ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

Structural Basis of Selective Aromatic Pollutant Sensing by the Effector Binding Domain of MopR, an NtrC Family Transcriptional Regulator

Shamayeeta Ray,[†] Menachem J. Gunzburg,[‡] Matthew Wilce,[‡] Santosh Panjikar,[*,‡,§] and Ruchi Anand[*,∥,⊥]

[†]IITB-Monash Research Academy, Mumbai 400076, Maharashtra, India

[‡]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia

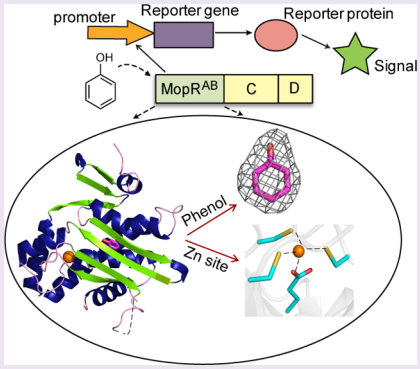
[§]Australian Synchrotron, Clayton, Victoria 3168, Australia

[∥]Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, Maharashtra, India

[⊥]Wadhwani Research Center for Bioengineering, IIT Bombay, Mumbai 400076, India

Supporting Information

ABSTRACT: Phenol and its derivatives are common pollutants that are present in industrial discharge and are major xenobiotics that lead to water pollution. To monitor as well as improve water quality, attempts have been made in the past to engineer bacterial *in vivo* biosensors. However, due to the paucity of structural information, there is insufficiency in gauging the factors that lead to high sensitivity and selectivity, thereby impeding development. Here, we present the crystal structure of the sensor domain of MopR ($MopR^{AB}$) from *Acinetobacter calcoaceticus* in complex with phenol and its derivatives to a maximum resolution of 2.5 Å. The structure reveals that the N-terminal residues 21−47 possess a unique fold, which are involved in stabilization of the biological dimer, and the central ligand binding domain belongs to the "nitric oxide signaling and golgi transport" fold, commonly present in eukaryotic proteins that bind long-chain fatty acids. In addition, $MopR^{AB}$ nests a zinc atom within a novel zinc binding motif, crucial for maintaining structural integrity. We propose that this motif is crucial for orchestrated motions associated with the formation of the effector binding pocket. Our studies reveal that residues W134 and H106 play an important role in ligand binding and are the key selectivity determinants. Furthermore, comparative analysis of MopR with XylR and DmpR sensor domains enabled the design of a MopR binding pocket that is competent in binding DmpR-specific ligands. Collectively, these findings pave way towards development of specific/broad based biosensors, which can act as useful tools for detection of this class of pollutants.



Phenol and its derivatives are harmful to both human and aquatic life as they are carcinogenic, mutagenic, and embryotoxic in nature, and their exposure, even in small quantities, can be lethal. Many soil bacteria such as *Pseudomonas* sp. possess regulatory pathways that can degrade these xenobiotics, produced through both natural and man-made industrial discharge.[1,2] In the face of a hostile environment, catabolic degradation of these compounds that can be utilized as carbon sources is stimulated *via* activation of highly regulated bacterial transcription factors, which include XylR, DmpR, and MopR.[2,3] On the basis of these proteins, extensive efforts have been devoted over the past two decades towards the design of sensitive *in vivo* biosensors.[4−6] Random mutagenesis and domain swapping experiments in the sensor region were also carried out, and hybrid sensors with varying sensitivity were created.[7−10] However, owing to the poor solubility of these proteins, they could not be effectively purified in an *in vitro* setting. Therefore, structure determination remained elusive, and exact selectivity determinants remained undetermined. Few homology models have been proposed; but the sensor domain exhibits very low sequence similarity (∼12%) to the characterized templates, posing limitations and leading to inaccuracy.[11,12] Hence, the architecture of the aromatic pollutant binding pocket and the role of peripheral residues in conferring selectivity and sensitivity in these regulators remain hard to pin down, and the field is faced with stagnation with respect to the development of broad based/specific biosensors. Another widely used approach for detection and degradation of aromatic pollutants from wastewater is employing semiconductor based electrochemical and photocatalytic oxidation methods, where xenobiotics like phenol and its derivatives can be sensed and degraded. However, these methods also have selectivity and sensitivity limitations, and ongoing efforts are underway to improve their performance.[13−20]

Received: January 9, 2016
Accepted: June 6, 2016
Published: June 30, 2016

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357−2365

XylR, DmpR, and MopR are mechanoenzymes possessing a modular architecture and belong to the NtrC superfamily of bacterial enhancer binding proteins (bEBPs;[21−23] Figure 1).

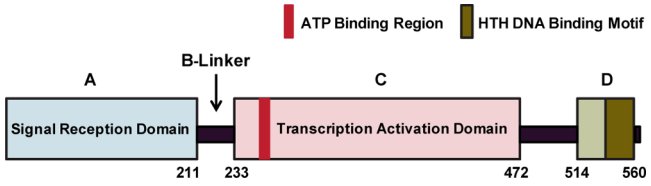


Figure 1. Domain organization of NtrC family proteins. The signal reception (A) domain is colored light blue. The central transcription activation (C) domain is colored pink. The A and C domains are joined by the B linker. The ATP binding region in the C domain is colored red. The DNA-binding (D) domain is colored light green with the HTH DNA-binding motif colored bottle green.

