



# Structural basis for DNA recognition by the transcription regulator MetR

Avinash S. Punekar,<sup>a</sup> Jonathan Porter,<sup>b</sup> Stephen B. Carr<sup>a\*</sup> and Simon E. V. Phillips<sup>a\*</sup>

<sup>a</sup>Research Complex at Harwell, Rutherford Appleton Laboratory, Harwell Oxford, Didcot OX11 0FA, England, and  
<sup>b</sup>Pfizer Ltd, Ramsgate Road, Sandwich CT13 9ND, England. \*Correspondence e-mail: stephen.carr@rc-harwell.ac.uk, simon.phillips@rc-harwell.ac.uk

Received 19 February 2016

Accepted 22 April 2016

Edited by W. N. Hunter, University of Dundee, Scotland

**Keywords:** methionine biosynthesis; MetR; LysR-type transcriptional regulator; DNA recognition; helix–turn–helix; molecular replacement; phasing; *phenix.mr\_rosetta*; *HADDOCK*; DNA binding.

**PDB reference:** DNA-binding domain of MetR, 5fo5

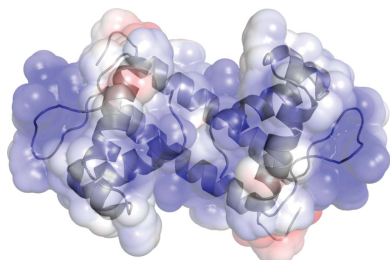
**Supporting information:** this article has supporting information at journals.iucr.org/f

MetR, a LysR-type transcriptional regulator (LTTR), has been extensively studied owing to its role in the control of methionine biosynthesis in proteobacteria. A MetR homodimer binds to a 24-base-pair operator region of the *met* genes and specifically recognizes the interrupted palindromic sequence 5'-TGAA-N<sub>5</sub>-TTCA-3'. Mechanistic details underlying the interaction of MetR with its target DNA at the molecular level remain unknown. In this work, the crystal structure of the DNA-binding domain (DBD) of MetR was determined at 2.16 Å resolution. MetR-DBD adopts a winged-helix–turn–helix (wHTH) motif and shares significant fold similarity with the DBD of the LTTR protein BenM. Furthermore, a data-driven macromolecular-docking strategy was used to model the structure of MetR-DBD bound to DNA, which revealed that a bent conformation of DNA is required for the recognition helix  $\alpha 3$  and the wing loop of the wHTH motif to interact with the major and minor grooves, respectively. Comparison of the MetR-DBD–DNA complex with the crystal structures of other LTTR-DBD–DNA complexes revealed residues that may confer operator-sequence binding specificity for MetR. Taken together, the results show that MetR-DBD uses a combination of direct base-specific interactions and indirect shape recognition of the promoter to regulate the transcription of *met* genes.

## 1. Introduction

Methionine biosynthesis in *Escherichia coli* has been extensively studied and involves two essential pathways. The first is for homocysteine biosynthesis (catalyzed by the enzymes MetA, MetB and MetC) and the second is for the biosynthesis of *N*<sup>5</sup>-methyltetrahydrofolate (catalyzed by the enzymes GlyA and MetF); they merge during the transfer of a methyl group from *N*<sup>5</sup>-methyltetrahydrofolate to homocysteine (catalyzed by either the cobalamin-dependent MetH or the cobalamin-independent MetE), producing methionine (reviewed in Weissbach & Brot, 1991; Ferla & Patrick, 2014). Intracellular methionine homeostasis is maintained by the tight regulation of the genes encoding enzymes in the methionine metabolic pathway. In *Escherichia coli*, methionine-metabolism genes are controlled by two specific transcription factors: the *S*-adenosylmethionine-dependent repressor MetJ and the homocysteine-dependent activator MetR (Ferla & Patrick, 2014; Leyn *et al.*, 2014).

MetR was first discovered as a transactivator of *metE* and *metH* expression (Urbanowski *et al.*, 1987), with subsequent investigations showing that it is also a transcriptional regulator of other methionine-biosynthesis genes: *metA* and *metF* in *Salmonella typhimurium* (Mares *et al.*, 1992; Cowan *et al.*, 1993) and *glyA* in *E. coli* (Lorenz & Stauffer, 1995). Expression of the *metR* gene is negatively regulated by its own gene



product MetR and by the repressor MetJ (Urbanowski & Stauffer, 1987). Interestingly, MetR also regulates the transcription of non-*met* genes such as *hmp*, which encodes flavohaemoglobin to protect bacteria from nitric oxide stress (Flatley *et al.*, 2005), and virulence-associated genes in *Pseudomonas aeruginosa* (Yeung *et al.*, 2009) and *Vibrio cholerae* (Bogard *et al.*, 2012).

The *E. coli* MetR protein (UniProt ID P0A9F9;  $M_r = 35\,629$ ) is a homodimer in solution (Maxon *et al.*, 1990) that binds to a 24-base-pair DNA sequence in the operator regions of the *met* genes, specifically recognizing an 8 bp interrupted palindromic repeat with consensus 5'-TGAA-N<sub>5</sub>-TTCA-3' (Maxon *et al.*, 1989; Cai, Maxon *et al.*, 1989; Leyn *et al.*, 2014). Similar MetR binding sites have also been found in the transcription-control regions of the *met* genes of *Salmonella typhimurium* (Urbanowski & Stauffer, 1989b; Byerly *et al.*, 1991). MetR-mediated regulation of *metR*, *metE* and *metH* genes is modulated by the co-inducer homocysteine (Urbanowski & Stauffer, 1987, 1989a; Cai, Redfield *et al.*, 1989), which increases the affinity of MetR for the target DNA sequence, for example in the transcriptional control regions of the *glyA* gene of *E. coli* (Lorenz & Stauffer, 1995) and the *metE* gene of *Streptococcus mutans* (Sperandio *et al.*, 2007).

MetR is a member of the well characterized LysR-type transcriptional regulator (LTTR) protein family, which is one of the largest in prokaryotes (Schell, 1993). LTTRs are structurally conserved proteins that bind a co-inducer to function either as activators or repressors of gene expression in various biological pathways. A typical LTTR comprises an N-terminal DNA-binding domain (DBD) connected *via* a linker helix (LH) to a C-terminal effector-binding domain (EBD). The DBD consists of a classical winged-helix–turn–helix (wHTH) motif, while the EBD is structurally similar to periplasmic binding proteins, consisting of two distinct  $\alpha/\beta$  subdomains linked by two interdomain strands, producing a co-inducer-binding hinge region/cleft between the two  $\alpha/\beta$  subdomains. LTTRs generally adopt a homo-oligomeric structure to recognize the consensus DNA sequence T-N<sub>11</sub>-A during transcriptional regulation (Schell, 1993; Maddocks & Oyston, 2008). Several crystal structures of diverse LTTRs have been characterized, including full-length CbnR (Muraoka *et al.*, 2003), CrgA (Sainsbury *et al.*, 2012), ArgP (Zhou *et al.*, 2010), BenM (Ruangprasert *et al.*, 2010), TsaR (Monferrer *et al.*, 2010), AphB (Taylor *et al.*, 2012) and OxyR (Jo *et al.*, 2015), the C-terminal EBDs of BenM (Craven *et al.*, 2009) and the MetR-like NMB2055 (Sainsbury *et al.*, 2012), and the N-terminal DBD-LH of BenM (Alanazi *et al.*, 2013). Currently, there is no structure of a full-length LTTR in complex with DNA; however, the crystal structure of the complex between BenM-DBD-LH and DNA (BenM-DBD-LH–DNA) has been solved (Alanazi *et al.*, 2013).

