Published online 15 June 2016

Nucleic Acids Research, 2016, Vol. 44, No. 14 ***6981–6993***
doi: 10.1093/nar/gkw543

Structural basis of thiol-based regulation of formaldehyde detoxification in *H. influenzae* by a MerR regulator with no sensor region

Rafael M. Couñago[1,2,3,*,†], Nathan H. Chen[1,2,†], Chiung-Wen Chang[1,2,3,†], Karrera Y. Djoko[1,2], Alastair G. McEwan[1,2,*] and Bostjan Kobe[1,2,3,*]

[1]School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, Qld 4072, Australia, [2]Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, Qld 4072, Australia and [3]Institute for Molecular Bioscience, University of Queensland, Brisbane, Qld 4072, Australia

Received December 12, 2015; Revised May 12, 2016; Accepted June 03, 2016

ABSTRACT

Pathogenic bacteria such as *Haemophilus influenzae,* a major cause of lower respiratory tract diseases, must cope with a range of electrophiles generated in the host or by endogenous metabolism. Formaldehyde is one such compound that can irreversibly damage proteins and DNA through alkylation and cross-linking and interfere with redox homeostasis. Its detoxification operates under the control of HiNmlR, a protein from the MerR family that lacks a specific sensor region and does not bind metal ions. We demonstrate that HiNmlR is a thiol-dependent transcription factor that modulates *H. influenzae* response to formaldehyde, with two cysteine residues (Cys54 and Cys71) identified to be important for its response against a formaldehyde challenge. We obtained crystal structures of HiNmlR in both the DNA-free and two DNA-bound forms, which suggest that HiNmlR enhances target gene transcription by twisting of operator DNA sequences in a two-gene operon containing overlapping promoters. Our work provides the first structural insights into the mechanism of action of MerR regulators that lack sensor regions.

INTRODUCTION

Bacterial cells must quickly adapt to changing environmental conditions, a process that is mostly based on the modulation of gene expression (1). The MerR-family of transcription factors controls stress responses to a diverse range of effectors, including xenobiotics (BmrR), redox stress (SoxR, NmlR) and metal-ion overload (MerR, ZntR, CueR) (2,3). MerR-family regulators are modular proteins consisting of a structurally similar N-terminal, helix–turn–helix (HTH) DNA-binding domain (DBD) and a dissimilar C-terminal sensor region that can range in size from a few residues, as seen for the metal-sensing regulators such as CueR, to hundreds of amino acids, as observed for the xenobiotic-sensing proteins such as BmrR. Members of this family assemble into head-to-tail dimers via a coiled-coil formed by a long central helical domain (4–10). In both *holo* and *apo* forms, MerR-family proteins bind tightly and specifically to quasi-palindromic inverted repeat (IR) DNA sequences within the target gene promoter. These regulators modulate gene expression by an unprecedented DNA untwisting mechanism (8,11–14). Binding of signaling compounds to the C-terminal sensor region prompts the unwinding of the bound DNA molecule via the re-orientation of the protein N-terminal DNA-binding domains (DBDs), a movement that is accommodated by the central coiled-coil.

The overall result of this remarkable conformational change is to re-orient the −10 and −35 elements in a fashion that allows open complex formation by RNA polymerase (RNAP). For transcription initiation to occur, regions 2.4 and 4.2 in the RNAP σ subunit must contact, simultaneously, the −10 and −35 elements from the promoter, respectively. In strong, 'optimal' promoters, spacer regions between these elements consist of 17–18 base pairs (bp), which places them on the same face of a B-form DNA molecule. By contrast, MerR-family regulons have elongated (19–20 bp) spacer regions, placing the −10 and −35

*To whom correspondence should be addressed. Tel: +617 3365 2132; Fax: +617 3365 4699; Email: b.kobe@uq.edu.au
Correspondence may also be addressed to Rafael M. Couñago. Tel: +55 19 3521 5234; Fax: +617 3365 4699; Email: rafael.counago@sgc.ox.ac.uk
Correspondence may also be addressed to Alastair G. McEwan. Tel: +617 3346 4110; Fax: +617 3365 4699; Email: a.mcewan@uq.edu.au
†These authors contributed equally to this work as First Authors.
Present addresses:
Rafael M. Couñago, Protein Kinase Chemical Biology Center, Structural Genomics Centre, UNICAMP, Campinas, SP 13083, Brazil.
Chiung-Wen Chang, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA.

elements on opposite faces of the DNA molecule, thus hindering gene transcription (2). Each DBD in a MerR-family protein dimer recognizes one half-site (9 bp) in a 19–20 bp quasi-palindromic, inverted repeat (IR) sequence, generally found nested inside the spacer region between the −10 and −35 elements. As sites recognized by MerR-family proteins and RNAP do not overlap, these proteins are thought to form a ternary complex with the nucleic acid, in which the transcription factor and the polymerase occupy opposite faces of the DNA molecule. Structural studies have also suggested that, in a MerR protein dimer, intimate contacts between the C-terminal sensing region from one protomer and the DBD from the second protomer play an important role during these structural rearrangements (4,7,10,14). Thus, to date, the function of MerR-family proteins is thought to rely on the unusual promoter architecture and on conformational changes initiated by the activation of a C-terminal sensor region (7,14,15).

Nevertheless, several bacteria encode MerR-family proteins that do not conform to this paradigm. These proteins do not have a recognizable C-terminal metal-binding site or a large sensor region and bind to IR sequences that flank the -35 promoter elements of their target genes. These unusual MerR-family proteins are known as AdhR in Gram-positive organisms, including *Lactobacilli*, *Clostridia* and *Bacillus subtilis*; and as NmlR (*Neisseria* MerR-like regulator) in Gram-negative bacteria. Within this latter group, NmlR appears restricted to the pathogens *Neisseria gonorrhoeae*, *N. meningitidis* and *Haemophilus influenzae* (16). In place of a C-terminal sensor region, function of NmlR-type regulators seem to rely on conserved Cys residues located at the hinge region between the N-terminal DBD and the central dimerization helix. Modification of reactive cysteines is emerging as an effective stress-sensing mechanism in transcriptional regulators (17,18). AdhR and NmlR regulons commonly include genes involved in neutralizing cytotoxic compounds produced via oxidative stress. To better characterize these unusual MerR-family proteins, we investigated the structure and function of the NmlR protein from *H. influenzae* (HiNmlR). Our data suggests that HiNmlR is a thiol-dependent transcription factor that modulates response to formaldehyde and functions by DNA unwinding, despite the lack of C-terminal sensor region.

MATERIALS AND METHODS

Construction of *H. influenzae* recombinant strains

Haemophilus influenzae strain Rd KW20 was routinely grown on brain heart infection (BHI, Becton Dickinson) agar supplemented with hemin (10 μg.ml^{-1}) and NAD$^+$ (10 μg.ml^{-1}) at 310 K. The *nmlR* gene (KEGG entry HI0186) was inactivated following transformation of *H. influenzae* strain Rd KW20 with a linear plasmid containing the pUC4K (Amersham) kanamycin resistance cassette flanked by genomic regions (∼1 kb) found up- and downstream from *nmlR*. Complementation of the *nmlR*-null strain was achieved by insertion of the wild type or various site-directed variants of the *nmlR* gene into the HI0601.1 open reading frame of *H. influenzae* Rd KW20. All mutant strains were confirmed by DNA sequencing. Oligonucleotide primer sequences can be found in Supplementary Table S1.

Sensitivity of *H. influenzae* strains to reactive aldehydes

Disc-diffusion susceptibility assays were performed by placing paper discs containing 10 μl of each stress reagent onto a layer of cells spread on BHI agar. The zone of clearing around the disc was measured after overnight incubation. For plating efficiency (growth) assays, fresh lawns of cells from an overnight agar plate were solubilized in BHI broth (OD600 of 0.4–8 × 10^7 CFU/ml), and serial dilutions (5 μl each) were plated on BHI solid medium containing increasing concentrations of formaldehyde (0–2.0 mM).

Measurement of AdhC activity

Overnight cultures from supplemented BHI agar were suspended into BHI broth and spread onto supplemented BHI agar with or without 0.8 mM formaldehyde. After overnight incubation, the cells were harvested and suspended into 2 ml of 50 mM Hepes (pH 7.0). Cells were lysed using a French pressure cell press (two passages at 16000 psi) (Thermo Scientific) and the lysate was clarified by centrifugation. Glutathione-dependent alcohol dehydrogenase activity was assayed using HMGSH (*S*-hydroxymethylglutathione) as the substrate. HMGSH was prepared by mixing equimolar amounts of GSH and formaldehyde (19). Briefly, 50 μl of cell lysate was added to a solution of 1 mM HMGSH and 4 mM NAD+ in 50 mM Hepes (pH 7.0) buffer. The reaction was monitored at 340 nm for the production of NADH at 310 K. The initial rates of reaction were obtained 30 s after the addition of cell lysate. All experiments were performed in triplicate. Results were standardized against the total protein content found in the cell lysate as measured using a BCA protein assay (BCA QuantiPro Kit, Sigma).