The bEBPs generally stimulate transcription of downstream genes by binding to specific enhancer like DNA elements known as upstream activator sequences (UAS), situated about 100 to 200 bp ahead of the promoter they activate.[21,22] In the presence of appropriate environmental cues, the bEBPs along with DNA bending integration host factors activate the σ54 RNA polymerase (RNAP) complex.[4,10,24] These bEBPs are composed of three major domains, the N-terminal signal reception domain, the central ATPase domain, and the C-terminal DNA binding domain[25] (Figure 1). The N-terminal signal reception domain, or the A-domain, comprises of around 210 amino acids and directly responds to external signals.[22] This domain is comparatively less conserved among the different NtrC family members (~12%).[23] The central transcriptional activation domain, or the C-domain, is the most conserved across the family (~40%) and comprises the ATPase domain (AAA+ domain) that provides energy required for the activation of the RNAP-σ54-promoter complex.[26] The A and the C domains are connected by a short 20-residue linker (B linker).[27] Previous biochemical studies have shown that this linker region is crucial in mediating movement of the A domain on ligand binding and facilitating derepression of the C domain.[28−30] The C terminal DNA binding domain or the D domain contains a helix−turn−helix motif that facilitates the latching of the protein onto the UAS; it is proposed to possess a structure similar to the FIS protein (PDB ID: 1FIA).[22]

Here, we present the crystal structure of the A domain and the B linker of MopR (MopR^{AB}). The structure in complex with various aromatic alcohols not only helps in mapping the residues involved in recognition of the sensor molecule but also sheds light on the structural basis of ligand specificity. To decipher the mode of selection of the effector and the stabilization provided by individual residues, mutagenesis of

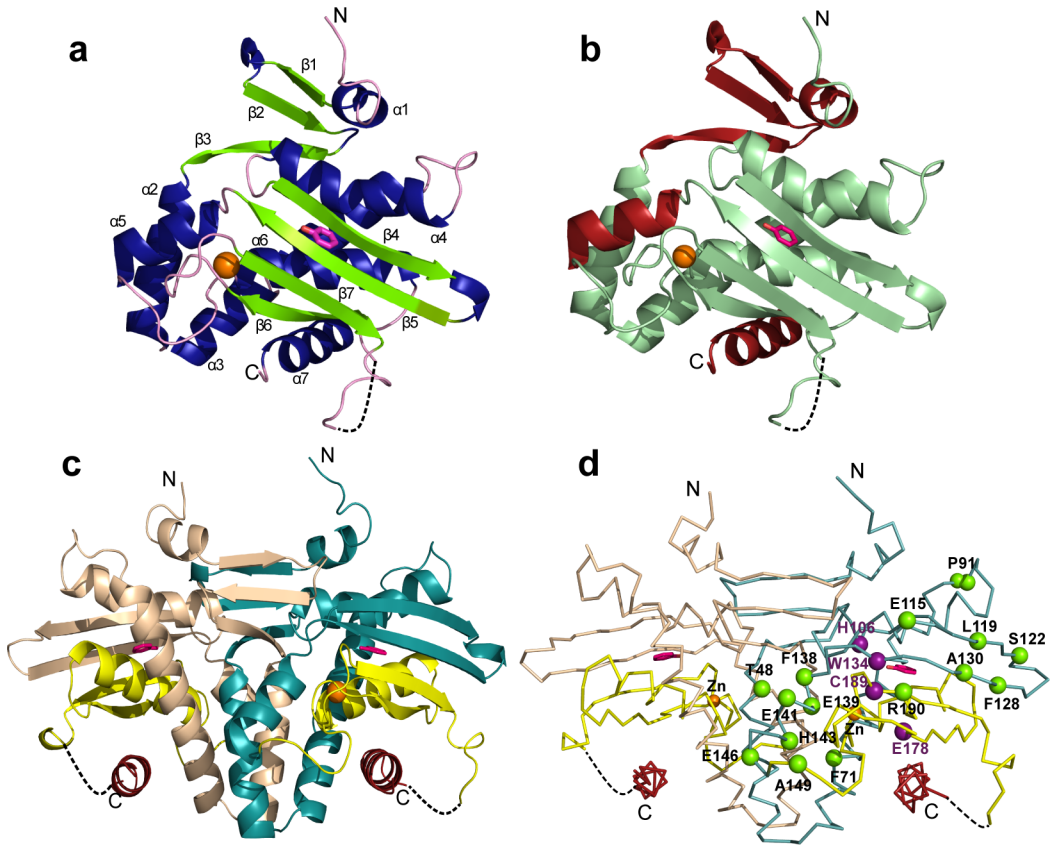


Figure 2. Structure of the MopR^{AB}−phenol complex. (a) MopR^{AB} monomer structure highlighting the secondary structural elements, with β-sheets in lightgreen, α-helices in blue, and the loop regions in pink. (b) MopR^{AB} monomer highlighting the unique regions in red. The rest of the monomer represented in pale green belongs to the "ligand binding domain involved in nitric oxide signaling and golgi transport" fold. (c) Overall structure of the MopR^{AB} dimer, with the N-terminal dimerization domain (A_N) in deep teal and wheat in both subunits, the Zn-binding domain (A_{Zn}) in yellow, and the C-terminal B-linker region in red. The missing region is represented as black dotted line. (d) Mapping of all the reported DmpR and XylR mutations (green beads) on a single monomer of the MopR^{AB} dimer and the MopR mutations (this study) as deep purple beads. The domains are colored the same as in panel c. In all panels, the bound phenol is represented as magenta stick and the Zn atom as orange sphere.