In this study, we present the crystal structure of the DBD-LH of MetR (residues 1–92) from *E. coli*. The MetR-DBD-LH structure was solved by a combination of molecular replacement (MR) and *Rosetta* rebuilding (*phenix.mr\_rosetta*) as implemented in *PHENIX* (Adams *et al.*, 2010). Additionally, we have used the BenM-DBD-LH–DNA complex as a refer-

Table 1

Macromolecule-production information for MetR-DBD-LH.

UniProt identifier	P0A9F9
Source organism	<i>E. coli</i>
Expression host	<i>E. coli</i> BL21( $\lambda$ DE3)
Complete amino-acid sequence of the construct produced	MIEVKHLKTLQALRNCGLAAAAATLHQTQSALS-HQFSDLEQRLGFRLFVRKSQLRFTPPQGEILL-QLANQVLPQISQALQACNEPQQTRLR

ence structure in molecular-docking simulation studies to build a model of the MetR-DBD-LH–DNA complex, giving insight into the recognition of DNA by MetR.

## 2. Materials and methods

### 2.1. Protein expression and purification

The MetR-91 plasmid encoding MetR-DBD-LH (Table 1) expression was a gift from Mark L. Urbanowski and Professor George V. Stauffer (Department of Microbiology, University of Iowa, USA). MetR-DBD-LH was recombinantly expressed in *E. coli* BL21( $\lambda$ DE3) pre-transformed with the *lac* repressor plasmid pMS421. The cells were grown in lysogeny broth (LB) culture medium supplemented with ampicillin ( $200\,\mu\text{g l}^{-1}$ ) and spectinomycin ( $50\,\mu\text{g l}^{-1}$ ) at 310 K to an OD<sub>600</sub> of  $\sim 0.6$ . Expression was induced by adding isopropyl  $\beta$ -D-1-thiogalactopyranoside to a final concentration of 1 mM and continued incubation at 310 K for an additional 4 h. Cells were harvested by centrifugation, resuspended in 10 ml buffer *A* [100 mM Tris–HCl pH 7.4, 100 mM KCl, 2 mM EDTA, 10% (w/v) glycerol, 2 mM DTT] and lysed by sonication on ice. The cell lysate was cleared by centrifugation at 43 000g for 30 min at 277 K. All subsequent purification steps were performed at 277 K. Solid ammonium sulfate was slowly added to the cell lysate to an initial concentration of 25% and stirred for 2 h. The precipitate was removed by centrifugation at 12 000g for 10 min. Additional ammonium sulfate was added to the supernatant to achieve a final saturation of 35% with stirring continued for a further 15 min. The resulting precipitate was pelleted by centrifugation at 12 000g for 10 min. The precipitate was dissolved in 4 ml buffer *B* [25 mM Tris–acetate pH 6.4, 50 mM KCl, 0.5 mM Mg(CH<sub>3</sub>COO)<sub>2</sub>·4H<sub>2</sub>O, 1 mM DTT] and dialyzed using Spectra/Por dialysis tubing (3500 Da molecular-weight cutoff) against buffer *B* overnight. The dialyzed protein sample was centrifuged at 12 000g for 10 min, filtered through a 0.2  $\mu\text{m}$  syringe filter and loaded onto a CM-52 cation-exchange column (1  $\times$  20 cm and 10 ml bed volume) pre-equilibrated with buffer *B* at a flow rate of 0.3 ml min<sup>−1</sup>. The column was washed with buffer *B* at the same flow rate until the *A*<sub>280</sub> reached the baseline. MetR-DBD-LH was eluted using four column volumes (40 ml) of buffer *B* with a linear salt gradient (50–350 mM KCl) and *A*<sub>280</sub> was continuously monitored. Peak fractions were analyzed on 12% Tricine polyacrylamide gels and pure fractions of MetR-DBD-LH were pooled, dialyzed against the storage buffer (10 mM Tris–HCl pH 7.2, 50 mM KCl, 1 mM MgCl<sub>2</sub>, 10 mM DTT), concentrated to 3 mg ml<sup>−1</sup>

Table 2

X-ray crystallographic statistics for MetR-DBD-LH.

Values in parentheses are for the outer shell.

Data collection and processing	
Diffraction source	PX7.2, SRS, Daresbury Laboratory
Wavelength (Å)	1.488
Temperature (K)	300
Detector	MAR Research image plate
Crystal-to-detector distance (mm)	175
Rotation range per image (°)	2
Total rotation range (°)	180
Exposure time per image (s)	120
Space group	C2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	27.76, 55.12, 49.99
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 98.82, 90
Resolution range (Å)	8.92–2.16 (2.24–2.16)
Total No. of reflections	31780 (1749)
No. of unique reflections	3976 (379)
Completeness (%)	98 (99)
Multiplicity	8 (4.6)
$\langle I/\sigma(I) \rangle$	21.28 (11.03)
$R_{\text{f.i.m.}}$	0.125 (0.204)
$CC_{1/2}^\dagger$	0.996 (0.955)
Overall <i>B</i> factor from Wilson plot (Å <sup>2</sup> )	15.8
Matthews coefficient (Å <sup>3</sup> Da <sup>−1</sup> )	1.83
Solvent content (%)	33
Molecules per asymmetric unit	1
Refinement	
No. of reflections, working set	3973 (378)
No. of reflections, test set	398 (38)
Final $R_{\text{cryst}}^\ddagger$	0.191 (0.186)
Final $R_{\text{free}}^\S$	0.235 (0.202)
No. of non-H atoms	
Protein	710
Ion	1
Ligand	4
Water	38
Total	753
R.m.s. deviations	
Bonds (Å)	0.002
Angles (°)	0.53
Average <i>B</i> factors (Å <sup>2</sup> )	
Protein	23.3
Ligand	35.4
Water	25.9
Ramachandran plot	
Favoured regions (%)	98
Additionally allowed (%)	2
Outliers	0
<i>MolProbity</i> statistics	
Overall score	1.83
Clashscore	6.15
Rotamer outliers (%)	3.8
PDB entry	5fo5

$^\dagger$   $CC_{1/2}$  is Pearson's correlation coefficient between random half data sets.  $^\ddagger$   $R_{\text{cryst}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$ .  $^\S$   $R_{\text{free}}$  is calculated for 10% of the total reflections that were randomly chosen and excluded from all refinement cycles.

as estimated by the Lowry method and stored at 193 K. The final yield of MetR-DBD-LH was 10 mg per litre of culture.

## 2.2. Crystallization, data collection and processing

The MetR-DBD-LH crystals were grown by mixing 1  $\mu$ l protein solution (3 mg ml<sup>−1</sup> in 10 mM Tris–HCl pH 7.2, 50 mM KCl, 1 mM MgCl<sub>2</sub>, 10 mM DTT) and 1  $\mu$ l precipitant solution (50 mM HEPES pH 7.2, 1.2 M ammonium sulfate, 200 mM KCl, 50 mM MgCl<sub>2</sub>) using sitting-drop vapour diffusion at 298 K. Large crystals (250 × 100 × 40  $\mu$ m) were

observed within one week. Flash-cooling of MetR-DBD-LH crystals in the presence of different cryoprotectants (PEG 400, glycerol, ethanediol and saturated xylitol solution) resulted in a large increase in mosaicity and a loss of diffraction. Therefore, single crystals were mounted in glass capillaries and X-ray diffraction data were collected at room temperature at station 7.2 of the Synchrotron Radiation Source at Daresbury Laboratory, England using a MAR Research image-plate detector system. Partial native data sets were collected from multiple crystals. Indexing and integration was performed using *XDS* (Kabsch, 2010). *POINTLESS* (Evans, 2011) was used to confirm the Laue group and *XSCALE* (Kabsch, 2010) was used to scale and merge the partial data sets. X-ray data-collection and processing statistics are summarized in Table 2.