Recombinant expression and purification of NmlR and mutants

PCR amplification of *H. influenzae Rd KW20 nmlR* (KEGG entry HI0186) and ligation-independent cloning to insert the gene into the N-terminal hexahistidine tag-containing vector pMCSG7 (20) were used to generate pMCSG7-Hi0186 for recombinant expression of HiNmlR. Site-directed mutagenesis (Quickchange Lightning Kit, Agilent Technologies) was used to introduce specific mutations in the coding sequence of pMCSG7-Hi0186. Protein expression was performed in *Escherichia coli* BL21(DE3), by growing the cells in the autoinduction medium (21) for 18 h at 297 K. Cells were harvested and disrupted by sonication in lysis buffer (50 mM Tris–HCl (pH 8.0), 500 mM NaCl, 30 mM imidazole and 1 mM TCEP [tris(2-carboxyethyl)phosphine]). The soluble supernatant was clarified by centrifugation at 277 K for 60 min at 27 000 × *g*. HiNmlR was purified using a HisTrap FF Crude column (5 ml; GE Healthcare) equilibrated with lysis buffer and eluted with 150 mM imidazole in lysis buffer. The hexahistidine tag was removed by incubation with TEV (tobacco etch virus) 3C protease (overnight at 277 K) followed by a second passage on a HisTrap FF Crude column. HiNmlR was further purified on a HiLoad 26/60 Superdex 75

gel filtration column (GE Healthcare) equilibrated with 25 mM Tris–HCl (pH 8.0), 150 mM NaCl, 5% (w/v) glycerol and 1 mM TCEP. Purified HiNmlR was concentrated to 25 mg.ml^{-1} using Amicon Ultra-4 centrifugal devices (Millipore) and flash-cooled in a liquid nitrogen bath prior to storage at −193 K.

Measurement of cysteine thiols

Reduced forms of HiNmlR (50 μM) were prepared by incubation with excess DTT (100 molar equivalents) for 30 min at 310 K, followed by desalting into the buffer containing 20 mM Tris and 150 mM NaCl (pH 8). Iodoacetamide (50 molar equivalents) was added and incubated for 1 h at 310 K. The modified HiNmlR was analysed using ESI-MS (electrospray ionization mass spectrometry) data collected on a QSTAR Pulsar ESI-QqTOF (Applied Biosystems) connected to a 2.1 × 150 mm C18 (5 μm, Phenomenex) column, using an Agilent Binary 1100 HPLC system. The positive ion (ESI+) mode with a cone voltage 25–35 V and a flow rate of 5 μl acetonitrile per min was used. The average molar mass of each sample was obtained by applying a deconvolution algorithm to the recorded spectrum, which was calibrated with horse heart myoglobin (16951 Da).

Microscale thermophoresis

Single-stranded DNA oligonucleotides were annealed in 20 mM Tris–HCl (pH 7.4), 2 mM $MgCl_2$, 50 mM NaCl in 100 μl reactions containing 10 μM Cy3-labeled forward oligonucleotide and 10.8 μM unlabeled reverse oligonucleotide. The annealing mixture was checked by obtaining comparable absorbance values at 260 nm for double-stranded (ds) DNA, and 550 nm for Cy3-labeled molecules. The interaction between annealed dsDNA and wild-type HiNmlR protein was confirmed by electrophoretic mobility shift assays (EMSAs). Quantitative binding assays were carried out using a Nanotemper Monolith NT.115 (settings: LED power: 100%, IR-laser power: 20%, 298 K). Each set of affinity measurements for one pair of dsDNA oligonucleotides contained 16 reactions, performed in 20 mM Tris–HCl (pH 7.6), 1.5 mM $MgCl_2$, 0.5 mM EGTA, 10% glycerol, 300 mM KCl, 0.5 mg/ml bovine serum albumin (BSA) and freshly added 10 mM TCEP. The Cy3-labeled dsDNA oligonucleotide was kept at a constant concentration of 100 nM. Unlabeled protein was titrated in a 1:1-dilution series (starting concentration of 94.8 μM). The reaction mixture was kept at 298 K for 10 min before loading into standard capillaries (NT.115). Prior to measurements, mixtures were further incubated for 15 min at 298 K in the Nanotemper. The data was analyzed using the NT-analysis acquisition software, which plots a binding curve using the normalized fluorescence of the labeled dsDNA at different concentrations of the unlabeled protein. Binding assays were performed in triplicate and the mean value was calculated.

Protein crystallization, structure determination and refinement

Protein:DNA co-crystallization experiments employed 10 mg.ml^{-1} and a 1:1.2 protein:DNA molar ratio. Following sparse matrix crystallization screens and optimization, large crystals (300 μm x 200 μm x 100 μm) of HiNmlR in complex with wild type or fully-symmetrical P*adhC-estD/nmlR* IR operator sequence (P*adhC-estD/nmlR(wt)* and P*adhC-estD/nmlR(sym)*, respectively) were obtained in 20% (w/v) PEG 3350, 0.4 M sodium nitrate, 0.1 M Bis–Tris propane (pH 7.5) and 2 mM TCEP after 3–5 days at 293 K. Crystals for the DNA-free Se-methionine-containing HiNmlR protein were grown in 0.1 M Bis–Tris (pH 6.5) and 25% PEG 3350. Prior to data collection, crystals were mounted in nylon loops (Hampton Research) and flash-cooled by rapid immersion in liquid nitrogen. Protein:DNA complex crystals were briefly transferred to a solution of 20% (w/v) PEG 3350, 0.4 M sodium nitrate, 0.1 M Bis–Tris propane (pH 7.5), 2 mM TCEP and 16% (v/v) glycerol prior to flash-cooling. Diffraction data on single native crystals were collected at the wavelength of 0.95 Å at the Australian Synchrotron or at UQROCX facility at the University of Queensland using a rotating Cu anode X-ray generator (1.54 Å wavelength). A single-wavelength anomalous dispersion (SAD) data-set was collected on a single crystal of Se-Met-incorporated HiNmlR protein (0.5° oscillation and 360° rotation). Data-sets were processed with XDS (22) and scaled and merged in Aimless within the CCP4 suite (23) (native data-sets) or XSCALE within XDS (Se-Met HiNmlR data-set). The HiNmlR:DNA complex structures were determined by molecular replacement (MR) using Phaser (search model PDB ID 3GPV) (24). The structure of the DNA-free, Se-Met HiNmlR was determined via MR-SAD in Phaser (24), using the DNA-bound structure as the search model and the anomalous signal from Se. Automatic model building was performed with Phenix AutoBuild (25,26). Structure refinement was performed using Buster (Global Phasing) iteratively with manual model building in Coot (27). Data processing and structure refinement statistics can be found in Supplementary Table S2. Resolution cut-offs were determined based on percentage of correlation between intensities from random half-data-sets at 0.1% significance level (28). Model manipulations were performed using Chimera (29). The program CURVES+ (30) was used to measure the widths of major and minor grooves. Buried surface areas upon complex formation were calculated using PISA (31). Electrostatic surface representations were created with APBS (32) following the assignment of atomic charge and radius information with PDB2PQR (33). All structural representations were prepared with PyMOL (Schrödinger, LLC). Protein domain motions were analyzed using DynDom (34). The coordinates and structure factors have been deposited in the Protein Data Bank with ID 5D8C (wild-type HiNmlR:P*adhC-estD/nmlR(wt)* DNA complex); 5E01 (wild-type HiNmlR:P*adhC-estD/nmlR(sym)* DNA complex); 5D90 (wild-type HiNmlR). Sequence patterns were identified using the Pattern Locator server (http://www.cmbl.uga.edu/software/patloc.html) (35).

RESULTS

HiNmlR is a regulator of formaldehyde detoxification in *H. influenzae*

In *N. meningitidis* and *H. influenzae*, NmlR controls expression of an operon consisting of two genes, *adhC* and

estD, involved in the glutathione-dependent detoxification of toxic formaldehyde to formate (Supplementary Figure S1A,C) (36–38). To establish the *in vivo* function of HiNmlR (KEGG entry HI0186), we constructed an *nmlR*-null mutant in *H. influenzae* strain Rd KW20 and tested its sensitivity towards reactive aldehydes using a disc-diffusion assay. Although insensitive to most aldehydes tested here, growth of the *nmlR* mutant was impaired by formaldehyde, as indicated by a larger zone of clearance in the presence of this aldehyde. Complementation of *nmlR* via ectopic expression of *nmlR* on the *H. influenzae* chromosome restored the zone of clearance to the same level as in wild-type cells. These data confirm that NmlR is required for defense against formaldehyde toxicity (Figure 1A).

HiNmlR is a thiol-dependent transcription factor

In view of the recognized mode of action of MerR-family regulators it would be expected that HiNmlR would operate as an inducer in response to formaldehyde. However, the lack of a recognizable C-terminal sensor region in the NmlR proteins and the presence of conserved Cys residues (at positions 54, 71 and 95) (Supplementary Figure S2) suggested that HiNmlR might function as a thiol-dependent transcription factor. Therefore, we systematically tested the contribution of individual HiNmlR Cys residues to *H. influenzae* growth at increasing formaldehyde concentrations (0–2.0 mM). As for the disc-diffusion assays above, the *nmlR*-null strain showed severely decreased growth at higher formaldehyde concentrations compared to wild-type cells. Likewise, mutation of HiNmlR Cys54 to alanine resulted in severely decreased growth, regardless of the residues occupying positions 71 and 95. By contrast, single alanine mutations at positions Cys71 and Cys95 did not result in any noticeable reduction in cell growth (Figure 1B). Thus, these experiments suggest that Cys54 plays a critical role during HiNmlR-mediated response to formaldehyde.