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357−2365

Table 1. Important Mutations in DmpR and XylR and Their Significance with Reference to MopR

DmpR	XylR	MopR	location	effect (as reported)	reference
F42L		T48	loop between β3 and α2 (dimer interface)	novel response to para-substituted phenols	Sarand *et al.*
	F65L	F71	α3 (dimer interface)	improved response to nitrotoluenes (NT) and cognate effectors	Garmendia *et al.*
	P85S	P91	loop between α3 and α4	transcription activation in absence of effectors	Delgado *et al.*
R109C		E115	β4 (near binding pocket)	novel response to broad range of para-substituted phenols	Sarand *et al.*
L113 V		L119	β4 (near binding pocket)	novel response to para-substituted phenols	Sarand *et al.*
D116N		S122	β4 (near binding pocket)	novel response to broad range of para-substituted phenols	Sarand *et al.*
F122L		F128	near binding pocket	novel response to para-substituted phenols at low levels	Sarand *et al.*
	V124A	A130	β5 (near binding pocket)	improved effector response profile including NTs and biphenyls	Garmendia *et al.*
F132L		F138	α5 (dimer interface)	reduced effector response profile	Skärfstad *et al.*
E133K		E139	α5 (dimer interface)	completely inactive, no effector response.	Shingler and Pavel
E135K	D135N	E141	α5 (dimer interface)	novel response to broad range of para-substituted phenols (DmpR)	Pavel *et al.*
				transcription activation in absence of effectors (XylR)	Delgado *et al.*
C137Y		H143	α5 (dimer interface)	novel response to para-substituted phenols at low levels	Sarand *et al.*
D140K		E146	α5 (dimer interface)	transcription activation in absence of effector, novel response to para-substituted phenols	Shingler and Pavel
	L143P	A149	near α5 and zinc site	improved effector response profile	Skärfstad *et al.*
	E172K	E178	zinc site	novel response to m-NT	Delgado and Ramos; Garmendia *et al.*
R184W	L184I	R190	near zinc site	improved response to 2,4-dcp, a poor DmpR effector (DmpR)	Shingler *et al.*
				improved response to NTs (XylR)	Garmendia *et al.*

selected pocket residues and subsequent ligand-binding experiments were performed. Based on the structural information, preliminary attempts to tune the selectivity of the biosensing pocket were also undertaken. Together, these studies provide a framework for the design of future biosensing devices with appropriate sensitivity and specificity.

RESULTS AND DISCUSSION

MopR^{AB} crystal structure is the first structure of a sensor domain to be solved for the XylR-NtrC subfamily. Previously, in the absence of any structure, many research groups had concentrated their efforts on homology modeling of the DmpR/XylR sensor domain.[11,12] However, due to the inaccuracy of these models and lack of ligand-complex structure, most of the predicted active site residues were not congruent with the actual MopR^{AB}–phenol binding pocket (Supporting Discussion and Supporting Information Figure S1). Moreover, over the past two decades, several mutagenesis studies in the signal-sensing domain of DmpR and XylR have been performed, backed by both *in vivo* and *in vitro* experiments[7−9,31−34] with the aim to detect the exact aromatic effector pocket. The broader aim of these studies was to design effective biosensors for degradation of the xenobiotics, but in the absence of any available structure these experiments could not adopt rational design. Hence, the crystal structure of MopR^{AB} is a break-through revelation in the field as it provides accurate details of the exact ligand-binding mode and the pocket architecture of these regulators.

The recombinant MopR^{AB} gene (residues 1−229) exists as a dimer in solution (Supporting Information Figure S2) and crystallized only in the presence of aromatic ligands. The structure of MopR^{AB} in complex with 3-chlorophenol (3-cp; PDB ID: 5KBH) was determined using zinc multiwavelength anomalous dispersion (MAD) method and was refined to 2.55 Å. Subsequently, MopR^{AB} structures in complex with phenol (PDB ID: 5KBE), o-cresol (PDB ID: 5KBG), and catechol (PDB ID: 5KBI) were determined using the difference Fourier method and refined between 2.9 and 2.5 Å (Supporting Information Table S1). The final refined structure of the MopR^{AB}–phenol complex contains residues 15−226 in subunit A and residues 14−221 in subunit B (Figure 2c).

Monomeric Structure of MopR^{AB}. The MopR^{AB} monomer consists of seven α-helices and seven β-strands (Figure 2a and Supporting Information Figure S2). The β-strands in the N-terminal region form a three-stranded antiparallel β-sheet that is sandwiched by the two helices α1 and α2. The other four-stranded antiparallel β-sheet along with α3, α4, and α6 forms a groove that serves as recognition site for the endogenous phenolic compounds; it also nests a Zn atom. It has been previously reported[7] that the ligand-binding and effector-specificity subregion mainly resides within the residues 116−192 (MopR numbering) of the sensing domain. However, the structure reveals that the actual binding pocket in the MopR^{AB} structure comprises a larger fraction of amino acids and lies between residues 99−191 (MopR numbering). Analysis of the structure using DALI revealed that the N-terminal residues 21−47 in MopR^{AB} possess a unique fold, whereas the rest of the protein is similar to the "ligand binding domain involved in nitric oxide signaling (NOS) and golgi transport" fold (Figure 2b). The top hit is an archeal protein MJ_1460 (PDB ID: 2OSO) of unknown function with an RMSD of 2.2 Å (aligning 116 residues). Surprisingly, it was observed that the previously unknown Zn site discovered in MopR^{AB} was almost structurally conserved in MJ_1460, but with one of the cystines replaced by a histidine. The other DALI hits were all eukaryotic long chain fatty acid binding proteins that lack the zinc-binding motif. The eukaryotic proteins usually exist as a part of the mega transport protein particle (TRAPP) complex involved in the transfer of proteins between organelles such as endoplasmic reticulum-to-golgi and are known to play a role in membrane localization of the complex.[35] Overall, the analysis reveals that although MopR^{AB} has very little sequence identity (<12%) with all of them, the NOS portion of the fold is essentially conserved in all of them. A striking fact that emerged from comparison of MopR^{AB} with its DALI hits was that the MopR^{AB} is the only protein among