## 2.3. Structure solution and refinement

Attempts to solve the structure of MetR-DBD-LH using classical MR were unsuccessful and structure determination required a combination of automated molecular replacement (AutoMR) and *Rosetta* rebuilding using the *phenix.mr\_rosetta* module (DiMaio *et al.*, 2011; Terwilliger *et al.*, 2012; DiMaio, 2013) as implemented in *PHENIX* (Adams *et al.*, 2010). The amino-acid sequence of MetR-DBD-LH (Table 1) was submitted to (i) the *HHPred* server (Söding *et al.*, 2005) for global sequence alignment and (ii) the *Robetta* fragment server (Kim *et al.*, 2004). The alignment (.hhr) file from the *HHPred* server, the 3-mer and 9-mer fragment files from the *Robetta* fragment server, the X-ray diffraction data set and the MetR-DBD-LH amino-acid sequence file were input into *phenix.mr\_rosetta*. The first model in the alignment file (PDB entry 3fxq; the crystal structure of the LTTR TsaR; Monferrer *et al.*, 2010) was automatically selected as a search model for AutoMR. For each best MR solution, 100 *Rosetta* models were rebuilt and rescored. The highest scoring *Rosetta* solutions were processed using *phenix.autobuild* to produce a new best model for another round of AutoMR and rebuilding using *phenix.mr\_rosetta*. The final best model from *phenix.mr\_rosetta* and the density-modified phases were used for iterative cycles of maximum-likelihood refinement in *phenix.refine* (Adams *et al.*, 2010) and manual rebuilding in *Coot* (Emsley *et al.*, 2010). The final refined model was validated using *MolProbity* (Chen *et al.*, 2010). Structure-refinement and validation statistics for MetR-DBD-LH are presented in Table 2.

## 2.4. Sequence and structure analysis

Multiple sequence alignment was performed using *Clustal Omega* (Sievers *et al.*, 2011). *PyMOL* (Schrödinger) and *CCP4mg* (McNicholas *et al.*, 2011) were used for structure analysis and for producing figures. *APBS* (Baker *et al.*, 2001) and *PDB2PQR* (Dolinsky *et al.*, 2007) in *PyMOL* were used to calculate the surface electrostatic potential map. *Cealign* in *PyMOL* and the *POSA* server (Li *et al.*, 2014) were used for structure alignment and to calculate the root-mean-square deviation (r.m.s.d.).

Table 3

HADDOCK parameters and settings for docking.

All other parameters and settings in HADDOCK were kept with default values.

Interface/access level	Guru
MetR-DBD-LH	
Active residues (coloured red)	MIEVKHLKTLQALRNCGLAAAAATLHQTQS- ALSHQFSDLEQLRGFLFVRKSQPLRFTP- QGEILLQLANQVLPQISQALQACNEPQQT- RLR
Passive residues	6.5 Å around active residues
Histidine protonation state	Automatic
Semi-flexible segments	Automatic
MetR DNA-binding site	
Active residues (coloured red)	5'-GATCATGAAAGTCCCTTCACTTCGGC-3' 3'-CTAGTACTTTCAGGAAGTGAAGCCG-5'
Passive residues	6.5 Å around active residues
Semi-flexible segments	Automatic
Sampling parameters	
No. of structures for rigid-body docking	2000
No. of trials for rigid-body minimization	5
No. of structures for semi-flexible refinement	300
No. of structures for explicit solvent refinement	200
$\epsilon$	78
Advanced sampling parameters	
Initial temperature for third TAD cooling step with fully flexible interface	300
Factor for time step in TAD	4
No. of steps for 300 K phase	750

## 2.5. Modelling of the MetR-DBD-LH–DNA complex

Modelling of the MetR-DBD-LH–DNA complex was performed with the HADDOCK 2.2 webserver (Dominguez *et al.*, 2003; de Vries *et al.*, 2010; Wassenaar *et al.*, 2012). A pre-bent B-form DNA fragment containing the MetR DNA-binding site (Table 3) was built using the DNA coordinates from the crystal structure of the BenM-DBD-LH–DNA complex (PDB entry 4iht; Alanazi *et al.*, 2013) as a reference in 3D-DART (van Dijk & Bonvin, 2009). The coordinate files of a MetR-DBD-LH dimer and the pre-bent MetR binding-site DNA were used as the starting structures for docking simulations (docking parameters are described in Table 3). HADDOCK requires a set of ambiguous interaction restraints (AIRs) defining the ‘active’ and ‘passive’ residues at the macromolecular interaction interface. The MetR-DBD-LH structure was superposed onto the BenM-DBD-LH–DNA structure to identify potential active residues in the HTH motif of MetR-DBD (Table 3) that may interact with DNA. Active residues in the MetR DNA-binding site have been identified previously from DNase I protection footprinting experiments (Cai, Maxon *et al.*, 1989) and verified by site-directed single base-pair mutations (Byerly *et al.*, 1991). Passive residues in MetR-DBD-LH and the MetR DNA-binding site were defined as all solvent-accessible neighbours of active residues. The AIRs implemented in our docking simulations do not contain any residue-specific interactions at the MetR-DBD–DNA binding interface. The semi-flexible regions in the interacting macromolecules were defined automatically by considering all residues within 5 Å of

Table 4

MetR-DBD-LH structure homologues.

Protein	PDB code	DALI Z-score	C $\alpha$ atoms aligned	R.m.s.d. (Å)	Sequence identity (%)	Phaser MR solution†		
						RFZ	TFZ	LLG
BenM	4iht	13.6	86	1.5	31	5.1	4.7	26.9
OxyR	4x6g	12.7	88	1.9	26	5.4	4.3	27.5
CrgA	3bhg	12.6	88	2.0	24	N	N	N
CbnR	1iz1	12.0	87	2.1	28	N	N	N
TsaR	3fxq	11.8	85	2.0	27	4.0	3.7	15.3
AphB	3t1b	11.6	87	2.7	18	4.4	3.5	17.4
ArgP	3isp	11.5	87	2.4	36	4.4	3.5	16.9

† N, no MR solutions found.

another molecule. 2000 structures with various conformations were generated in the rigid-body docking step, followed by semi-flexible simulated annealing of the 300 solutions with the lowest intermolecular energy and final refinement in explicit solvent. The resulting 200 best structures were clustered using an interface-ligand root-mean-square deviation (i-l-r.m.s.d.) cutoff of 7.5 Å. The top-ranked cluster with the best HADDOCK score contained 101 structures. The best four structures (lowest energy structures with an r.m.s.d. of 0.8 Å over backbone atoms) from this cluster were analyzed using HADDOCK 2.2 and represent the model of the MetR-DBD-LH–DNA complex.

## 3. Results and discussion

### 3.1. Structure determination of MetR-DBD-LH

Classical MR was unable to solve the structure of MetR-DBD-LH. A variety of search models, the DBDs of LTTR homologues with significant sequence similarity, were tested using Phaser (McCoy, 2007) but no solutions were found (see Table 4 for detailed analysis). The structure of MetR-DBD-LH was eventually solved using phenix.mr\_rosetta (DiMaio *et al.*, 2011; Terwilliger *et al.*, 2012; DiMaio, 2013) and refined at

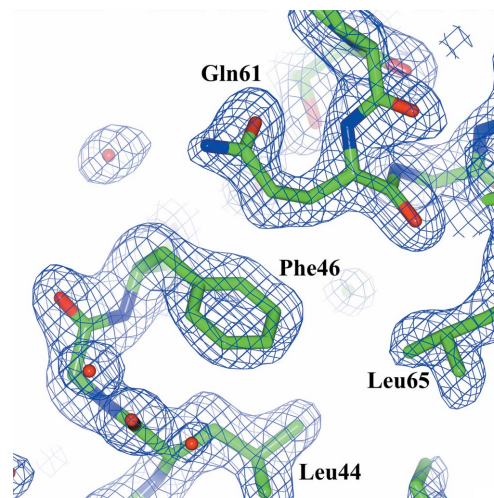


Figure 1  
The electron-density map of MetR-DBD-LH.



