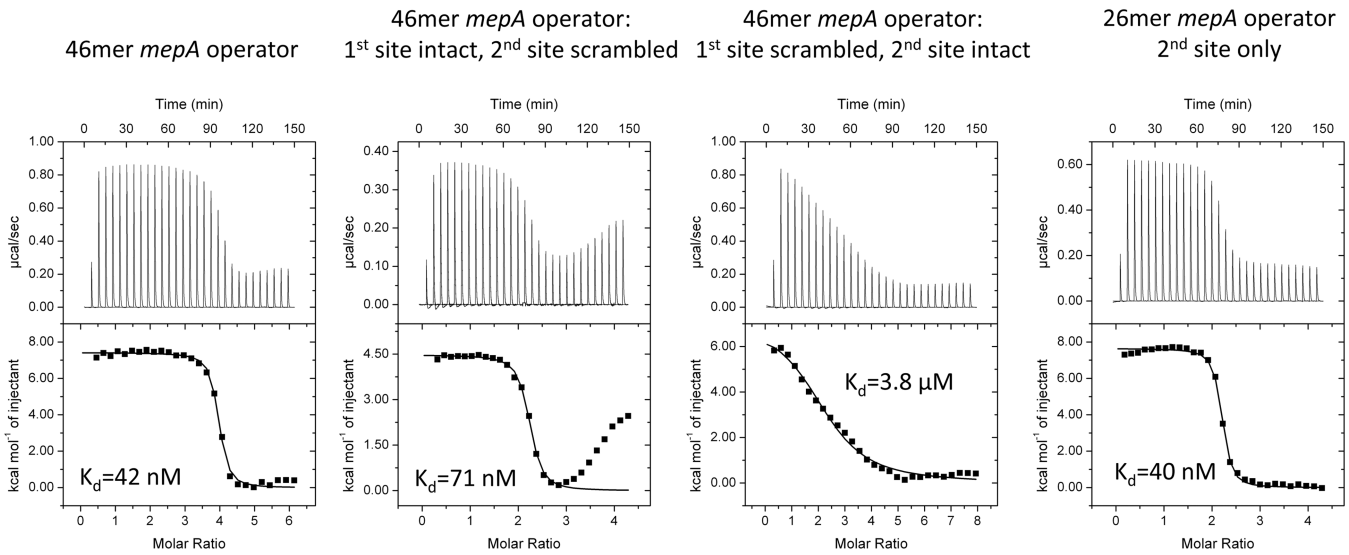

Figure 9. MepR binding to the *mepA* operator. In all experiments, MepR was injected into dsDNA. The DNA concentration was 10 µM, the concentration of MepR in the first experiment was 272 µM, in the second and fourth, 200 µM, and in the third, 370 µM. A stoichiometry of 4 corresponds to 2 dimers of MepR binding to a single dsDNA molecule; a stoichiometry of 2 corresponds to 1 dimer of MepR binding to a single dsDNA molecule. In MepR binding to the 46 bp *mepA* operator in which the 2nd site is scrambled, only the first transition was fit. The sequences of the oligodeoxynucleotides and fitted thermodynamic parameters are provided in Supplementary Table S1.

and the fitted K_d (42 nM) is very close to the K_d of MepR binding to the *mepR* operator (36 nM). The thermogram is more complex for MepR binding to the 46 bp *mepA* operator in which the second binding site was scrambled (Figure 9). Specifically, a second less steep transition appears. We assigned the first transition, which displays a stoichiometry of 2 (1 MepR dimer per 1 DNA duplex) to MepR binding to the intact 1st *mepA* operator site. The second transition likely represents a binding mode observed for the non-specific MepR–DNA complex (Supplementary Figure S4). Interestingly, when the 1st site of the *mepA* operator is destroyed, MepR binding to the intact 2nd *mepA* site is dramatically impaired (K_d = 3.8 µM, n = 2.32, which implies the binding of 1 MepR dimer per one DNA duplex; Figure 9). These data corroborate previous findings of Kumaraswami *et al.* (21), which demonstrated that 5′ end truncations of the *mepA* operator affect MepR binding much more significantly than 3′ end truncations, and proposed that the 5′ end (the 1st) *mepA* site is the primary MepR binding site, while the 3′ end (the 2nd) *mepA* site is a lower affinity, secondary site. Intriguingly, MepR binds to a 26 bp oligodeoxynucleotide encompassing only the second *mepA* site with high affinity (K_d = 40 nM). Thus, our data show that MepR binding to an individual *mepA* 2nd site is indistinguishable from binding to the *mepR*