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357−2365

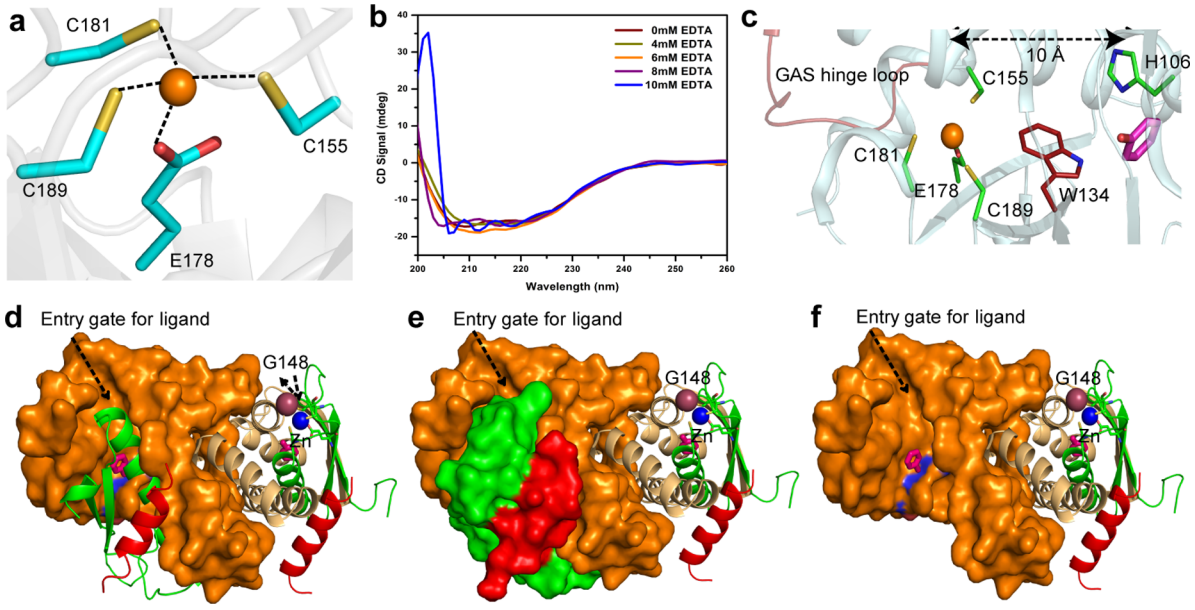


Figure 3. Zinc binding and its role in signal transduction. (a) Architecture of the Zn-binding site. Zn atom is represented as an orange sphere; the carbon atoms of the Zn site residues are in sky blue, oxygen in red, and sulfur in yellow. (b) CD data showing gradient chelation of $MopR^{AB}$ with EDTA. (c) Global view highlighting the ligand and the zinc-binding region. The conserved residue W134 that makes interaction with phenol bridges the two sites. The proposed flexible regions including the GAS hinge loop and W134 are represented in firebrick red, and the rest of the protein is in pale cyan. The zinc binding and ligand binding residues are colored as green sticks; phenol is in magenta, and zinc as orange spheres. (d–f) Proposed ligand entry gate in $MopR^{AB}$. A_N is represented in orange, A_{Zn} in green, and the B-linker region in red. Phenol is represented as a magenta stick and zinc as a blue sphere, and G148 is highlighted as a purple sphere. A_{Zn} and the linker region are represented as cartoons in panel d and as surfaces in panel e, capping the ligand. Panel f represents $MopR^{AB}$ without A_{Zn} and a linker region, which makes the ligand visible. Conformational change in A_{Zn} and the B-linker region is crucial toward opening and closing of the entry gate.

these that exists as a dimer. Closer examination of the structure shows that MopR has a unique helix insertion, α5, which is replaced by a long loop in all the other DALI hits. The helix α5 along with the exclusive β-sheet dominated N-terminal region of $MopR^{AB}$ enables it to assemble as a dimer (Figure 2b). Therefore, we believe that both these regions, unique to MopR, play an important role in defining its specific function and possibly are involved in crosstalk with the ATPase and the DNA binding domains thereby, facilitating transcription regulation.

Dimerization Interface. $MopR^{AB}$ is present as a dimer in the asymmetric unit with the interface formed mainly by the intertwining of the β3 strands from the two participating subunits (Figure 2c) with the residues 41–47 (NRMLLMH) of both subunits forming several hydrogen bonds as well as hydrophobic contacts. Furthermore, β3 interacts with the loop region between β5 and α5 of the other subunit, strengthening the dimer. The dimerization region also involves significant contacts of α5 from one monomer, with α2 of the adjacent monomer. Moreover, the loop region between β1 and β2 is in close proximity of helix α4 from the other subunit, thereby reinforcing the rigid dimer. The interface has a combined buried surface area of 2350 $Å^2$ and is stabilized by a number of hydrophobic contacts and hydrogen bonds, including three salt bridges (Supporting Information Table S2).