2.16 Å resolution (Fig. 1). For successful phasing by MR, the r.m.s.d. between the search model and target structures is suggested to be in the range 1.5–2 Å (Terwilliger *et al.*, 2012). Usually, the r.m.s.d. decreases when the search models have high sequence identity to the target structures (Evans & McCoy, 2008). A *BLAST* search against the Protein Data Bank (PDB) using the MetR-DBD-LH sequence suggested that the crystal structure of the LTTR protein ArgP (PDB entry 3isp; Zhou *et al.*, 2010) would be the best search model, with 36% sequence identity. A pairwise comparison of the final MetR-DBD-LH and ArgP-DBD-LH structures gave an r.m.s.d. of 2.4 Å for 87 aligned C $\alpha$  atoms, suggesting why MR failed when ArgP-DBD-LH and fragments thereof were used as a search model. MR also failed with BenM-DBD-LH (PDB entry 4iht; Alanazi *et al.*, 2013) as a search model, which had a pairwise r.m.s.d. of 1.5 Å with the final MetR-DBD-LH structure for 86 aligned C $\alpha$  atoms. Possible reasons for the failure of MR using this better search model include the poor completeness (89%) of the MetR-DBD-LH X-ray data in the

low-resolution shell and structural differences between the search model and MetR-DBD-LH. Structural comparison by superimposition of the LTTR-DBDs shows dynamic mobility in the wing region and the C-terminal end of LH (Fig. 2*b*), which would impede MR. Removal of the flexible regions in an attempt to improve the likelihood of successful MR phasing also failed, possibly owing to the fact that HTH motif alone comprises about only 45% of the overall scattering mass of the MetR-DBD-LH asymmetric unit.

Successful phasing by MR required improvement of the search-model quality (optimization of side-chain and backbone coordinates) through multiple rounds of model rebuilding and electron-density map-guided refinement using *phenix.mr\_rosetta*. The final best model from *phenix.mr\_rosetta* ( $R_{\text{cryst}} = 0.229$  and  $R_{\text{free}} = 0.319$ ) and the density-modified phases were used for further cycles of maximum-likelihood refinement in *phenix.refine* (Adams *et al.*, 2010) and manual rebuilding in *Coot* (Emsley *et al.*, 2010), producing a final model with  $R_{\text{cryst}} = 0.191$  and  $R_{\text{free}} = 0.235$  (Table 2). The

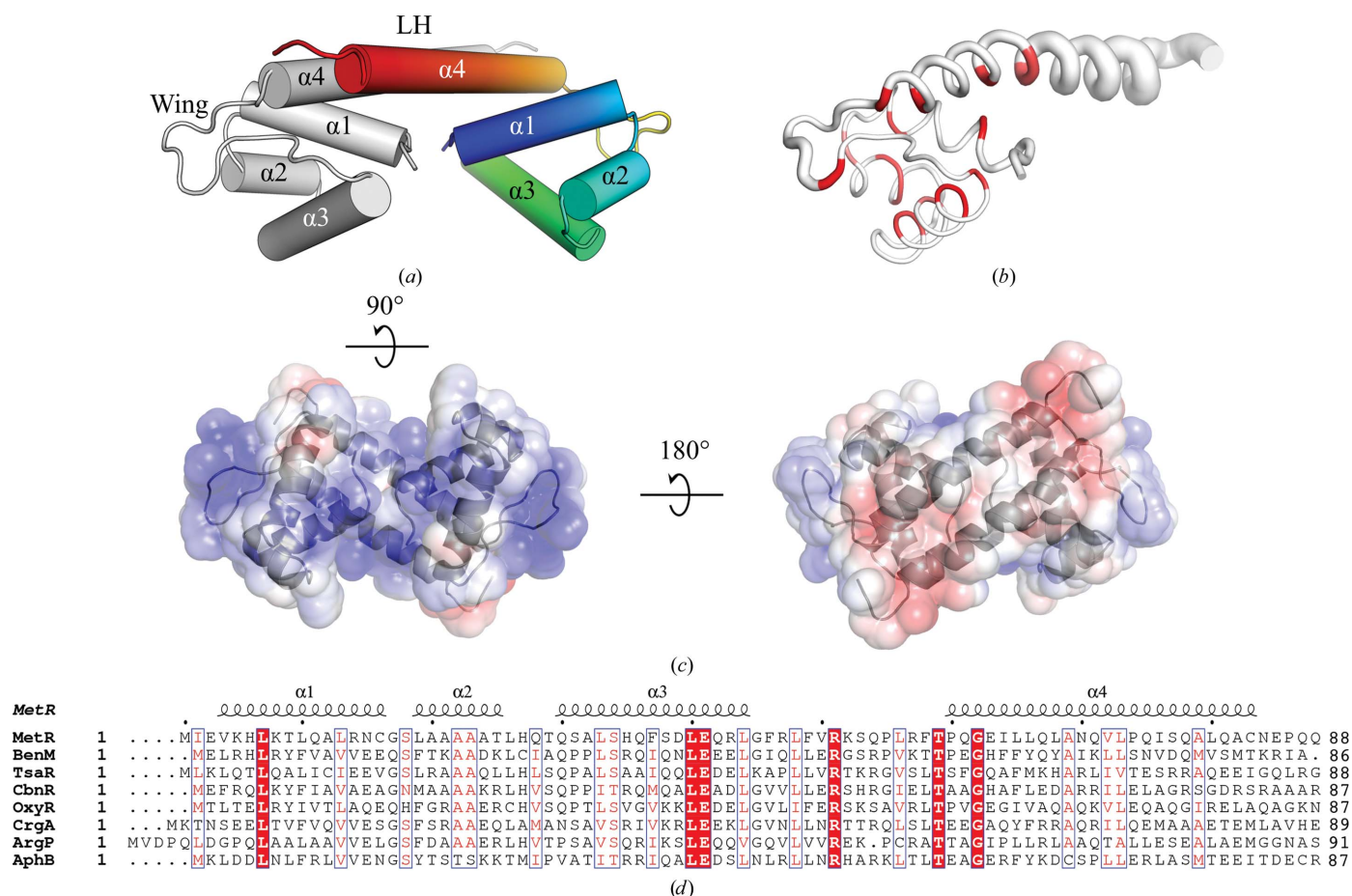


Figure 2

(a) Structure of MetR-DBD-LH in cartoon representation with helices shown as solid cylinders. One molecule is coloured in rainbow mode (blue to red from the N-terminus to the C-terminus) and the symmetry-related molecule is coloured grey. (b) MetR-DBD-LH in tube representation coloured according to the level of sequence conservation shown in (d), from white (low identity) to red (100% identity). The tube radius is proportional to the mean C $\alpha$  r.m.s.d. deviation between MetR-DBD-LH and the structural homologues in (d). The image was created with *ENDscript* 2.0 (Robert & Gouet, 2014). (c) Electrostatic surface potential of MetR-DBD-LH. The colour spectrum ranges from deep red (−3kT) to deep blue (+3kT). (d) Structure-based sequence alignment of MetR with BenM (PDB entry 4iht), TsaR (PDB entry 3fxq), CbnR (PDB entry 1iz1; Muraoka *et al.*, 2003), OxyR (PDB entry 4x6g; Jo *et al.*, 2015), CrgA (PDB entry 3hhg; Sainsbury *et al.*, 2012), ArgP (PDB entry 3isp; Zhou *et al.*, 2010) and AphB (PDB entry 3t1b; Taylor *et al.*, 2012). The image was created with *ESPrpt* 3.0 (Robert & Gouet, 2014). Strictly conserved residues are shown in white on a red background and conservative substitutions are shown in red on a white background. The secondary structure of MetR is shown above the alignment.

overall  $R_{\text{cryst}}$  and  $R_{\text{free}}$  values are slightly higher than the  $R$  factors in the highest resolution shell, possibly owing to the contribution of the high ammonium sulfate concentration from the bulk solvent, as reported previously (Phillips, 1980). Of the 92 residues present in the construct (Table 1), amino acids 1–88 were clearly resolved, but the last four residues had no visible electron density, probably owing to the flexibility of the C-terminus. The electron density for the side chains of Arg43, Arg51 and Lys52 was also poor, indicating that they are disordered.