We then tested the relevance of individual Cys residues in the activation of the P*adhC-estD* operon by measuring AdhC activity (via HMGSH-mediated reduction of NAD^+) upon growth in media supplemented with a sub-lethal concentration of formaldehyde (0.8 mM) (Figure 1C). AdhC activity has been shown previously to correlate with the level of gene expression (16,36,37). Interestingly, under these conditions, the *nmlR*-null mutant displayed low levels of AdhC activity, similar to the basal levels observed for the unchallenged wild-type strain. Wild-type cells challenged with 0.8 mM formaldehyde displayed a ∼30-fold increase in AdhC activity, compared to unchallenged cells. Similar results were observed for the C95A mutant strain (∼30-fold increase upon formaldehyde challenge), whereas activation of AdhC in C71A cells upon formaldehyde challenge was somewhat lower (∼20-fold increase). These results suggest that Cys71 but not Cys95 in HiNmlR also have a role in activating *adhC* transcription during formaldehyde challenge. By contrast, unchallenged C54A mutant cells displayed considerably lower AdhC activity compared to wild-type cells under similar conditions. Moreover, AdhC activity remained low in C54A cells even after a formaldehyde challenge. Similar results were obtained for double and triple mutants containing the HiNmlR C54A mutation. Interestingly, cells expressing HiNmlR containing only Cys54 (C71A/C95A mutant) were unresponsive to formaldehyde challenge (Figure 1C), but could sustain basal levels of AdhC activity similar to those observed for unchallenged wild-type cells. Together, these observations suggest a critical role for Cys54 in HiNmlR function, and a lesser but important role for Cys71 in the activation of the P*adhC-estD* operon under formaldehyde stress.

To determine if these Cys-to-Ala mutations were affecting the ability of HiNmlR to bind to its operator, recombinant HiNmlR with a C-terminal His tag was expressed and purified. The purified HiNmlR was found to contain three cysteine thiols in their reduced state, as determined through alkylation using iodoacetamide followed by MS analysis (Supplementary Figure S3). Microscale thermophoresis assays revealed that Cys-to-Ala mutations had only a small impact on HiNmlR binding affinity for the target DNA (a 24-bp oligonucleotide corresponding to the wild-type P*adhC-estD/nmlR* IR operator sequence) (Table 1 and Supplementary Figure S4). This indicated that the phenotypes observed for the C54A and C71A mutants shown above were not due to the inability of HiNmlR to bind and recognize its operator, but were more likely due to defects in their abilities to sense formaldehyde and induce gene expression.

Structure of the HiNmlR:P*adhC-estD/nmlR* promoter DNA complex

In addition to lacking a recognizable sensor region, HiNmlR also interacts with an unusual DNA target (Supplementary Figure S1A,B, Figure 2A). To better understand how HiNmlR interacts with the target DNA and the role of its cysteine residues in gene activation, we determined the crystal structures of full-length HiNmlR with and without DNA. Co-crystals were obtained for full-length wild-type HiNmlR bound to a 18-bp double-stranded (ds) DNA oligonucleotide consisting of the wild-type genomic sequence of the target P*adhC-estD/nmlR* promoter IR (16 bp) flanked by single C and G residues (at the 5′ and 3′ ends, respectively; termed P*adhC-estD/nmlR(wt)*; Supplementary Table S1) and for the same protein in complex with a modified, fully symmetrical IR sequence (termed P*adhC-estD/nmlR(sym)*; Supplementary Tables S1 and S2).

Conformational changes notwithstanding (discussed below), both DNA-free and DNA-bound structures display the prototypical modular organization observed for MerR-family proteins: a winged, helix–turn–helix (HTH) DBD (helices α1–α4; residues 1–74) and a dimerization helix (helix α5; residues 75–119). The C-terminal region of HiNmlR (residues 119–135) is mostly disordered; the nine C-terminal residues (127-135) could not be modeled into electron density maps and were thus omitted from the final model. In the co-crystals, a HiNmlR dimer was bound to a single dsDNA molecule (Figure 2A and B). Both P*adhC-estD/nmlR(wt)* and P*adhC-estD/nmlR(sym)* DNA-bound structures are very similar (Supplementary Figure S5) and all subsequent analyses are performed with the one containing the wild-type (asymmetrical) IR site (P*adhC-estD/nmlR(wt)*).

Cysteine residues 54 and 71 localize to α3 and α4, respectively, of the DBD, whereas Cys95 is in the central region of the dimerization helix α5. In the crystal, these two dimeriza-

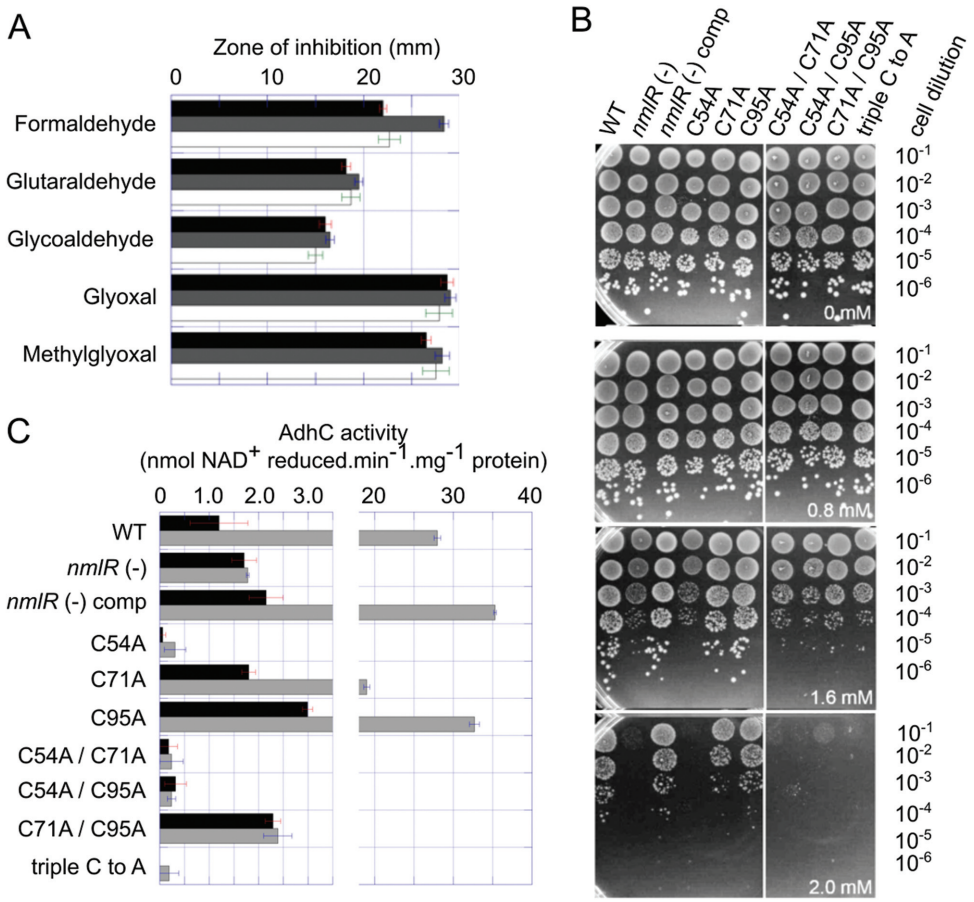

Figure 1. HiNmlR and formaldehyde defense in *H. influenzae* Rd KW20. (**A**) Disc-diffusion assays of the wild-type (WT), black bars; *nmlR*-null (−), gray bars; and *nmlR*-complemented ((−) comp), white bars; strains in the presence of reactive aldehydes as indicated. (**B**) Formaldehyde sensitivity of wild-type and mutant *nmlR* variants on solid media containing fixed concentrations of formaldehyde. Cell dilutions are indicated on the right. (**C**) AdhC activity levels in the wild-type and mutant *nmlR* variants grown on solid media in the absence and presence of 0.8 mM formaldehyde. 'Triple C to A' refers to the triple mutant C54A/C71A/C95A. Results are the average of three independent experiments (error bars indicate the standard error of the mean).

Table 1. DNA-binding affinity constants (nM)

HiNmlR protein	ds 24-bp P*adhC-estD*/*nmlR* IR operon			
	Wild-type	G5′C[a]	C5G[a]	G5′C/C5G[a]
Wild-type	25.2 ± 2.3	245.0 ± 26.1	252.0 ± 17.4	4690.0 ± 311.0
C54A	91.1 ± 15.6	n/d[b]	n/d	n/d
C71A / C95A	33.0 ± 3.4	n/d	n/d	n/d
C54A / C95A	47.4 ± 6.7	n/d	n/d	n/d
C54A / C71A / C95A	55.3 ± 5.3	n/d	n/d	n/d

[a]Numbering for P*adhC-estD* template strand.
[b]n/d, not determined.

tion helices form a long and closely-packed coiled-coil (Figure 2B). The protein DBD is formed by four tightly packed helices displaying an extensive network of hydrophobic interactions that also involve α5′ from the other protomer (Figure 2B and C). Inspection of electron density maps did not suggest cysteine residues were modified in any way (Supplementary Figure S6).