Many of the previously reported mutations that were shown to have a significant effect on ligand binding (Table 1, Figure 2d) were mapped onto the structure of $MopR^{AB}$, and the results show that most of them were present along the dimeric interface. For example, residues F132, E133 E135, C137, and D140 in DmpR[7,8] lie on the dimerization helix α5 that packs with the adjacent helix α2, and mutation in this region can result in disruption of the dimer. Analysis of the $MopR^{AB}$ structure clearly reveals that the D140K mutation, as reported previously,[7] results in a direct steric clash with the main chain of α2. Structural analysis also showed that E135 forms a salt bridge with the conserved residue R36 (in DmpR) that lies in β3 of the neighboring subunit (Supporting Information Table S2). It is thus not surprising that the mutation E135K[9,32] has been shown to be extremely detrimental as this mutation forces the two positively charged residues to come into close proximity, leading to destabilization of the biological dimer. *In vivo*, most of these dimerization mutants showed significant effects in downstream transcription regulation.[8] Studies on other NtrC family proteins that possess a modular architecture with a central ATPase domain have shown that upon activation, there is a major rearrangement of the overall structure leading to oligomerization of the ATPase unit to either a hexameric or a heptameric arrangement.[36,37] A similar scenario is likely in the case of MopR, where the dimeric form rearranges to a hexameric one, leading us to believe that this interface is fluid and undergoes major conformational change upon effector binding. In the light of the current structural data and in congruence with the evidence that rearrangement of the dimeric interface is a major step toward transcription activation, it is not surprising that in the absence of the structure, many of the random mutations performed along the dimeric interface by other researchers resulted in significant effects. Therefore, it appears that mutants causing destabilization of the dimeric interface weakens the transcription control exerted by the sensor domain, and the observed result is accelerated downstream activation.

Novel Zn-Binding Site. One of the most striking features of the crystal structure of $MopR^{AB}$ is the presence of Zn in both

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357–2365

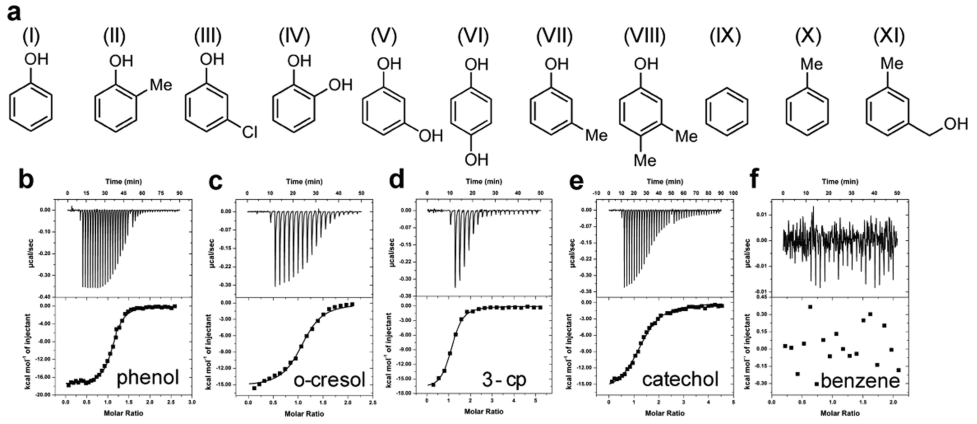


Figure 4. Binding analysis of MopR^{AB} with different aromatic effectors. (a) Structure of all the aromatic effectors used in this work: (I) phenol, (II) o-cresol, (III) 3-chlorophenol, (IV) catechol, (V) resorcinol, (VI) hydroquinone, (VII) m-cresol, (VIII) 3,4-dimethylphenol (3,4-dmp), (IX) benzene, (X) toluene, (XI) 3-methylbenzylalcohol (3-mba). (b–f) ITC data of MopR^{AB} with phenol, catechol, and benzene, respectively. The curves that correspond to raw data are shown in the top panel and the curve fit in the bottom panel. Data were fit using one set of site model, and the thermodynamic parameters obtained from the curve fitting are given in Table 2.

subunits, despite its absence in the crystallization condition or in the protein buffer. The presence of Zn was identified in MopR^{AB} crystals using X-ray fluorescence spectroscopy. The presence of Zn was exploited for the structure determination using Zn-MAD phasing. The structure reveals that the Zn atom is nested between helices α6, β6, and β7, at a distance of 10 Å from the aromatic effector, forming a rigid domain; it does not have any direct interactions with the bound ligand. This structure provides the first example, to our knowledge, of a tetrahedrally coordinated Zn bound to three cysteines and one glutamate residue (Figure 3a). This Zn-motif is conserved among a number of NtrC family members, including DmpR and XylR. Due to the unique signature of the motif, it remained unnoticed until later in multiple sequence alignments (MSAs). Furthermore, the absence of a water molecule at the fourth coordination position highlights the fact that this Zn atom does not play a catalytic role; rather, it is structural in nature.[38] The CD data of the apo MopR^{AB} show the signature ω-shaped helical pattern, which is gradually lost with increasing concentration of EDTA (Figure 3b). Unlike catalytic Zn, which generally leaches out at a concentration range of 1–3 mM, here the helicity is only completely disrupted by 10 mM EDTA. Additionally, E178A and C189A point mutations were performed to disrupt the tetrahedral coordination of Zn. These mutant proteins exhibited extremely low solubility, indicating that the presence of the intrinsic Zn is very important for maintaining the structural integrity of MopR^{AB}. Hence, the previously reported E172K mutation in XylR (E178 in MopR^{AB}) also resulted in a dramatic loss in ligand binding as the residue directly interacts with the Zn atom (Figure 3a).[31,32]