### 3.2. Structure of MetR-DBD-LH

The MetR-DBD-LH crystals belonged to the monoclinic space group  $C2$  and had a solvent content of 33%. The structure of MetR-DBD-LH contains a monomer in the asymmetric unit. Rotation about the crystallographic twofold rotation axis produces the biologically active MetR-DBD-LH homodimer (Fig. 2*a*) with an antiparallel arrangement of the two protomers. The DBD (residues 1–59) adopts a winged-helix–turn–helix (wHTH) fold consisting of helix  $\alpha 1$  (residues

4–16), a turn (residues 17–18), helix  $\alpha 2$  (residues 19–25), a second turn (residues 26–29), helix  $\alpha 3$  (residues 30–44) and a large flexible loop or ‘wing’ (residues 45–59). The DBD is followed by a long LH (residues 60–85). Nonpolar residues on the LHs make extensive hydrophobic contacts that stabilize the dimerization interface, with a buried surface of 1261 Å<sup>2</sup>, similar to the surface area buried at the dimer interface of various other LTTRs (Ruangprasert *et al.*, 2010). In MetR-DBD, the N-terminal end of helix  $\alpha 1$ , the outer surface of helix  $\alpha 3$  and the wing contain both positively charged and polar amino acids, suggesting that this is the DNA-binding surface (Fig. 2*c*, left view). The hydrophobic core of the DBD consists of a cluster of ten leucine and three phenylalanine residues. Interestingly, four leucine residues (Leu19, Leu26, Leu33 and Leu40) of the hydrophobic core are spaced seven residues apart in the primary sequence, resembling a heptad repeat, which has previously led to the misassignment of this region as a leucine-zipper motif (Maxon *et al.*, 1990). *In vivo* site-directed mutagenesis experiments on MetR had previously shown that a single mutant L33P or a double mutant L26S or L40Q lead to loss of the MetR activity (Maxon *et al.*, 1990).

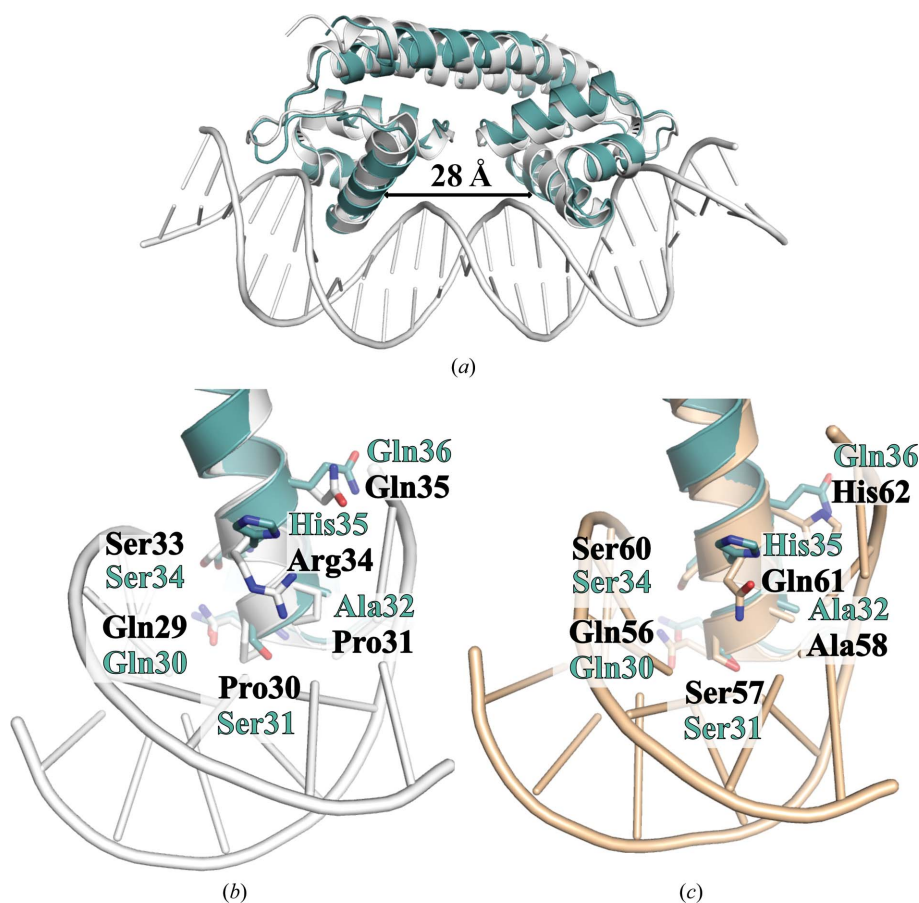


Figure 3

Comparison between MetR-DBD and DNA-bound DBDs of BenM and NoIR. The MetR-DBD-LH-DNA model (teal) was aligned with the crystal structures of the BenM-DBD-LH-DNA complex (PDB entry 4iht, grey) and the NoIR-DBD-LH-DNA complex (PDB entry 4on0, wheat; Lee *et al.*, 2014). (a) The MetR-DBD-LH-DNA model superposes on the BenM-DBD-LH-DNA complex with an r.m.s.d. of 1.9 Å for 168 aligned C $\alpha$  atoms. For clarity, only the DNA chains of the BenM-DBD-LH-DNA complex are shown as cartoons. The arrow indicates the distance between two recognition helices (helix  $\alpha 3$ ) in the compared structures. MetR-DBD shows similarities and differences in the recognition-helix interactions with DNA when compared with the recognition helices of (b) BenM-DBD and (c) NoIR-DBD. For clarity, only the DNA chains of NoIR-DBD-LH-DNA complex are shown as cartoons. The amino-acid side chains are shown in stick representation and are coloured by atom type.

Our structure shows that Leu26, Leu33 and Leu40 are part of a strong hydrophobic interaction network that maintains the overall architecture and stability of the HTH motif.

A database search using the *DALI* server (Holm & Rosenström, 2010) identified the DBDs of LTTRs such as BenM, TsaR, CbnR, OxyR, CrgA, AphB and ArgP as the closest structural homologues of MetR-DBD-LH (Table 4). This is expected given the sequence similarity between MetR-DBD-LH and its structural homologues (Fig. 2*d*). A comparison of the structures of LTTR-DBDs shows that immediately following the HTH motifs is the wing connecting helix  $\alpha 3$  and LH. Most LTTR-DBDs, except for BenM-DBD, contain a  $\beta$ -hairpin in the wing region. MetR-DBD also lacks the  $\beta$ -hairpin in the wing region, which is oriented similar to that in BenM-DBD. The crystal structure of the BenM-DBD-LH-DNA complex (PDB entry 4iht) was the highest *DALI* hit (*Z*-score = 13.6) for MetR-DBD-LH. The high structural similarity between MetR-DBD-LH and BenM-DBD-LH allowed us to use PDB entry 4iht as a reference structure to model the MetR-DBD-LH-DNA complex.