The overall architecture of individual domains in the HiNmlR:DNA complex is similar to the one observed for the equivalent regions of other MerR proteins. Differences are mostly restricted to the 'wing' loop and the α4 helix (Figure 2D). Nevertheless, superposition of the DBD from different MerR proteins reveals that, in a protein dimer, there are pronounced differences in the relative positions of the DNA-interacting helices α2- and α2′ (Figure 2D). These differences are most likely due to disparate activation states amongst analyzed structures. The relative orientation of α2/α2′ in the HiNmlR dimer is more similar to the one observed for the repressor state of CueR than to the activator states of CueR and two other MerR factors, SoxR and BmrR (Figure 2D).

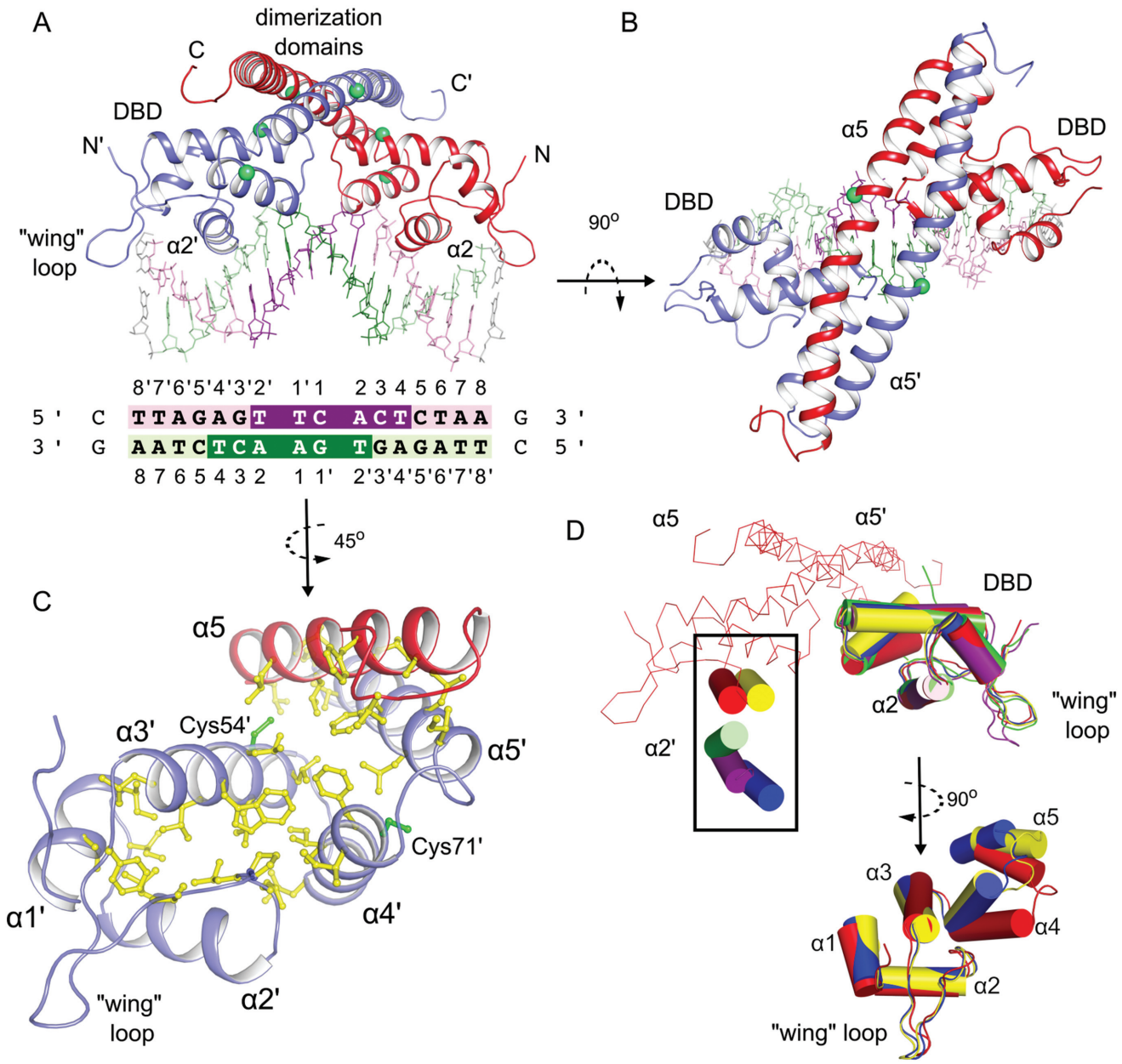

Figure 2. Crystal structure of the HiNmlR dimer bound to the P*adhC-estD/nmlR* IR operator (P*adhC-estD/nmlR(wt)*) DNA. (**A**) Overall structure of the protein:DNA complex. Protein monomers (ribbons) are indicated in different colors. DNA-binding domain (DBD), 'wing' (α2-α3) loop and dimerization domain (α5-helix) are indicated; prime denotes the equivalent regions from the other protomer. Operator DNA fragment (stick representation) with template strands for the P*adhC-estD* operon and the P*nmlR* gene shown in light pink and green, respectively; the -35 elements are shown in darker shades of the same colors. Nucleotides in white are not part of the wild-type IR sequence, shown at the bottom in bold and color-boxed as above. (**B**) Coiled-coil formed by dimerization helices (α5 and α5′) from the two protomers in the dimer. C95 and C95' (Cα) are indicated by green spheres. (**C**) Packing in the DBD and dimerization domain showing hydrophobic residues (yellow in stick representation). Conserved Cys residues (Cys54′ and Cys71′) important for HiNmlR function are shown in green. Protomers are colored as in panel A. (**D**) Structural comparisons of MerR proteins. Relative position of α2′-helix (boxed) following superposition of equivalent Cα atoms in the DBDs from the other protomer (top). Superimposed proteins: HiNmlR (this work; red), activator complexes for SoxR (PDB ID 2ZHG (10); green), BmrR (PDB ID 1EXI (7); purple) and CueR (PDB ID 4WLW (14); repressor complex for CueR (PDB ID 4WLS (14); yellow). The DBD from one protomer and the α2′ helix from the other protomer are shown as ribbons. Additional regions in HiNmlR are shown as a Cα trace and omitted in CueR, SoxR and BmrR for clarity. Bottom - organization of the DBD for repressor (HiNmlR; red; and CueR; yellow) and activator (Ag^{+}:CueR, blue) complexes.

Promoter recognition

MerR proteins utilize residues in their α2-helix to engage the major groove in the target DNA. The DNA-bound structure of HiNmlR suggests that α2-helix residues Tyr17, Arg20, Phe21 and Lys24 are important for DNA recognition. Residues Tyr17 and Phe21 make van der Waals contacts to Cyt5 and Cyt3, respectively, and their phenolic rings are found perpendicular to the pyrimidine rings of the nucleotides. The side-chains from these amino-acids also provide a hydrophobic cavity that accommodates the methyl group from Thy4. Residue Lys24 can engage Gua5′, opposite to Cyt5, and Ade6′ via hydrogen bonds. Finally, the bulky and basic side-chain of Arg20 favors pyrimidine over purine nucleotides at positions 7′ and 8′. Further protein:DNA contacts involve less specific interactions between HiNmlR residues in the 'wing' loop and the α2-helix, and the phosphate and sugar groups of the P*adhC-estD/nmlR(wt)* molecule (Figure 3A,B). Interestingly, HiNmlR does not engage the central region of the DNA and the -35 elements for both P*adhC* and P*nmlR* promoters (spanning positions 2–4′ in the DNA molecule) and these, presumably, remain accessible to the cellular transcription machinery via the DNA′s major groove (Figure 3C).

Individual MerR-family proteins recognize different IR DNA sequences. Unique among MerR regulons, the -35 promoter elements for both NmlR-regulated genes reside within the IR central region (Supplementary Figure S1B). Based on our structural observations and on the high degree of conservation among members of this protein sub-family (Supplementary Figure S2A), we identified the recognition sequence for NmlR-type regulators as 8′-YYNGA**G**NNNNCTCNRR-8 (where Y is a pyrimidine, R is a purine and N is any nucleotide; central bases are underlined; bold indicates position −31). The identified IR pattern can be used to locate NmlR-type regulons in bacterial genomes (Supplementary Table S3).

Out of the four α2-helix residues in HiNmlR involved in specific interactions with the DNA (Tyr17, Arg20, Phe21 and Lys24), positions 17 and 24 are the least conserved among the MerR-family members (Supplementary Figure S2B). Conversely, these residues favor a G-C base pair at position 5′ in the HiNmlR regulon, which is usually replaced with an A-T or T-A base-pair in other MerR regulons (Supplementary Figure S1B and D). We thus employed microscale thermophoresis to test the importance of the Lys24-to-Gua5′ and the Tyr17-to-Cyt5 interactions for the DNA recognition by HiNmlR. Indeed, single (G5′C or C5G) and double (G5′C and C5G) mutations to the P*adhC-estD/nmlR* IR operator sequences (24 bp in length, Supplementary Table S1) decreased binding affinity by ∼10- and ∼190-fold, respectively, compared to the value obtained for the wild-type sequence (25.2 ± 2.33 nM) (Table 1 and Supplementary Figure S4).