operator (K_d = 36 nM, Figure 1) or full *mepA* operator (K_d = 42 nM, Figure 9). In the context of the *mepA* operator, however, the binding of MepR to the 1st site is absolutely necessary for high affinity MepR binding to the 2nd site, which implies positive cooperativity in the overall MepR binding to the *mepA* operator. In accord, calculation of the Hill coefficient using the equation $n_H = 2 / [1+(K'/K)^{1/2}]$ (here K′ is the dissociation constant for the intact *mepA* operator, and K is the dissociation constant for the *mepA* operator, in which the 1st site was destroyed), previously employed by Ramos *et al.* (50) to study the cooperativity of binding of the TtgR repressor to its operator site, results in a value of 1.81 supporting our hypothesis that binding shows positive cooperativity. It is likely, that this cooperativity is achieved via DNA-mediated allosteric effects (51), since the model of the MepR–*mepA* operator complex reveals no obvious direct interactions between the two MepR dimers. In order to test this hypothesis and provide a more detailed mechanism of MepR binding to the *mepA* operator, the structure of the MepR–*mepA* complex along with additional biochemical studies are necessary and beyond the scope of the current work.

MarR proteins as global and local transcription regulators

Several MarR proteins act as global regulators thereby affecting the expression of multiple genes. These include SarA from *Staphylococcus aureus* (52) and *Enterococcus faecalis* SlyA (15). On the other hand, some MarR family members regulate only one operon and are often called 'local' regulators. These include *B. subtilis* OhrR (53), *E. coli* MarR (5) and *Pseudomonas aeruginosa* MexR (54). Recently, Dolan *et al.* (33) reported the structure of *S. enterica* global regulator SlyA bound to its cognate DNA and hypothesized that SlyA uses 'indirect readout' to identify its binding sites in numerous operator sites. The authors linked the small number of direct electrostatic base–amino acid side chain interactions between the recognition helix and the major groove (in the SlyA–DNA complex there is only one bidentate hydrogen bond between Arg65 and G_{14}) with the ability of the repressor to recognize multiple operator sites that are variable in length and sequence, and suggested that the lack of direct, specific protein–DNA interactions is a hallmark feature of global virulence regulators. Our structure of MepR in complex with its cognate DNA is inconsistent with this assumption, as there are no hydrogen bonds between the amino acid side chains of the recognition helix and the bases of the GTTAG signature sequence motif (Figures 3 and 5). Rather, the specificity of MepR binding is determined solely by the combination of van der Waals contacts, which are typically not powerful discriminators of DNA bases. In accord, a transcriptome analysis of a *mepR* deletion strain suggests that MepR does not mediate any regulatory effects outside of the *mepRAB* operon (G.W. Kaatz, unpublished data). Therefore, it is likely that the number of the specific protein–DNA major groove hydrogen bonds is not the only criterion to distinguish between local and global regulators in the MarR family. It will be of great interest to discern those factors that would allow a robust prediction of the global or local gene regulatory functions of MarR family members.

The organization of the *mepR* and *mepA* operators

The structure of MepR in complex with *mepR* operator DNA and our DNA-binding data indicate that the GTTA G signature sequence is relatively insensitive to single substitution mutations. It predicts that changing the last guanine to thymine or cytosine, or the third thymine to cytosine, producing the sequences GTTA***T***, GTTA***C*** and GT***C***AG, will not introduce any unfavourable or remove any favourable protein–DNA contacts, and, theoretically, will not impair MepR binding. We have shown that even those mutations in the GTTAG motif, which remove one or two van der Waals contacts, do not affect the affinity of binding significantly (only a 2–3-fold increase in K_d per change, Figure 6). Intriguingly, neither base of the first G–C base pair of the signature motif is contacted by the protein, which is concordant with the lack of conservation of this base pair in the two half-sites of the *mepA* operator (Figure 8). Nevertheless, the GTTAG signature sequence is very well preserved in the *mepR* and *mepA* operators ((19), Figures 3 and 8). The reason for this preservation likely arises from the organization of the *mepR* and *mepA* upstream promoter regions, which have to provide not only high affinity binding sites for MepR, but also maintain optimal –35 and –10 promoter elements for RNA polymerase binding. Thus, in the *mepR* operator, the last guanine nucleotide in the first half-site (GTTA**G,** Figure 3) is also part of the TSS, while in the *mepA* operator, the –35 box, TAGTTA, overlaps significantly with the first half-site of the first *mepA* operator site (**GTTA**G, Figure 8). Thus, the signature sequence is maintained in order to serve as binding sites for two proteins.