### 3.3. A model of the MetR-DBD-LH-DNA complex

In WHTH motif-containing transcription factors, helix  $\alpha 3$  serves as the recognition helix that inserts into the major groove of the substrate DNA (Aravind *et al.*, 2005). In LTTRs, the length of the LH determines the relative position of the recognition helices (Schell, 1993), and the crystal structure of the BenM-DBD-LH-DNA complex (PDB entry 4iht) is currently the only example of an LTTR bound to DNA (Alanazi *et al.*, 2013). The BenM dimer binds to a 25 bp DNA sequence in the operator region and recognize the consensus DNA sequence ATAC-N<sub>7</sub>-GTAT. In BenM, the recognition helices in DBD are separated by 28 Å (Fig. 3*a*) and interaction of BenM with its DNA substrate does not alter the backbone conformation of DBD but induces a 45–47° bend in the DNA,

resulting in a subtle structural rearrangement of the DNA backbone to accommodate the recognition helices in the adjacent major grooves. DNase I protection footprinting experiments have shown that binding of the MetR dimer to its operator site protects 24 bp of DNA and recognizes the consensus DNA sequence 5'-TGAA-N<sub>5</sub>-TTCA-3' (Maxon *et al.*, 1989; Cai, Maxon *et al.*, 1989; Leyn *et al.*, 2014). The high structural similarity between the MetR-DBD-LH and BenM-DBD-LH dimers (Fig. 3*a*) and the similarity of their substrate size suggest that the MetR-DBD-LH dimer would also induce a similar bend angle on binding to DNA. However, the variation in the base preferences in the half-sites of the BenM and MetR DNA-binding sequences suggests that MetR and BenM would utilize different amino acid–base contacts in their respective substrate major and minor grooves for sequence-specific DNA recognition.

To understand the molecular basis for DNA recognition by MetR, we modelled the MetR-DBD-LH-DNA complex using molecular-docking simulations in the data-driven program *HADDOCK*. The structure of the MetR-DBD-LH in the final docked model remained essentially identical to the starting model (r.m.s.d. of 0.6 Å over 176 aligned C $\alpha$  atoms). The two recognition helices in MetR-DBD-LH are separated by 28 Å and our docked model shows that the two recognition helices of the MetR-DBD-LH dimer occupy consecutive major grooves of the DNA containing the palindromic recognition half-sites (TGAA and TTCA), while the wings make contacts in the minor grooves.

The MetR-DBD homodimer makes a total of ten base and 12 phosphate contacts with the DNA (Fig. 4). The molecular surface of MetR-DBD at the DNA-binding interface of each monomer contain clusters of positively charged amino acids (Lys5 and Lys8 in helix  $\alpha 1$ , His27 in the loop after helix  $\alpha 2$ , His35 and Arg43 in helix  $\alpha 3$  and Arg47, Arg51, Lys52 and Arg57 in the wing region; Fig. 2*c*, left view) for making extensive interactions with the negatively charged phosphate

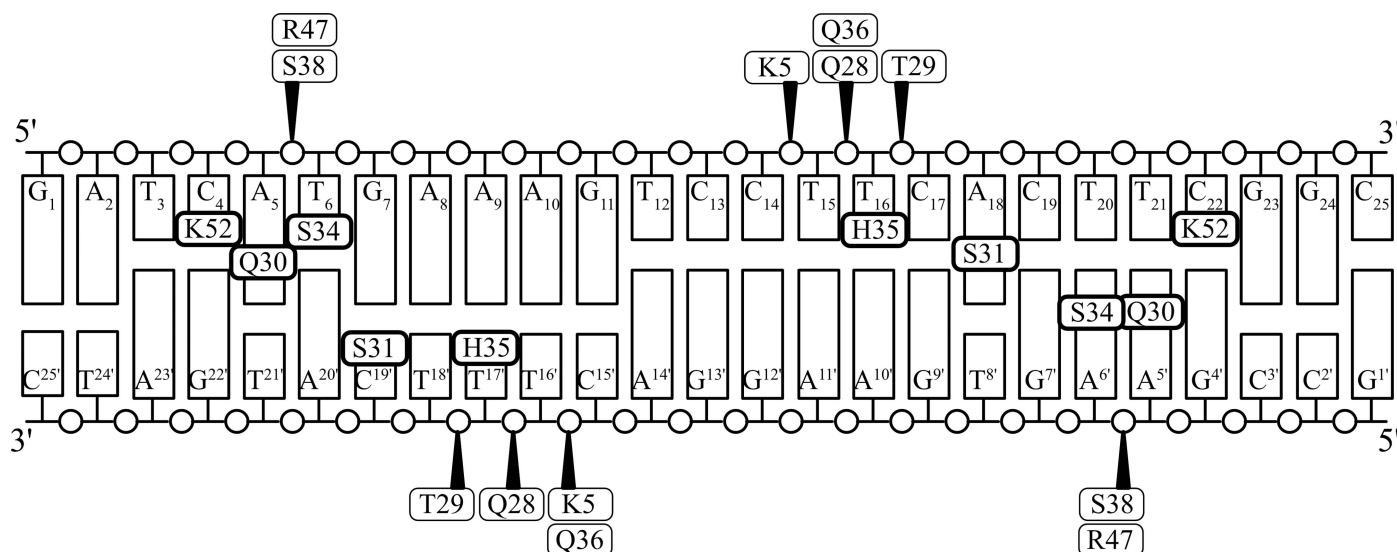


Figure 4

Schematic representation of MetR-DBD-DNA interactions. Thin-lined rounded rectangles with arrowheads represent amino acids that make DNA-backbone contacts and thick-lined rounded rectangles represent amino acids that contact DNA bases.



backbone of the DNA. Residues Lys5, Gln36 and Ser38 from the HTH motif and Arg47 from the wing region are equivalent to Arg4, Gln35, Gln37 and Gln46, respectively, in the BenM-DBD-LH-DNA structure (Alanazi *et al.*, 2013), which make similar hydrogen-bond interactions with the phosphate backbone (Fig. 4). Additionally, the dipole moment of helix  $\alpha 2$  in MetR further stabilizes the interface by interaction with the DNA backbone.

In the BenM-DBD-LH-DNA complex, residues Ala28, Gln35, Gln37, Asn38 and Arg50 perform indirect readout of sequence-positioned phosphates (Alanazi *et al.*, 2013). The equivalent residues in MetR are Thr29, Gln36, Ser38, Asp39 and Arg51, respectively. Thr29 in both monomers is hydrogen-bonded to the backbone phosphate groups of cytosine C17 (in TTCA) and thymines T16' and T17' in the complementary strand (Fig. 4). Ser34 is well conserved in LTTRs (Fig. 2d) and the C $\beta$  atom of Ser34 may recognize the methyl group of thymidine T6 (in TGAA), similar to the role suggested for Ser33 in BenM. Additionally, clamping of the DNA phosphate backbone of the BenM recognition site through hydrogen bonds from the hydroxyl O atom of Ser33 and the  $\epsilon$ -NH of Arg50 have been suggested to enhance the sensitivity of sequence readout (Alanazi *et al.*, 2013). Our model suggests that Ser34 and Arg51 would play similar roles in MetR. In our MetR-DBD-LH-DNA model Gln28 in monomer A hydrogen-bonds to thymine T16 (TTCA) in the template

strand, but in monomer B it interacts with T17' in the complementary strand (Fig. 4). Interestingly, other LTTR-DBDs contain a hydrophobic amino acid at this position (Fig. 2d), suggesting that Gln28 may play an additional role in the indirect readout of the operator sequence.