Since zinc is so unique to the XylR, DmpR, and MopR subfamily of NtrC regulators, it presumably plays a role in overall arrangement of the full-length MopR complex and in transducing the effector binding. As mentioned earlier, MopR^{AB} only crystallized in the presence of a ligand. This is probably due to the fact that a ligand-bound protein adopts a stable conformation and possesses a drastically reduced intrinsic flexibility. This leads us to hypothesize that in the apo full-length protein, the MopR^{AB} domain has an open form, and the binding of the ligand prompts the compaction of this domain. It would have been much easier to understand the conformational changes if the structure of the apo MopR^{AB} domain would have been determined. In its absence, we propose that the ligand-binding cavity is transient in nature, and the interaction of the ligand with the W134 residue triggers the Zn-binding domain (A_{Zn}: residue 148–202) to move and pack against the A_N domain (1–147). This creates the ligand-binding cavity, the so-called closed form (Figure 3c). In addition, the intertwining of β1–β3, several hydrogen bonds (including salt bridges), and hydrophobic contacts at the dimer interface provide structural stability (Supporting Information Table S2) and keep A_N more rigid against the significant movement of A_{Zn} that is induced by the ligand.

The question of how MopR^{AB} switches between the open and closed forms requires the location of a hinge region in the protein, which often would be located at the center of the motion. The structures and MSAs were used to trace this region. Relating the protein sequence to sites of structural flexibility has been previously analyzed in great detail by using a protein motion database from which a number of correlations between hinges and sequence features have been established.[39] Based on those studies, it has been predicted that the smaller residues (glycine, alanine, and serine) are preferred and hinges are likely to occur in turns or random coils and tend to be on the surface, since steric clashes would often prevent them from being in the core. The simplest hinge consists of a single point on the chain separating two rigid regions. It was observed that the turn in MopR^{AB} at the end of the A_N contains a conserved G148, followed by nonconserved A149 and S150 residues. The location, environment, and solvent accessibility of these residues suggest that they serve as the likely hinge. Moreover, it has been reported that the mutation of L143P in XylR (A149 in MopR) showed a differential effector response.[7] This highlights the fact that introducing rigidity in this region has an effect on the movement associated with A_{Zn}, which harbors part of the ligand-binding cavity (Figure 3d–f), and most likely plays a role in closing and opening of the cavity. In the full-length apo protein, A_{Zn} may be involved in repression of the AAA+ domain. Ligand binding serves as a trigger for domain movement, initiating derepression by facilitating change in the oligomerization state (dimeric to hexameric). This subsequently prompts ATP hydrolysis, leading to downstream transcriptional activation of the phenol degrading catabolic gene cluster.

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357–2365

Table 2. Binding Ananlysis of MopR^{AB} and Its Mutants to Different Aromatic Effectors

MopR^{AB} construct	aromatic compounds	ΔH (kcal/mol)	$T\Delta S$(kcal/mol)	ΔG(kcal/mol)	K_d (μM)	N (no. of sites)
native	phenol	−17.56 ± 0.13	−8.64 ± 0.18	−8.92 ± 0.06	0.46 ± 0.06	1.11 ± 0.01
native	3-chlorophenol	−17.19 ± 0.22	−8.64 ± 0.09	−8.55 ± 0.13	0.55 ± 0.06	1.07 ± 0.01
native	o-cresol	−15.37 ± 0.21	−7.15 ± 0.06	−8.22 ± 0.15	0.78 ± 0.05	1.07 ± 0.01
native	m-cresol	−20.94 ± 0.34	−13.11 ± 0.17	−7.83 ± 0.17	1.80 ± 0.03	1.02 ± 0.02
native	resorcinol	−22.69 ± 0.63	−14.95 ± 0.02	−7.74 ± 0.61	2.16 ± 0.03	1.02 ± 0.09
native	hydroquinone	−14.85 ± 0.14	−7.29 ± 0.05	−7.56 ± 0.09	3.69 ± 0.17	1.03 ± 0.06
native	3,4-dimethylphenol[a]	−3.44 ± 0.15	−0.29 ± 0.06	−3.15 ± 0.09	4.02 ± 0.01	0.95 ± 0.09
native	catechol	−8.08 ± 0.07	−0.89 ± 0.08	−7.19 ± 0.02	4.12 ± 0.08	1.34 ± 0.05
H106A	phenol[a]	−14.14 ± 0.42	−7.15 ± 0.06	−6.99 ± 0.36	7.87 ± 0.01	1.13 ± 0.03
W134A	phenol	−16.52 ± 0.23	−7.15 ± 0.03	−9.37 ± 0.11	2.87 ± 0.01	1.14 ± 0.02
W134A-H106A	phenol[a]	−10.35 ± 0.32	−4.47 ± 0.06	−5.88 ± 0.26	38.46 ± 0.13	1.04 ± 0.03

[a]ITC data for these particular runs have C values < 20; hence the computed thermodynamic parameters are not resolved accurately due to weak binding affinity.