Distorted DNA structure

Re-orientation of α2/α2′ upon MerR protein activation introduces considerable torsional stresses and results in the bending and untwisting of the DNA molecule—ultimately reconfiguring the sub-optimal target promoter −10 and −35 elements to an orientation favorable to gene transcription by RNAP (7,8,10,14). Surprisingly, the co-crystal structure of the P*copA* operator bound to a CueR mutant unresponsive to its cognate signaling species (Cu^+ and Ag^+) showed that the DNA molecule is slightly bent and untwisted in the repressor state (14). Likewise, the co-crystal structures obtained here reveal that the P*adhC-estD/nmlR* IR operator sequence in complex with HiNmlR adopts a bent conformation with local untwisting (Figure 4A-C). HiNmlR bends the P*adhC-estD* operator DNA and introduces a ∼21° kink between its two IR half-sites (Figure 4C), supporting the idea that the HiNmlR:DNA crystal structure represents the repressor state of the complex. Likewise, the DNA kink introduced by HiNmlR is more similar to the one introduced by the repressor mutant of CueR than to the extreme remodeling of P*copA* induced by the metal-bound wild-type protein (Figure 4C). Binding to HiNmlR reduces the distance between the two IR half-sites in P*adhC-estD/nmlR(wt)* by ∼3.0 Å. As the −35 elements of both the P*adhC-estD* operon and the P*nmlR* gene are contained within the operator IR sequence, the shortening between −10 and −35 elements induced by HiNmlR for these genes corresponds to ∼1.5 Å (or <1 bp). These observations further suggest that the structure of HiNmlR bound to the P*adhC-estD/nmlR* IR operator sequence captured in the crystals corresponds to the repressor state of the HiNmlR regulon. Despite these distortions, nucleotides at positions 1 and 1′ maintained Watson–Crick base-pairs (Figure 4D, Supplementary Figure S5B).

Structural rearrangements upon DNA binding

The structure of the DNA-free form of HiNmlR reveals the structural changes in HiNmlR associated with DNA binding (Supplementary Table S2, Figure 5A,B). Structural analyses reveal that DNA binding re-orients the major groove-interacting α2- and α2′ helices in the protein dimer, reducing the distance between these two helices from ∼39.0 to ∼35.0 Å (Figure 5A). This movement is accomplished by the inward rotation (by ∼9.5° each) of two structurally-equivalent rigid bodies pivoted on a hinge formed by the central coiled-coil residues 94'-96' and 95–97 (Figure 5B). Each rigid body in the HiNmlR dimer consists of residues 1–89 from one subunit plus residues 109′–125′ from the other subunit (and *vice versa*), and their rotation does not disrupt the protein hydrophobic core described above for the DNA-bound structure (Figure 2C).

Proposed interactions with RNAP and transcriptional control

Comparing the DNA molecule from the HiNmlR co-crystal structure to its idealized B-form suggests that binding of HiNmlR induces a slight untwisting of the promoter DNA and brings -10 and -35 boxes slightly closer to the DNA configuration expected for optimally spaced (17-bp) promoters (Figure 6A–C). Nevertheless, docking of the HiNmlR:DNA complex onto the structure of the RNAP σ4 domain suggest that the slight reconfiguration of the P*adhC-estD* operator induced by HiNmlR results in the movement of the promoter −10 box away from its interacting partner, the RNAP σ2 domain. By contrast, similar docking studies suggest that both the repressor and

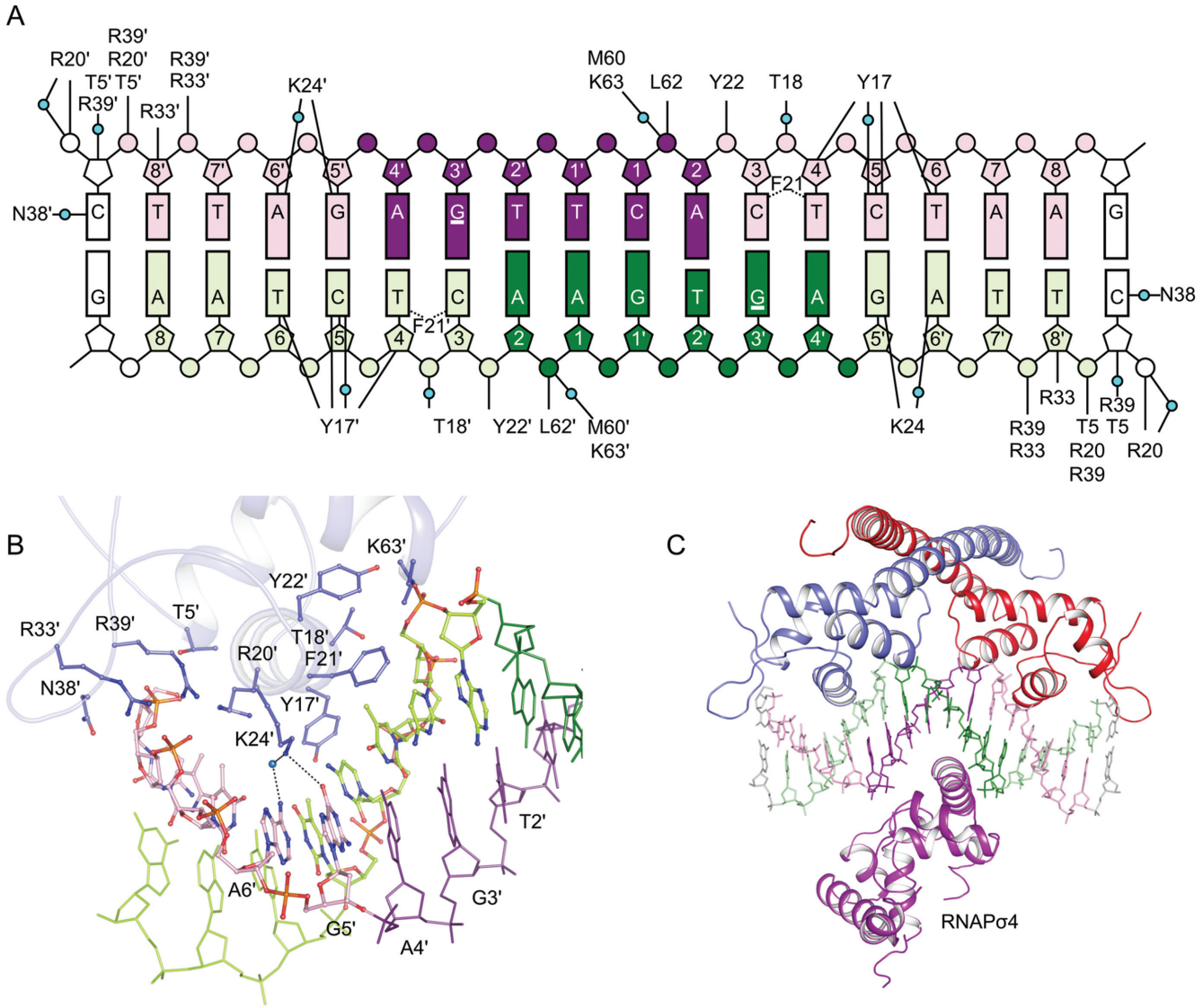

Figure 3. Protein:DNA contacts. (**A**) Schematic representation of HiNmlR:P*adhC-estD/nmlR* IR operator (P*adhC-estD/nmlR(wt)*) contacts. Protein:DNA hydrogen bonds are shown by solid lines connecting the DNA to the indicated amino-acids; water-mediated hydrogen bonds are indicated by a solid line broken by a blue sphere. Stacking between aromatic residues and nucleotide bases is represented as dashed lines. DNA bases are depicted as boxes, sugar groups as pentagons and phosphate groups as circles. Underlined nucleotides occupy the −31 position, which is important for RNAP binding. (**B**) Details of the protein:DNA interactions depicted in panel A. Hydrogen bonds between Lys24′ and DNA bases A6′ and G5′ are indicated by dashed lines. (**C**) Modeling of the interaction between the HiNmlR:DNA complex and the RNAP σ4 domain. The nucleotides occupying positions -35 to -30 in the structure of the DNA-bound σ^A domain 4 from *Thermus aquaticus* RNAP (PDB ID 1KU7 (46), shown as a purple ribbon) were superimposed onto the equivalent region in the P*adhC-estD* operon template strand (positions 2–4'—shown in purple stick representation) present in the HiNmlR:P*adhC-estD/nmlR* complex structure. DNA coloring and numbering as in Figure 2A.

activator states of CueR rotate the −10 box of its target P*copA* promoter towards the RNAP σ2 domain (Figure 6D). These discrepancies can be explained by the atypical architecture of the HiNmlR regulon. The −35 element in the P*adhC-estD* promoter is embedded within the HiNmlR recognition site and, as a consequence, the protein binds directly opposite to the RNAP σ4 domain. On the other hand, CueR and other MerR proteins bind at the center of the promoter element and opposite to RNAP β subunit, which is found between $\sigma 2$ and $\sigma 4$ in the *holo* enzyme structure (Figure 6E and F). These docking studies further support the idea that the HiNmlR co-crystal structure obtained here represents the repressor state of the regulon.