In the BenM-DBD-LH-DNA structure, Gln29, Pro30, Pro31 and Arg34 at the N-terminal end of the recognition helix  $\alpha 3$  were identified as critical for making base-specific interactions with the operator recognition site (Alanazi *et al.*, 2013). The equivalent residues in MetR are Gln30, Ser31, Ala32 and His35, respectively (Fig. 3b). Gln30 is a positionally conserved glutamine found in most LTTR proteins (Fig. 2d) that recognizes the adenine base at position 5 in the major groove (see later). In BenM, the side chains of Pro30 and Pro31 provide a complementary surface for direct readout of the bases through tight van der Waals interactions by a mechanism which is poorly understood (Alanazi *et al.*, 2013). In MetR, Ala32 has similar hydrophobic characteristics to Pro31, whereas the polar side chain of Ser31 differs from Pro30, producing different DNA contacts within the major groove (Fig. 3b). Interestingly, the amino-acid sequence QSALSHQ in the  $\alpha 3$  helix of MetR is very similar to another LTTR protein, NoIR, of the ArsR/SmtB family, which contains the amino-acid sequence QSALSQH in the DNA-recognition helix (Fig. 3c). In NoIR, Gln56, Ser57, Ser60 and Gln61 make hydrogen-bond interactions with the nucleotides

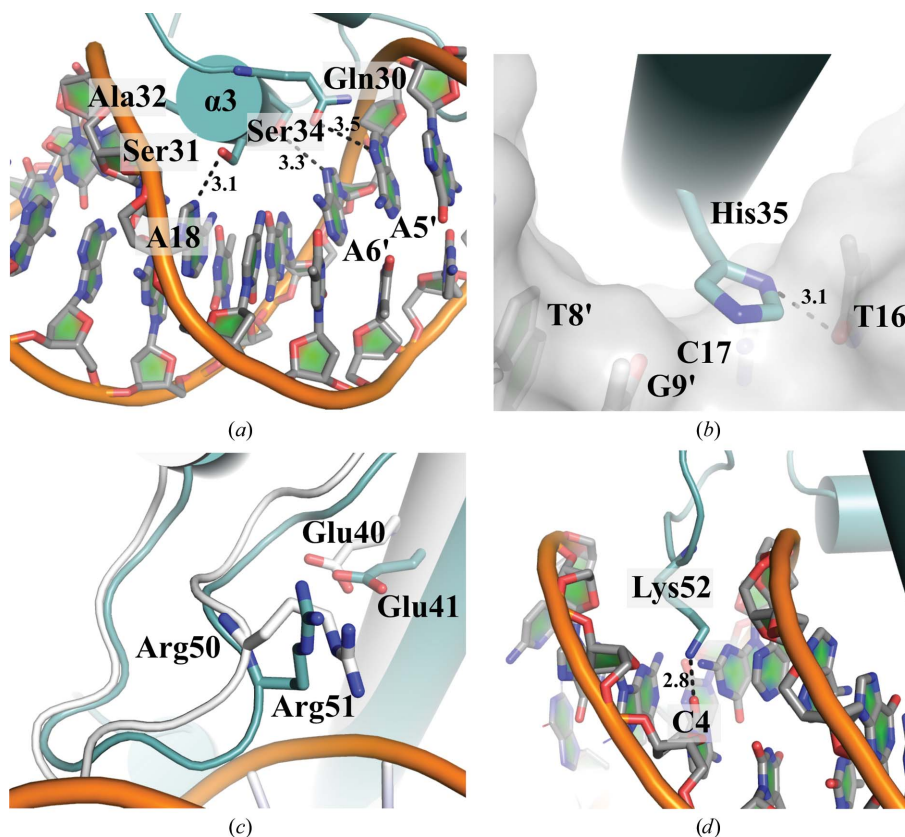


Figure 5

(a) Residues at the N-terminal end of recognition helix  $\alpha 3$  and (b) the side chain of His35 interact with nucleotides in the major groove. (c) The well conserved arginine and glutamic acid residues in MetR (teal) and BenM (grey). (d) Lys52 in the wing region penetrates to interact with the nucleotides in the minor groove. The hydrogen bonds are indicated by black dotted lines and the bond distances are shown in Å.



in the major groove and mutation of these amino acids resulted in a loss of DNA binding (Lee *et al.*, 2014). In the MetR-DBD dimer Gln30 makes a hydrogen-bond interaction with the N7 atoms of adenosines A5 and A5', and Ser31 interacts with the N7 atom of A18 in the template strand (TTCA) and the N4 atom of C19', the complementary base to guanosine G7 in the template strand (TGAA; Fig. 5*a*). Thus, our MetR-DBD-LH-DNA model shows that most of the base-specific contacts in the major groove come from Gln30, Ser31 and Ser34, which directly interact with nucleotides (Fig. 5*a*) located at the outer ends of the operator recognition site. This result corroborates previous findings that site-directed mutagenesis of the bases at the outer ends of the palindromic recognition half-sites drastically reduce MetR binding (Byerly *et al.*, 1991).

His35 in MetR could be a critical determinant of operator site recognition as it occupies a position similar to Arg34 in BenM and Gln61 in NoIR (Figs. 3*b* and 3*c*). A comprehensive structural analysis of DNA-binding proteins has shown that histidine side chains make bifurcated hydrogen-bond interactions, mainly with guanine, and van der Waals contacts by ring-stacking interactions to differentiate thymine from cytosine, thereby providing some degree of specificity through direct and indirect readout mechanisms (Luscombe *et al.*, 2001; Rohs *et al.*, 2010). In our MetR-DBD-LH-DNA model, the imidazole ring of His35 fits snugly into the major groove in a position to make base-specific contacts through hydrogen-bond and van der Waals interactions (Fig. 5*b*). The O4 atoms of thymines T16 and T17' are hydrogen-bonded to His35 N<sup>δ1</sup> in monomers *A* and *B*, respectively, which orients the His35 imidazole ring to provide a geometrically complementary surface for making van der Waals contacts with the edges of the bases of each DNA half-site lining the major-groove floors and walls.

Structural comparison between the apo DBDs of CbnR, TsaR and BenM and the BenM-DBD-LH-DNA complex shows that Glu40 and Arg50 in BenM-DBD are critical for positioning of the wing to contact the DNA (Alanazi *et al.*, 2013). In the MetR-DBD-LH structure the conserved residues Glu41 and Arg51 occupy equivalent positions (Fig. 5*c*). Lys52 at the tip of the wings probes the minor groove to make base-specific contacts and van der Waals interactions with the neighbouring base pairs (Fig. 5*d*). In particular, N<sup>ε</sup> of Lys52 in both subunits makes a single hydrogen bond to the exocyclic O2 of two cytosines, C4 and C22. A comprehensive structural analysis of protein-DNA complexes showed that positively charged amino-acid side chains in the wing region frequently recognize sequence-dependent variation in the minor-groove shape and electrostatic potential (Rohs *et al.*, 2009). Structure-based sequence alignment shows that His35 and Lys52 in MetR are not highly conserved among other LTTRs (Fig. 2*d*), suggesting that their interactions with the major and minor groove, respectively, may be important for operator site discrimination by MetR.

Our model of the MetR-DBD-LH-DNA complex shows how promoter recognition is achieved by a combination of direct base readout and indirect readout of the shape of the

major groove. It also suggests that binding of the DBD is sufficient to induce a bend in the DNA. Structural studies of other LTTRs (Ezezika *et al.*, 2007; Ruangprasert *et al.*, 2010; Zhou *et al.*, 2010; Taylor *et al.*, 2012) provide evidence that the binding of small effector molecules to the EBD triggers conformational changes in the protein that alter the LTTR structure from an inactive to an active state. This adjusts the relative positions of the DBDs with respect to each other, which could be expected to alter the local conformation of the DNA, but currently no structure of a full-length LTTR bound to DNA is available to confirm this. It is also unclear whether activation of transcription is a direct result of the change in conformation of the protein, the DNA or both, and further investigation is required to fully understand how gene expression is modulated by this important class of transcription factors.