Conclusions and perspectives

In this work we provide the molecular details of the transcription repression mechanism of the multidrug efflux operon *mepRA* by the MarR regulator MepR. The structure of MepR in complex with its *mepR* operator DNA reveals that although MepR binds DNA similarly to other MarR proteins, this repressor uses a unique mechanism of binding site recognition based solely on van der Waals interactions between the amino acid residues of the recognition helix and the bases of the major groove encompassing the GTTAG signature sequence. Using the MepR–*mepR* operator structure we modelled a MepR–*mepA* operator complex, which contains two MepR dimer-binding sites. This model suggests that the MepR dimers do not interact with each other in MepR–*mepA* operator complex, and the observed cooperative binding of MepR to the second *mepA* operator site is likely caused by DNA-mediated allosteric effects (51).

We also identified two protein–protein interaction networks that contribute to or stabilize high affinity DNA binding by MepR, which include the His14′:Arg59 pair and Arg10′:His35:Phe108 triad. In both networks, the residues interact via cation–π or π–π interactions. In most crystal structures of MarR proteins bound to cognate effectors (PcaV-protocatechuate (PDB accession code **4fht**)

(30), ST1710-salicylate (PDB accession code **3gez**) (22), MTH313-salicylate (PDB accession code **3bpx**) (24), *S. aureus* MarR-kanamycin (PDB accession code **4emo**), the ligand is located in a cleft surrounded by helices α1, α2 and α5. By inference, we propose that residues Arg10 (α1), His35 (α2) and Phe108 (α5) are involved in MepR-ligand interaction. The substrates of MepR are cationic lipophilic compounds and frequently possess delocalized π-electron systems, which would allow them to interact with these residues and, hence, disrupt this triad. As a consequence of the disruption of this interaction network, the MepR–DNA affinity would be decreased (Supplementary Figure S3A) and lead to the induction of MepR. In accord, residue Arg15 of *S. coelicolor* PcaV, which corresponds to the MepR residue Arg10, is critical for both ligand and DNA binding (30). Structures of MepR-ligand complexes will be necessary to confirm the importance of Arg10 and other residues of the triad and the His14′:Arg59 pair in drug binding and induction.

ACCESSION NUMBERS

Coordinates and structure factors for MepR–DNA complex structures described here have been deposited in the Protein Data Bank with accession codes **4LLL** and **4LLN**.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the beamline scientists at the Advanced Light Source, beamline B.L. 5.0.2 and Dr Nam K. Tonthat for their assistance with X-ray diffraction data collection. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

FUNDING

Duke University Medical School (in part); VA Biomedical Laboratory Research and Development [IO1BX000465]. Funding for open access charge: Personal resources.

Conflict of interest statement. None declared.