DISCUSSION

MerR-family proteins regulate transcription by DNA untwisting (12,39,40). Recently, Philips and co-workers showed that the allosteric signal in the metal regulator CueR propagates from the protein C-terminal metal-binding site to the hinge loop, triggering the re-orientation of the protein DNA-binding domains to remodel the operator DNA and allow transcription by RNAP (14).

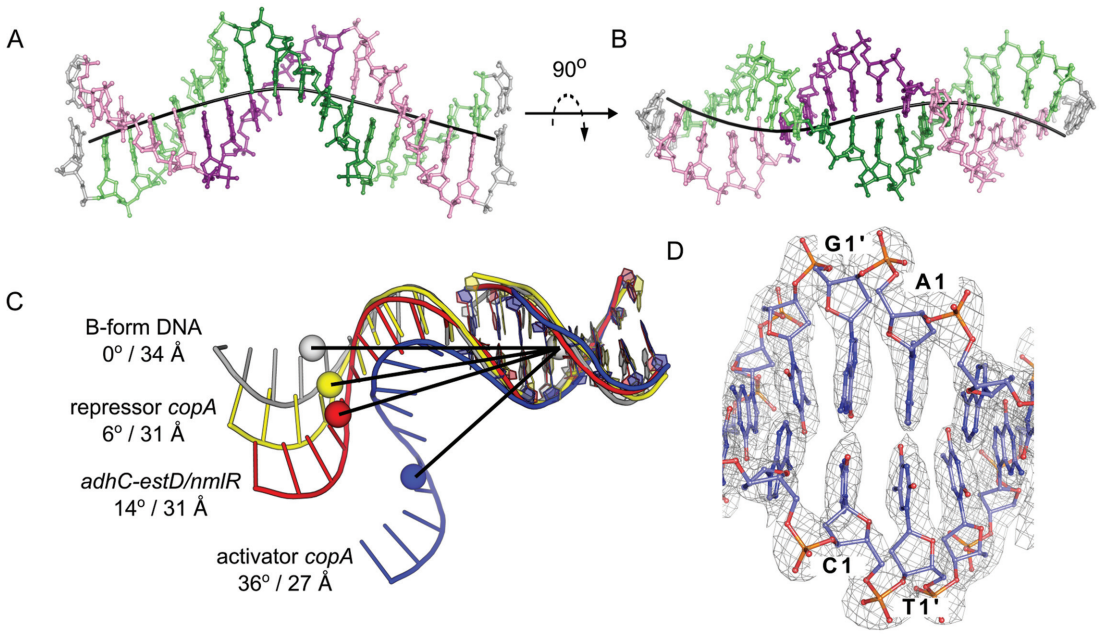

Figure 4. HiNmlR-induced distortion of the P*adhC-estD/nmlR* IR operator (P*adhC-estD/nmlR(wt)*) DNA. (**A**) 'Side' and (**B**) 'top' views of the overall structure of the P*adhC-estD/nmlR* IR operator DNA (stick representation). Nucleotides are colored as in Figure 2. The helical axis is shown as a black line. (**C**) Comparison of DNA-bending angles and operator shortening (indicated) for B-DNA (gray), the repressor (yellow) and activator (blue) P*copA* complexes and the P*adhC-estD* operator (red). IR half-sites were used for superposition. (**D**) Simulated annealing omit electron density map (white mesh) corresponding to the central nucleotides (positions 1 and 1') in the P*adhC-estD/nmlR* IR operator DNA (P*adhC-estD/nmlR(wt)*, contoured at 1.5σ). Nucleotides are shown in stick representation. The whole DNA molecule was omitted from map calculation.

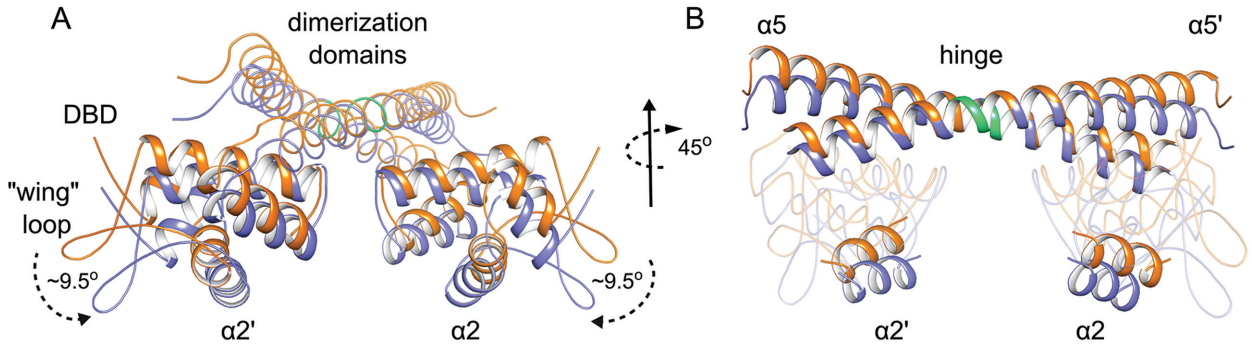

Figure 5. Structural rearrangements in HiNmlR upon DNA binding. Differences in DNA-binding (**A**) and dimerization (**B**) domains between DNA-bound (blue) and DNA-free (orange) HiNmlR structures following superposition of 'hinge' residues (94′–96′ and 95–97, shown in green). (**A**) DBDs are shown in ribbon representation. Dimerization domains are shown as loops and faded out for clarity. (**B**) Dimerization domains and DNA major groove-interacting α2 and α2'-helices are shown as ribbons; the remainder of DBDs is shown as loops and faded out for clarity.

Prior to these studies, DNA-bound structures of activated BmrR and SoxR also suggested transcriptional regulation by MerR proteins relied on intimate contacts between the protein DNA-binding and C-terminal sensor regions (4,7,8,10). Our data suggest that HiNmlR, a MerR protein lacking a recognizable sensor region, activates gene expression by a similar DNA untwisting mechanism. But in HiNmlR, modification of conserved cysteine residues located in the helices α3 (Cys54) and α4 (Cys71) by reactive species generated during the organism's response to formaldehyde are likely to be the trigger for the conformational changes leading to the activation of the target P*adhC-estD* operon.

The facile chemistry of thiol groups makes Cys residues well-suited to function as sensors for changes in the redox state of a cellular environment, and many transcriptional regulators in bacteria display 'sensing' Cys residues (17,18), including OxyR - the major bacterial hydrogen peroxide sensor (41). In this protein, the sensing Cys residues (Cys199 and Cys208) make up a redox center that is able to respond to a range of distinct stresses from reactive oxygen and nitrogen species. Different Cys modifications, including the formation of sulfenic acids, S-nitrosothiols and glutathione mixed disulfides, have been shown to form on Cys199 of OxyR, with each resulting in functionally distinct structure and transcriptional activity (41). *E. coli* contains a single cysteine-containing formaldehyde sensor FrmR, a member of the CsoR family, but the formaldehyde-modified form has not been identified (42,43).