## Acknowledgements

We thank Mark L. Urbanowski and Professor George V. Stauffer (Department of Microbiology, University of Iowa) for the generous gift of the MetR-91 plasmid encoding MetR-DBD-LH.

## References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Alanazi, A. M., Neidle, E. L. & Momany, C. (2013). *Acta Cryst.* **D69**, 1995–2007.
- Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. (2005). *FEMS Microbiol. Rev.* **29**, 231–262.
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
- Bogard, R. W., Davies, B. W. & Mekalanos, J. J. (2012). *MBio*, **3**, e00236–12.
- Byerly, K. A., Urbanowski, M. L. & Stauffer, G. V. (1991). *J. Bacteriol.* **173**, 3547–3553.
- Cai, X.-Y., Maxon, M. E., Redfield, B., Glass, R., Brot, N. & Weissbach, H. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 4407–4411.
- Cai, X.-Y., Redfield, B., Maxon, M., Weissbach, H. & Brot, N. (1989). *Biochem. Biophys. Res. Commun.* **163**, 79–83.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Cowan, J. M., Urbanowski, M. L., Talmi, M. & Stauffer, G. V. (1993). *J. Bacteriol.* **175**, 5862–5866.
- Craven, S. H., Ezezika, O. C., Haddad, S., Hall, R. A., Momany, C. & Neidle, E. L. (2009). *Mol. Microbiol.* **72**, 881–894.
- Dijk, M. van & Bonvin, A. M. J. J. (2009). *Nucleic Acids Res.* **37**, W235–W239.
- DiMaio, F. (2013). *Acta Cryst.* **D69**, 2202–2208.
- DiMaio, F., Terwilliger, T. C., Read, R. J., Wlodawer, A., Oberdorfer, G., Wagner, U., Valkov, E., Alon, A., Fass, D., Axelrod, H. L., Das, D., Vorobiev, S. M., Iwai, H., Pokkuluri, P. R. & Baker, D. (2011). *Nature (London)*, **473**, 540–543.
- Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G. & Baker, N. A. (2007). *Nucleic Acids Res.* **35**, W522–W525.
- Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. (2003). *J. Am. Chem. Soc.* **125**, 1731–1737.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
- Evans, P. & McCoy, A. (2008). *Acta Cryst.* **D64**, 1–10.

- Ezekika, O. C., Haddad, S., Clark, T. J., Neidle, E. L. & Momany, C. (2007). *J. Mol. Biol.* **367**, 616–629.
- Ferla, M. P. & Patrick, W. M. (2014). *Microbiology*, **160**, 1571–1584.
- Flatley, J., Barrett, J., Pullan, S. T., Hughes, M. N., Green, J. & Poole, R. K. (2005). *J. Biol. Chem.* **280**, 10065–10072.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Jo, I., Chung, I.-Y., Bae, H.-W., Kim, J.-S., Song, S., Cho, Y.-H. & Ha, N.-C. (2015). *Proc. Natl Acad. Sci. USA*, **112**, 6443–6448.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kim, D. E., Chivian, D. & Baker, D. (2004). *Nucleic Acids Res.* **32**, W526–W531.
- Lee, S. G., Krishnan, H. B. & Jez, J. M. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 6509–6514.
- Leyn, S. A., Suvorova, I. A., Kholina, T. D., Sherstneva, S. S., Novichkov, P. S., Gelfand, M. S. & Rodionov, D. A. (2014). *PLoS One*, **9**, e113714.
- Li, Z., Natarajan, P., Ye, Y., Hrabe, T. & Godzik, A. (2014). *Nucleic Acids Res.* **42**, W240–W245.
- Lorenz, E. & Stauffer, G. V. (1995). *J. Bacteriol.* **177**, 4113–4120.
- Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. (2001). *Nucleic Acids Res.* **29**, 2860–2874.
- Maddocks, S. E. & Oyston, P. C. F. (2008). *Microbiology*, **154**, 3609–3623.
- Mares, R., Urbanowski, M. L. & Stauffer, G. V. (1992). *J. Bacteriol.* **174**, 390–397.
- Maxon, M. E., Redfield, B., Cai, X.-Y., Shoeman, R., Fujita, K., Fisher, W., Stauffer, G., Weissbach, H. & Brot, N. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 85–89.
- Maxon, M. E., Wigboldus, J., Brot, N. & Weissbach, H. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 7076–7079.
- McCoy, A. J. (2007). *Acta Cryst.* **D63**, 32–41.
- McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). *Acta Cryst.* **D67**, 386–394.
- Monferrer, D., Tralau, T., Kertesz, M. A., Dix, I., Solà, M. & Usón, I. (2010). *Mol. Microbiol.* **75**, 1199–1214.
- Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K. & Senda, T. (2003). *J. Mol. Biol.* **328**, 555–566.
- Phillips, S. E. V. (1980). *J. Mol. Biol.* **142**, 531–554.
- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
- Rohs, R., Jin, X., West, S. M., Joshi, R., Honig, B. & Mann, R. S. (2010). *Annu. Rev. Biochem.* **79**, 233–269.
- Rohs, R., West, S. M., Sosinsky, A., Liu, P., Mann, R. S. & Honig, B. (2009). *Nature (London)*, **461**, 1248–1253.
- Ruangprasert, A., Craven, S. H., Neidle, E. L. & Momany, C. (2010). *J. Mol. Biol.* **404**, 568–586.
- Sainsbury, S., Ren, J., Saunders, N. J., Stuart, D. I. & Owens, R. J. (2012). *Acta Cryst.* **F68**, 730–737.
- Schell, M. A. (1993). *Annu. Rev. Microbiol.* **47**, 597–626.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- Söding, J., Biegert, A. & Lupas, A. N. (2005). *Nucleic Acids Res.* **33**, W244–W248.
- Sperandio, B., Gautier, C., McGovern, S., Ehrlich, D. S., Renault, P., Martin-Verstraete, I. & Guédon, E. (2007). *J. Bacteriol.* **189**, 7032–7044.
- Taylor, J. L., De Silva, R. S., Kovacicova, G., Lin, W., Taylor, R. K., Skorupski, K. & Kull, F. J. (2012). *Mol. Microbiol.* **83**, 457–470.
- Terwilliger, T. C., DiMaio, F., Read, R. J., Baker, D., Bunkóczi, G., Adams, P. D., Grosse-Kunstleve, R. W., Afonine, P. V. & Echols, N. (2012). *J. Struct. Funct. Genomics*, **13**, 81–90.
- Urbanowski, M. L. & Stauffer, G. V. (1987). *J. Bacteriol.* **169**, 5841–5844.
- Urbanowski, M. L. & Stauffer, G. V. (1989a). *J. Bacteriol.* **171**, 3277–3281.
- Urbanowski, M. L. & Stauffer, G. V. (1989b). *J. Bacteriol.* **171**, 5620–5629.
- Urbanowski, M. L., Stauffer, L. T., Plamann, L. S. & Stauffer, G. V. (1987). *J. Bacteriol.* **169**, 1391–1397.
- Vries, S. J. de, van Dijk, M. & Bonvin, A. M. J. J. (2010). *Nature Protoc.* **5**, 883–897.
- Wassenaar, T. A. *et al.* (2012). *J. Grid Comput.* **10**, 743–767.
- Weissbach, H. & Brot, N. (1991). *Mol. Microbiol.* **5**, 1593–1597.
- Yeung, A. T. Y., Torfs, E. C. W., Jamshidi, F., Bains, M., Wiegand, I., Hancock, R. E. W. & Overhage, J. (2009). *J. Bacteriol.* **191**, 5592–5602.
- Zhou, X., Lou, Z., Fu, S., Yang, A., Shen, H., Li, Z., Feng, Y., Bartlam, M., Wang, H. & Rao, Z. (2010). *J. Mol. Biol.* **396**, 1012–1024.