REFERENCES

1. Wilkinson,S.P. and Grove,A. (2006) Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Curr. Issues Mol. Biol.*, **8**, 51–62.
2. Ellison,D.W. and Miller,V.L. (2006) Regulation of virulence by members of the MarR/SlyA family. *Curr. Opin. Microbiol.*, **9**, 153–159.
3. Grkovic,S., Brown,M.H. and Skurray,R.A. (2002) Regulation of bacterial drug export systems. *Microbiol. Mol. Biol. Rev.*, **66**, 671–701.
4. Schumacher,M.A. and Brennan,R.G. (2002) Structural mechanisms of multidrug recognition and regulation by bacterial multidrug transcription factors. *Mol. Microbiol.*, **45**, 885–893.
5. Alekshun,M.N. and Levy,S.B. (1997) Regulation of chromosomally mediated multiple antibiotic resistance: the mar regulon. *Antimicrob. Agents. Chemother.*, **41**, 2067–2075.
6. Alekshun,M.N. and Levy,S.B. (1999) The mar regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends Microbiol.*, **7**, 410–413.
7. Alekshun,M.N. and Levy,S.B. (1999) Alteration of the repressor activity of MarR, the negative regulator of the *Escherichia coli* marRAB locus, by multiple chemicals in vitro. *J. Bacteriol.*, **181**, 4669–4672.
8. Alekshun,M.N., Levy,S.B., Mealy,T.R., Seaton,B.A. and Head,J.F. (2001) The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 A resolution. *Nat. Struct. Biol.*, **8**, 710–714.
9. Vazquez-Torres,A. (2012) Redox active thiol sensors of oxidative and nitrosative stress. *Antioxid. Redox. Signal.*, **17**, 1201–1214.
10. Newberry,K.J., Fuangthong,M., Panmanee,W., Mongkolsuk,S. and Brennan,R.G. (2007) Structural mechanism of organic hydroperoxide induction of the transcription regulator OhrR. *Mol. Cell*, **28**, 652–664.
11. Brugarolas,P., Movahedzadeh,F., Wang,Y., Zhang,N., Bartek,I.L., Gao,Y.N., Voskuil,M.I., Franzblau,S.G. and He,C. (2012) The oxidation-sensing regulator (MosR) is a new redox-dependent transcription factor in *Mycobacterium tuberculosis*. *J. Biol. Chem.*, **287**, 37703–37712.
12. Kim,H. and Choe,J. (2013) The X-ray crystal structure of PA1374 from *Pseudomonas aeruginosa*, a putative oxidative-stress sensing transcriptional regulator. *Biochem. Biophys. Res. Commun.*, **431**, 376–381.
13. Quade,N., Mendonca,C., Herbst,K., Heroven,A.K., Ritter,C., Heinz,D.W. and Dersch,P. (2012) Structural basis for intrinsic thermosensing by the master virulence regulator RovA of *Yersinia*. *J. Biol. Chem.*, **287**, 35796–35803.
14. Ellison,D.W., Lawrenz,M.B. and Miller,V.L. (2004) Invasin and beyond: regulation of *Yersinia* virulence by RovA. *Trends Microbiol.*, **12**, 296–300.
15. Michaux,C., Sanguinetti,M., Reffuveille,F., Auffray,Y., Posteraro,B., Gilmore,M.S., Hartke,A. and Giard,J.C. (2011) SlyA is a transcriptional regulator involved in the virulence of *Enterococcus faecalis*. *Infect. Immun.*, **79**, 2638–2645.
16. Chang,Y.M., Jeng,W.Y., Ko,T.P., Yeh,Y.J., Chen,C.K. and Wang,A.H. (2010) Structural study of TcaR and its complexes with multiple antibiotics from *Staphylococcus epidermidis*. *Proc. Natl Acad. Sci. USA*, **107**, 8617–8622.
17. Li,R., Manna,A.C., Dai,S., Cheung,A.L. and Zhang,G. (2003) Crystal structure of the SarS protein from *Staphylococcus aureus*. *J. Bacteriol.*, **185**, 4219–4225.
18. DeMarco,C.E., Cushing,L.A., Frempong-Manso,E., Seo,S.M., Jaravaza,T.A. and Kaatz,G.W. (2007) Efflux-related resistance to norfloxacin, dyes, and biocides in bloodstream isolates of *Staphylococcus aureus*. *Antimicrob. Agents Chemother.*, **51**, 3235–3239.
19. Kaatz,G.W., DeMarco,C.E. and Seo,S.M. (2006) MepR, a repressor of the *Staphylococcus aureus* MATE family multidrug efflux pump MepA, is a substrate-responsive regulatory protein. *Antimicrob. Agents Chemother.*, **50**, 1276–1281.
20. Kaatz,G.W., McAleese,F. and Seo,S.M. (2005) Multidrug resistance in *Staphylococcus aureus* due to overexpression of a novel multidrug and toxin extrusion (MATE) transport protein. *Antimicrob. Agents Chemother.*, **49**, 1857–1864.
21. Kumaraswami,M., Schuman,J.T., Seo,S.M., Kaatz,G.W. and Brennan,R.G. (2009) Structural and biochemical characterization of MepR, a multidrug binding transcription regulator of the *Staphylococcus aureus* multidrug efflux pump MepA. *Nucleic Acids Res.*, **37**, 1211–1224.
22. Kumarevel,T., Tanaka,T., Umehara,T. and Yokoyama,S. (2009) ST1710-DNA complex crystal structure reveals the DNA binding mechanism of the MarR family of regulators. *Nucleic Acids Res.*, **37**, 4723–4735.
23. McAleese,F., Petersen,P., Ruzin,A., Dunman,P.M., Murphy,E., Projan,S.J. and Bradford,P.A. (2005) A novel MATE family efflux pump contributes to the reduced susceptibility of