Identification of the Cognate Ligands and Ligand-Binding Architecture. A thermal shift (Tm) assay was carried out initially with a series of ligands to distinguish the aromatic binders from the nonbinders of MopR^{AB}. The ligands that were screened included both DmpR and XylR effectors (Figure 4a).[40,41] Based on the appreciable positive shift in Tm in the melting curves, it was concluded that XylR effectors are nonbinders and MopR has a preference for phenolic compounds (Supporting Information Figure S3). Further, to calculate the binding constants and to derive the thermodynamic parameters, ITC was performed (Figure 4b−f, Supporting Information Figure S4 and Table 2).

XylR effectors again showed no binding with MopR^{AB} (Figure 4f and Supporting Information Figure S4). ITC results show that the binding is highly enthalpy-driven. This could be due to the energy released during interaction of MopR^{AB} upon binding of the effector into the pocket. Of all compounds tested, the K_d value for phenol was the lowest (0.46 ± 0.06 μM), confirming that it is indeed the best binder of MopR^{AB}. Other phenol derivatives like 3-chlorophenol, o-cresol, and m-cresol also exhibited low K_d values ranging between 0.5 and 1.8 μM (Table 2), indicating high binding affinity of MopR^{AB} toward these compounds. Few bulkier phenols like 3,4-dimethyphenol and benzenediols (catechol, resorcinol, and hydroquinone) were also tested for binding, and the affinity of all these compounds toward MopR^{AB} was found to be comparatively lower than the smaller phenol derivatives, with their K_d values ranging between 2.5 and 4.2 μM (Table 2). From this observed trend in binding affinity, it can be concluded that MopR^{AB} has evolved to accept the smaller phenolic ligands better and is less efficient in effectively accommodating the larger sized moieties. Based on the ITC results, four of the phenolic binders (phenol, o-cresol, 3-chlorophenol, and catechol) were selected for cocrystallization with MopR^{AB}. From the ligand-bound crystal structures, it was observed that the binding pocket of each MopR^{AB} subunit has a diameter of ∼5 Å and is lined mostly by hydrophobic residues (Figure 5). Stacking interactions of residues F99, F132, Y161, Y165, and Y176 with the phenyl ring of the effectors are major contributors to the stabilization of the ligand. The phenolic effector is additionally counterpoised in the active site by interactions of the phenolic OH group by direct hydrogen bonding interaction with the nitrogen atoms of H106 and W134, present in helix α4 and β5, respectively (Figure 5). Analysis of the structure reveals that the hydroxyl group of the ligand is positioned in such a way that it makes a stronger

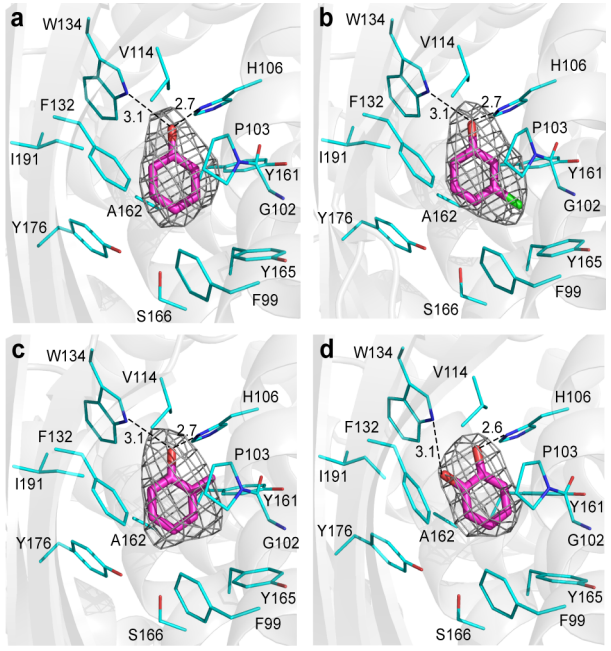


Figure 5. Representation of the binding pocket of MopR^{AB}. MopR^{AB} binding pocket in complex with (a) phenol, (b) 3-cp, (c) o-cresol, and (d) catechol. The phenolic effectors are colored in magenta, and the ligand-binding residues in sky blue. Oxygen and nitrogen atoms are in red and blue, respectively. F_o-F_c maps were contoured at 3.0σ for all of the ligands.

hydrogen bond with the nitrogen atom of H106 than with that of W134. A comparison of the various MopR^{AB}−effector complex structures revealed almost identical ligand binding conformations and pocket architectures (Supporting Information Figure S5). In the case of catechol, since it has two hydroxyl groups, there is a minor change in its orientation. The phenyl ring of catechol undergoes an 8° rotation with respect to that observed for phenol, thereby resulting in one of its OH groups interacting with H106 and the other with W134.