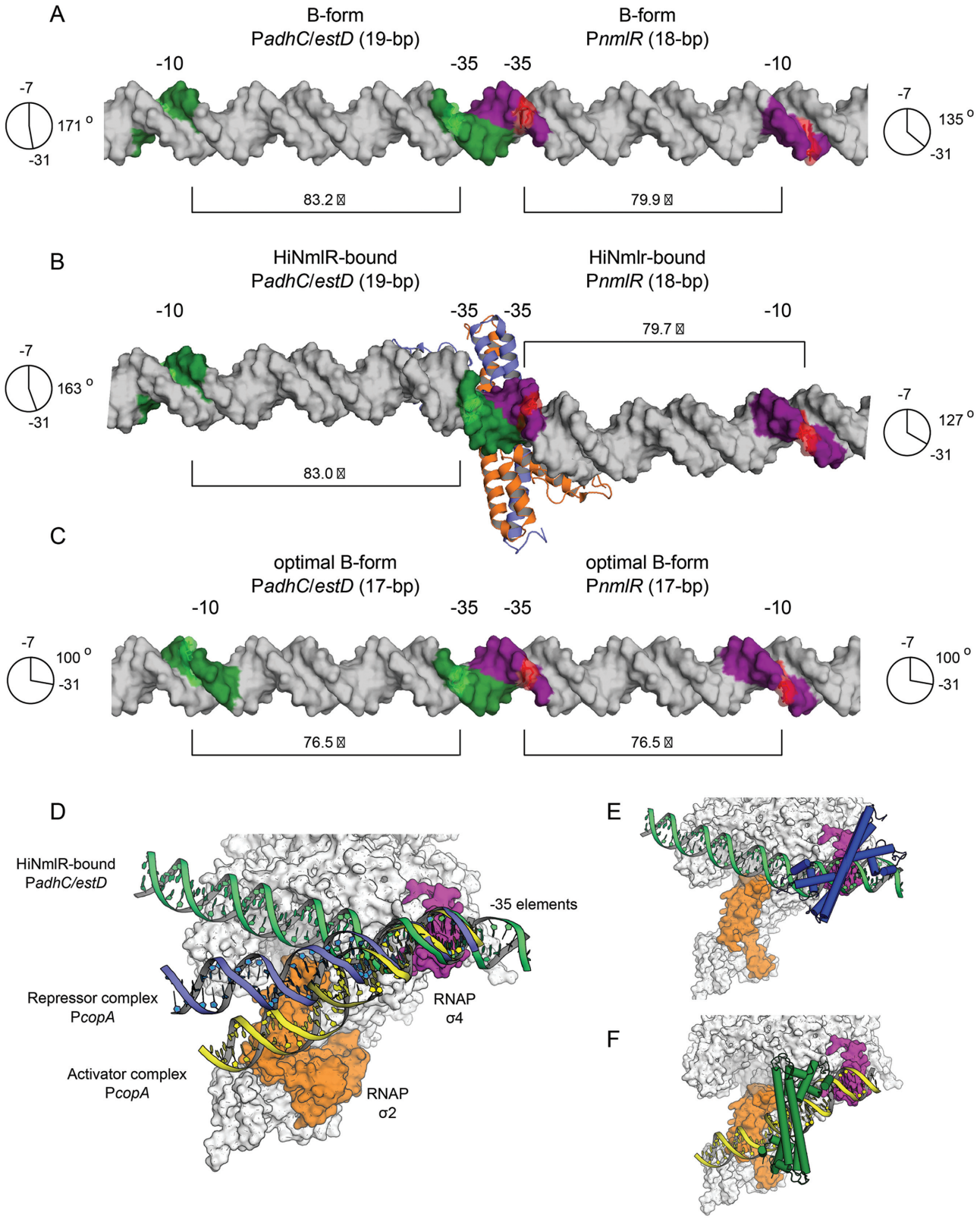

Figure 6. DNA remodelling by HiNmlR and transcriptional control by MerR proteins. (**A**) B-DNA model of P*adhC-estD* and P*nmlR*. (**B**) Repressor state of the HiNmlR regulon. The DNA from the crystal structure is extended with ideal B-DNA (gray) to include the −10 elements from both P*adhC-estD* and P*nmlR*. (**C**) B-form DNA model for optimally-spaced (17-bp) promoters in the HiNmlR regulon. In panels *A* to *C*, promoter elements for P*adhC-estD* are shown in green and for P*adhC* in purple. Radial plots indicate relative positions of -10 and -35 nucleotides that directly interact with RNAP σ4 and σ2 domains (highlighted). (**D**) Placement of HiNmlR-bound P*adhC-estD* (extended with ideal B-DNA) onto RNAP. CueR:P*copA* repressor and activator complexes were also modeled for comparison. (**E**) and (**F**) Modeling of ternary complexes show that HiNmlR (E) and CueR (F) bind opposite to different regions of RNAP.

As for OxyR, the activation mechanism of NmlR proteins is likely to be complex and how HiNmlR uses Cys residues to sense formaldehyde remains unclear. Neither expression of the recombinant protein at low concentrations of formaldehyde in the culture medium (Supplementary Figure S7), nor incubation of purified HiNmlR with low concentrations of formaldehyde or formaldehyde and glutathione, resulted in any modifications on the protein detectable by mass spectrometry (Supplementary Figure S8). In the cell, modifications to HiNmlR may be transient and quickly resolved, or formaldehyde may only interact with the protein indirectly, for example through formaldehyde carriers such as glutathione or tetrahydrofolate (18).

Nevertheless, our *in vivo* results show that HiNmlR requires Cys54 for the switch from repressor to activator states, with Cys71 and Cys95 possibly also playing a role in this transition. Residue Cys54 is absolutely conserved in the NmlR sub-family of MerR regulators and is required for transcriptional activation by the HiNmlR homolog in *B. subtilis*, AdhR (44). The structural data presented here reveals that Cys54, Cys71 and Cys95 locate to regions of the regulator molecule shown important for the conformational changes involved in the activation of other MerR proteins (12,39,40). Residue Cys71 locates to the HiNmlR helix α4 and is in close proximity to the hinge loop, which is thought to control the orientation of the DBD in relation to the dimerization helix α5 (14). The thiol group of Cys71 is also shielded from the solvent by the charged side-chains of Glu81 and Arg67 and these groups may play a role in stabilizing thiol negative charges and facilitating the attachment and resolution of thiol adducts. Residue Cys54 locates to the interface between the DBD and the dimerization helix. Modification of this residue would likely disturb the tight packing observed for this region of the molecule and change the observed configuration of DBD and the dimerization helix. Finally, Cys95 locates to middle of the dimerization coiled-coil, a region shown to accommodate some of the protein structural changes observed following binding to DNA. Despite being mostly inaccessible to the solvent in both DNA-bound and DNA-free structures, residues Cys54 and Cys71 were alkylated following an iodoacetamide treatment (in the absence of DNA), possibly reflecting the dynamic nature of HiNmlR structure in these regions. Previous work has also shown that HiNmlR counterparts in *B. subtilis* (AdhR) and *S. pneumoniae* can be S-nitrosylated *in vitro* at their (only) cysteine residue—equivalent to Cys54 in HiNmlR (45).

Regardless of their activation state, MerR proteins form a ternary complex with DNA and RNAP. The atypical NmlR regulon architecture has the regulator-recognized IR sites flanking the promoter −35 element, which, in turn, must interact with the RNAP σ4 domain. The structures presented here reveal that HiNmlR makes specific interactions with the Cyt nucleotide at position 3′ of the target IR DNA sequence, whereas RNAP σ4 domain interacts with the opposing guanine nucleotide (corresponding to position −31 within the -35 promoter sequence) (46). Thus, NmlR proteins bind their target DNA directly opposite to the RNAP σ4 domain. Binding sites for other MerR proteins are found in the spacer region between -10 and -35 elements, thus, in the ternary complex, these regulators bind opposite to the RNAP β subunit, which is found between the σ4 and σ2 domains. Thus, although the DNA remodeling introduced by CueR and HiNmlR in their resting states are similar, the different placement of HiNmlR on the polymerase-DNA complex results in the P*adhC-estD* -10 element resting much further away from the RNAP σ2 domain than the equivalent region in the P*copA* promoter. Moreover, the DNA remodeling observed in the CueR activator complex would not be enough to align the P*adhC-estD* -10 element with the RNAP σ2 domain. Obtaining the structure of the HiNmlR activator complex will shed light on how the modification of cysteine residues can translate to the large DNA untwisting and bending required to allow the interaction between P*adhC-estD* -10 box and RNAP.

Mutagenesis coupled with binding experiments confirms the relevance of protein:DNA contacts identified in our structure and helps us rationalize residue conservation of both NmlR-type proteins and cognate IR DNA targets. Residues in the α2 helix that contact the DNA (especially Tyr17, Arg20, Phe21 and Lys24) are highly conserved among the NmlR-type proteins, as are positions 5′ (Gua) on the IR operator sequence of NmlR regulons. This pattern of residue/nucleotide conservation observed for NmlR-type regulons and regulators is slightly different from the ones found for other MerR-family members and their target operator sequences. Such differences may be due to the presence of the embedded −35 promoter elements within NmlR IR operator sequences, which appear to be restricted to NmlR-type regulons within the MerR family (10,16).

We detected basal levels of AdhC activity in *H. influenzae* strains harbouring the wild-type NmlR protein and the *nmlR*-null mutants even without formaldehyde challenge. Considering that HiNmlR, like other MerR family proteins, is thought to act as a repressor in its resting state, the low basal level of activity in the wild-type bacterium suggests that low levels of activation of HiNmlR occurred, possibly due to background levels of formaldehyde being produced by the cell (18). On the other hand, the low basal levels of AdhC activity observed in *nmlR*-null mutant cells indicate that RNAP can transcribe from the P*adhC-estD* operon in the absence of HiNmlR. This 'leaky' transcription may be due to the RNAP from *H. influenzae* being less sensitive to sub-optimal promoter spacer length, as compared with the *E. coli* enzyme. Interestingly, binding of *Thermus aquaticus* RNAP σ4 domain to a canonical −35 element induced DNA untwisting to a similar level to the one observed here for *apo*-HiNmlR (46), and highlights the inherent flexibility of protein:DNA complexes. Alternatively, in the absence of HiNmlR, another transcriptional regulator may bind to the P*adhC-estD* promoter region and facilitate basal level transcription from this operon. There are at least two paralogs of HiNmlR in the *H. influenzae* genome (KEGG entries HI0293 [CueR] and HI1623 [NimR]). Both lack the conserved cysteine residues found in NmlR proteins.

In summary, we show that HiNmlR, a MerR-family regulator lacking a C-terminal sensor region, is a thiol-dependent transcription factor that modulates *H. influenzae* response to formaldehyde and contains reactive Cys residues. Our structural studies of both DNA-free and DNA-bound forms of HiNmlR suggest that despite the lack of the sensor region, the protein regulates transcrip-

tion through a DNA distortion mechanism similar to other MerR-family regulators.