laboratory-derived *Staphylococcus aureus* mutants to tigecycline. *Antimicrob. Agents Chemother.*, **49**, 1865–1871.
24. Saridakis,V., Shahinas,D., Xu,X. and Christendat,D. (2008) Structural insight on the mechanism of regulation of the MarR family of proteins: high-resolution crystal structure of a transcriptional repressor from *Methanobacterium thermoautotrophicum*. *J. Mol. Biol.*, **377**, 655–667.
25. Sun,F., Ding,Y., Ji,Q., Liang,Z., Deng,X., Wong,C.C., Yi,C., Zhang,L., Xie,S., Alvarez,S. *et al.* (2012) Protein cysteine phosphorylation of SarA/MgrA family transcriptional regulators mediates bacterial virulence and antibiotic resistance. *Proc. Natl Acad. Sci. USA*, **109**, 15461–15466.
26. Wilke,M.S., Heller,M., Creagh,A.L., Haynes,C.A., McIntosh,L.P., Poole,K. and Strynadka,N.C. (2008) The crystal structure of MexR from *Pseudomonas aeruginosa* in complex with its antirepressor ArmR. *Proc. Natl Acad. Sci. USA*, **105**, 14832–14837.
27. Chang,Y.M., Chen,C.K., Ko,T.P., Chang-Chien,M.W. and Wang,A.H. (2013) Structural analysis of the antibiotic-recognition mechanism of MarR proteins. *Acta Crystallogr. D Biol. Crystallogr.*, **69**, 1138–1149.
28. Guerra,A.J., Dann,C.E. III and Giedroc,D.P. (2011) Crystal structure of the zinc-dependent MarR family transcriptional regulator AdcR in the Zn(II)-bound state. *J. Am. Chem. Soc.*, **133**, 19614–19617.
29. Aoki,R., Takeda,T., Omata,T., Ihara,K. and Fujita,Y. (2012) MarR-type transcriptional regulator ChlR activates expression of tetrapyrrole biosynthesis genes in response to low-oxygen conditions in cyanobacteria. *J. Biol. Chem.*, **287**, 13500–13507.
30. Davis,J.R., Brown,B.L., Page,R. and Sello,J.K. (2013) Study of PcaV from *Streptomyces coelicolor* yields new insights into ligand-responsive MarR family transcription factors. *Nucleic Acids Res.*, **41**, 3888–3900.
31. Perera,I.C., Lee,Y.H., Wilkinson,S.P. and Grove,A. (2009) Mechanism for attenuation of DNA binding by MarR family transcriptional regulators by small molecule ligands. *J. Mol. Biol.*, **390**, 1019–1029.
32. Hong,M., Fuangthong,M., Helmann,J.D. and Brennan,R.G. (2005) Structure of an OhrR-ohrA operator complex reveals the DNA binding mechanism of the MarR family. *Mol. Cell*, **20**, 131–141.
33. Dolan,K.T., Duguid,E.M. and He,C. (2011) Crystal structures of SlyA protein, a master virulence regulator of *Salmonella*, in free and DNA-bound states. *J. Biol. Chem.*, **286**, 22178–22185.
34. Stevenson,C.E., Assaad,A., Chandra,G., Le,T.B., Greive,S.J., Bibb,M.J. and Lawson,D.M. (2013) Investigation of DNA sequence recognition by a streptomycete MarR family transcriptional regulator through surface plasmon resonance and X-ray crystallography. *Nucleic Acids Res.*, **41**, 7009–7022.
35. Birukou,I., Tonthat,N.K., Seo,S.M., Schindler,B.D., Kaatz,G.W. and Brennan,R.G. (2013) The molecular mechanisms of allosteric mutations impairing MepR repressor function in multidrug-resistant strains of *Staphylococcus aureus*. *mBio*, **4**, e00528-13.
36. Doublie,S. (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.*, **276**, 523–530.
37. Gasteiger,E., Hoogland,C., Gattiker,A., Duvaud,S., Wilkins,M.R., Appel,R.D. and Bairoch,A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: Walker,J.M. (ed.), *The Proteomics Protocols Handbook*. Humana Press, Totowa, NJ, pp. 571–607.
38. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr. Pt A*, **276**, 307–326.
39. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
40. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
41. Lavery,R., Moakher,M., Maddocks,J.H., Petkeviciute,D. and Zakrzewska,K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.*, **37**, 5917–5929.
42. Bateman,B.T., Donegan,N.P., Jarry,T.M., Palma,M. and Cheung,A.L. (2001) Evaluation of a tetracycline-inducible promoter in *Staphylococcus aureus* in vitro and in vivo and its application in demonstrating the role of sigB in microcolony formation. *Infect. Immun.*, **69**, 7851–7857.
43. Foster,T.J. (1998) Molecular genetic analysis of staphylococcal virulence. *Methods Microbiol.*, **27**, 433–454.
44. Privalov,P.L., Dragan,A.I. and Crane-Robinson,C. (2011) Interpreting protein/DNA interactions: distinguishing specific from non-specific and electrostatic from non-electrostatic components. *Nucleic Acids Res.*, **39**, 2483–2491.
45. Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
46. Saito,K., Akama,H., Yoshihara,E. and Nakae,T. (2003) Mutations affecting DNA-binding activity of the MexR repressor of mexR-mexA-mexB-oprM operon expression. *J. Bacteriol.*, **185**, 6195–6198.
47. Alekshun,M.N., Kim,Y.S. and Levy,S.B. (2000) Mutational analysis of MarR, the negative regulator of marRAB expression in *Escherichia coli*, suggests the presence of two regions required for DNA binding. *Mol. Microbiol.*, **35**, 1394–1404.
48. Liu,Y., Manna,A.C., Pan,C.H., Kriksunov,I.A., Thiel,D.J., Cheung,A.L. and Zhang,G. (2006) Structural and function analyses of the global regulatory protein SarA from *Staphylococcus aureus*. *Proc. Natl Acad. Sci. USA*, **103**, 2392–2397.
49. Schindler,B.D., Seo,S.M., Jacinto,P.L., Kumaraswami,M., Birukou,I., Brennan,R.G. and Kaatz,G.W. (2013) Functional consequences of substitution mutations in MepR, a repressor of the *Staphylococcus aureus* mepA multidrug efflux pump gene. *J. Bacteriol.*, **195**, 3651–3662.
50. Krell,T., Teran,W., Mayorga,O.L., Rivas,G., Jimenez,M., Daniels,C., Molina-Henares,A.J., Martinez-Bueno,M., Gallegos,M.T. and Ramos,J.L. (2007) Optimization of the palindromic order of the TtgR operator enhances binding cooperativity. *J. Mol. Biol.*, **369**, 1188–1199.
51. Kim,S., Brostromer,E., Xing,D., Jin,J., Chong,S., Ge,H., Wang,S., Gu,C., Yang,L., Gao,Y.Q. *et al.* (2013) Probing allostery through DNA. *Science*, **339**, 816–819.
52. Cheung,A.L., Bayer,A.S., Zhang,G., Gresham,H. and Xiong,Y.Q. (2004) Regulation of virulence determinants in vitro and in vivo in *Staphylococcus aureus*. FEMS Immunol. *Med. Microbiol.*, **40**, 1–9.
53. Fuangthong,M., Atichartpongkul,S., Mongkolsuk,S. and Helmann,J.D. (2001) OhrR is a repressor of ohrA, a key organic hydroperoxide resistance determinant in *Bacillus subtilis*. *J. Bacteriol.*, **183**, 4134–4141.
54. Poole,K., Tetro,K., Zhao,Q., Neshat,S., Heinrichs,D.E. and Bianco,N. (1996) Expression of the multidrug resistance operon mexA-mexB-oprM in *Pseudomonas aeruginosa*: mexR encodes a regulator of operon expression. *Antimicrob. Agents Chemother.*, **40**, 2021–2028.