ACCESSION NUMBERS

Protein Data Bank: 5D8C, 5E01 and 5D90.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Australian Synchrotron beamline scientists for the help with X-ray data collection, and members of Kobe lab for valuable suggestions. We acknowledge the use of the UQ-ROCX Facility and the assistance of Karl Byriel and Gordon King. We thank Chris Wood, Amanda Nouwens and Alun Jones for help with mass spectrometry experiments.

FUNDING

Australian Research Council (ARC) [DP0986578 to A.G.M.]; National Health and Medical Research Council (NHMRC) Program [565526 to A.G.M. and B.K. and 1071659 to B.K.]; B.K. is a NHMRC Research Fellow [1003326 and 1110971]. Funding for open access charge: NHMRC.
Conflict of interest statement. None declared.

REFERENCES

1. Galperin,M.Y. (2006) Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J. Bacteriol.*, **188**, 4169–4182.
2. Brown,N.L., Stoyanov,J.V., Kidd,S.P. and Hobman,J.L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol. Rev.*, **27**, 145–163.
3. McEwan,A.G., Djoko,K.Y., Chen,N.H., Counago,R.L., Kidd,S.P., Potter,A.J. and Jennings,M.P. (2011) Novel bacterial MerR-like regulators their role in the response to carbonyl and nitrosative stress. *Adv. Microb. Physiol.*, **58**, 1–22.
4. Kumaraswami,M., Newberry,K.J. and Brennan,R.G. (2010) Conformational plasticity of the coiled-coil domain of BmrR is required for bmr operator binding: the structure of unliganded BmrR. *J. Mol. Biol.*, **398**, 264–275.
5. Changela,A., Chen,K., Xue,Y., Holschen,J., Outten,C.E., O'Halloran,T.V. and Mondragon,A. (2003) Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science*, **301**, 1383–1387.
6. Chang,C.C., Lin,L.Y., Zou,X.W., Huang,C.C. and Chan,N.L. (2015) Structural basis of the mercury(II)-mediated conformational switching of the dual-function transcriptional regulator MerR. *Nucleic Acids Res.*, **43**, 7612–7623.
7. Heldwein,E.E. and Brennan,R.G. (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature*, **409**, 378–382.
8. Newberry,K.J. and Brennan,R.G. (2004) The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. *J. Biol. Chem.*, **279**, 20356–20362.
9. Godsey,M.H., Baranova,N.N., Neyfakh,A.A. and Brennan,R.G. (2001) Crystal structure of MtaN, a global multidrug transporter gene activator. *J. Biol. Chem.*, **276**, 47178–47184.
10. Watanabe,S., Kita,A., Kobayashi,K. and Miki,K. (2008) Crystal structure of the [2Fe-2S] oxidative-stress sensor SoxR bound to DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 4121–4126.
11. O'Halloran,T. and Walsh,C. (1987) Metalloregulatory DNA-binding protein encoded by the merR gene: isolation and characterization. *Science*, **235**, 211–214.
12. Frantz,B. and O'Halloran,T.V. (1990) DNA distortion accompanies transcriptional activation by the metal-responsive gene-regulatory protein MerR. *Biochemistry*, **29**, 4747–4751.
13. Outten,C.E., Outten,F.W. and O'Halloran,T.V. (1999) DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in Escherichia coli. *J. Biol. Chem.*, **274**, 37517–37524.
14. Philips,S.J., Canalizo-Hernandez,M., Yildirim,I., Schatz,G.C., Mondragon,A. and O'Halloran,T.V. (2015) Allosteric transcriptional regulation via changes in the overall topology of the core promoter. *Science*, **349**, 877–881.
15. O'Halloran,T.V., Frantz,B., Shin,M.K., Ralston,D.M. and Wright,J.G. (1989) The MerR heavy metal receptor mediates positive activation in a topologically novel transcription complex. *Cell*, **56**, 119–129.
16. Kidd,S.P., Potter,A.J., Apicella,M.A., Jennings,M.P. and McEwan,A.G. (2005) NmlR of Neisseria gonorrhoeae: a novel redox responsive transcription factor from the MerR family. *Mol. Microbiol.*, **57**, 1676–1689.
17. Luebke,J.L. and Giedroc,D.P. (2015) Cysteine sulfur chemistry in transcriptional regulators at the host-bacterial pathogen interface. *Biochemistry*, **54**, 3235–3249.
18. Chen,N.H., Djoko,K.Y., Veyrier,F.J. and McEwan,A.G. (2016) Formaldehyde Stress Responses in Bacterial Pathogens. *Front. Microbiol.*, **7**, 257.
19. Mason,R.P., Sanders,J.K., Crawford,A. and Hunter,B.K. (1986) Formaldehyde metabolism by Escherichia coli. Detection by in vivo 13C NMR spectroscopy of S-(hydroxymethyl)glutathione as a transient intracellular intermediate. *Biochemistry*, **25**, 4504–4507.
20. Eschenfeldt,W.H., Lucy,S., Millard,C.S., Joachimiak,A. and Mark,I.D. (2009) A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods Mol. Biol.*, **498**, 105–115.
21. Studier,F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.*, **41**, 207–234.
22. Kabsch,W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
23. Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 760–763.
24. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
25. Terwilliger,T.C., Grosse-Kunstleve,R.W., Afonine,P.V., Moriarty,N.W., Zwart,P.H., Hung,L.W., Read,R.J. and Adams,P.D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.*, **64**, 61–69.
26. 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.
27. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
28. Karplus,P.A. and Diederichs,K. (2012) Linking crystallographic model and data quality. *Science*, **336**, 1030–1033.
29. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
30. 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.
31. Krissinel,E. and Henrick,K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, **372**, 774–797.
32. Baker,N.A., Sept,D., Joseph,S., Holst,M.J. and McCammon,J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 10037–10041.

33. Dolinsky,T.J., Czodrowski,P., Li,H., Nielsen,J.E., Jensen,J.H., Klebe,G. and Baker,N.A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.*, **35**, W522–W525.
34. Hayward,S. and Berendsen,H.J. (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins*, **30**, 144–154.
35. Mrazek,J. and Xie,S. (2006) Pattern locator: a new tool for finding local sequence patterns in genomic DNA sequences. *Bioinformatics*, **22**, 3099–3100.
36. Kidd,S.P., Jiang,D., Jennings,M.P. and McEwan,A.G. (2007) Glutathione-dependent alcohol dehydrogenase AdhC is required for defense against nitrosative stress in Haemophilus influenzae. *Infect. Immun.*, **75**, 4506–4513.
37. Kidd,S.P., Jiang,D., Tikhomirova,A., Jennings,M.P. and McEwan,A.G. (2012) A glutathione-based system for defense against carbonyl stress in Haemophilus influenzae. *BMC Microbiol.*, **12**, 159.
38. Chen,N.H., Counago,R.M., Djoko,K.Y., Jennings,M.P., Apicella,M.A., Kobe,B. and McEwan,A.G. (2013) A glutathione-dependent detoxification system is required for formaldehyde resistance and optimal survival of Neisseria meningitidis in biofilms. *Antioxid. Redox Signal.*, **18**, 743–755.
39. Ansari,A.Z., Chael,M.L. and O'Halloran,T.V. (1992) Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature*, **355**, 87–89.
40. Ansari,A.Z., Bradner,J.E. and O'Halloran,T.V. (1995) DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature*, **374**, 371–375.
41. Kim,S.O., Merchant,K., Nudelman,R., Beyer,W.F. Jr, Keng,T., DeAngelo,J., Hausladen,A. and Stamler,J.S. (2002) OxyR: a molecular code for redox-related signaling. *Cell*, **109**, 383–396.
42. Herring,C.D. and Blattner,F.R. (2004) Global transcriptional effects of a suppressor tRNA and the inactivation of the regulator frmR. *J. Bacteriol.*, **186**, 6714–6720.
43. Higgins,K.A. and Giedroc,D. (2014) Insights into protein allostery in the CsoR/RcnR family of transcriptional repressors. *Chem. Lett.*, **43**, 20–25.
44. Nguyen,T.T., Eiamphungporn,W., Mader,U., Liebeke,M., Lalk,M., Hecker,M., Helmann,J.D. and Antelmann,H. (2009) Genome-wide responses to carbonyl electrophiles in Bacillus subtilis: control of the thiol-dependent formaldehyde dehydrogenase AdhA and cysteine proteinase YraA by the MerR-family regulator YraB (AdhR). *Mol. Microbiol.*, **71**, 876–894.
45. Reeves,B.D., Joshi,N., Campanello,G.C., Hilmer,J.K., Chetia,L., Vance,J.A., Reinschmidt,J.N., Miller,C.G., Giedroc,D.P., Dratz,E.A. *et al.* (2014) Conversion of S-phenylsulfonylcysteine residues to mixed disulfides at pH 4.0: utility in protein thiol blocking and in protein-S-nitrosothiol detection. *Org. Biomol. Chem.*, **12**, 7942–7956.
46. Campbell,E.A., Muzzin,O., Chlenov,M., Sun,J.L., Olson,C.A., Weinman,O., Trester-Zedlitz,M.L. and Darst,S.A. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell*, **9**, 527–539.