To decipher the importance of the structural basis of interaction of MopR^{AB} with its aromatic effectors, based on the inspection of the crystal structures, a series of single and double amino acid substitutions of the binding pocket residues were made. Residues selected for mutagenesis were H106 and W134, both of which form direct hydrogen bonds with the phenolic OH of all the ligands. The mutations H106A, W134A, and H106A-W134A were designed such that the side chains were

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357−2365

reduced to a bare minimum. The ITC results of these mutants with phenol confirm that disrupting either the H106 or W134 interaction with the ligand has a marked effect on the strength of binding (Table 2, Supporting Information Figure S6). Furthermore, phenol showed higher sensitivity toward H106A mutation (K_d increased by ~17-fold) than the W134A mutant (~6 fold increased K_d); the double mutation, W134A−H106A, resulted in a drastic change in the binding affinity (80 times poorer binding than native MopR^{AB}). In addition, to understand if the substitution of the cyclic side chains with an aliphatic one that is capable of forming hydrogen bond alters the effector−protein interaction, the residues selected above were also substituted with asparagine, and it was observed that both H106N and W134N mutations resulted in noticeably reduced affinity for phenol and the double mutant exhibiting a complete loss in binding. These observations highlight the significance of steric factors and proper positioning of the hydrogen bonding interactions that play an important role in optimization of the binding pocket interactions. Based on the above results, it is clear that H106 is the key phenolic anchor, whereas W134 acts as a sensor for the ligand binding pocket, which reports the presence of the ligand to the adjacent zinc binding domain involved in regulation of function.

Toward Design of Biosensor. Structural analysis of MopR^{AB} provides insights into how this protein sequesters specific effectors and which are the key residues involved in stabilization of the ligand. The selectivity and specificity of the MopR effectors is important for the correct activation of downstream pathways. In order to explain selective ligand specificity in DmpR and XylR, a comparative analysis of the MopR^{AB} binding pocket with DmpR and XylR was performed based on MSA (Supporting Information Figure S7). The analysis indicates that these proteins primarily have a conserved hydrophobic pocket, and most of the residues maintain an environment conducive for the particular aromatic moiety to be stabilized via stacking interactions (Figure 6a).

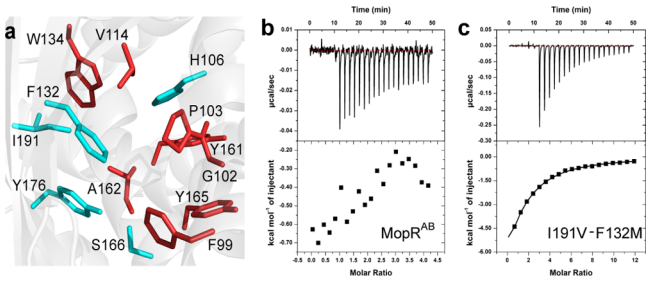


Figure 6. Structural basis of substrate specificity in MopR^{AB}. (a) Binding pocket of MopR^{AB} representing the conserved pocket residues of XylR, MopR, and DmpR, in red, and the variable residues, in cyan. (b,c) ITC data of native MopR^{AB} (b) and I191V−F132M mutant (c) with 2,5-dimethylphenol. Proteins in the sample cell were titrated against 2,5-dimethylphenol. The curves that correspond to raw data are shown in the top panel and the curve fit in the bottom panel. No fitting parameters could be evaluated for native MopR^{AB}. The mutant data were fit using one set of site model.

It appears that marginal hydrophilic nature is introduced by the residues H106, Y176, and S166 in the MopR pocket. Sequence comparison reveals that XylR lacks the key phenol anchoring residue, histidine, and most likely this results in inability of XylR to bind to any of the phenolic ligands. However, for the other phenolic effector, DmpR, this histidine residue is conserved (H100 in DmpR, H106 in MopR). Pocket lining residues S166, F132, I191, and Y176 (numbering according to MopR sequence) are variable across the three members and alteration in these residues in each of the closely related family members seems to be important for fine-tuning of specificity. It has been previously reported that DmpR exhibits better affinity toward bulkier phenolic effectors such as a few dimethylphenols.[40] Comparative analysis with MopR indicates that the effector binding pocket of DmpR has undergone subtle changes to accommodate these ligands. For example, the rigid aromatic moiety F132 is replaced by an aliphatic methoinine, which is capable of adopting several rotamers, some of which can facilitate in creating space in the binding region, and the bulkier I191 of MopR is replaced by a valine residue in DmpR. To confirm that these changes indeed help in shifting the binding profile, initial efforts to change the specificity profile of MopR^{AB} toward DmpR were undertaken, by mimicking the DmpR binding pocket. F132 and I191 in MopR^{AB} were mutated to methionine and valine, respectively. The ITC of a DmpR effector, 2,5-dimethylphenol, with the native and the mutant form reveals that native MopR^{AB} has almost no binding affinity for 2,5-dmp (Figure 6b), whereas the double mutant I191V−F132M exhibits significant affinity toward 2,5-dmp with a K_d of 11.79 ± 0.03 μM (Figure 6c). Therefore, it appears that the pocket residues are the key sensor determinants and a series of broad based biosensors can be designed by tuning MopR as a template that can sense a wide range of pollutants. The preliminary studies provide evidence toward the feasibility of designing single/multiple protein variants by engineering a ligand binding pocket for sensing a variety of aromatic effectors.

Conclusion. DmpR, XylR, and MopR serve as transcription regulators and are found in soil bacteria that are capable of degrading xenobiotics (benzene, toluene, phenol etc.). Because of the ability of these proteins to detect these pollutants specifically at very low levels, they can be used toward development of effective biosensors. Here, we described the crystal structure of the sensor domain of MopR^{AB} in complex with various aromatic effectors. The structure of the sensor domain, the first in the family, helped in characterization of the pocket architecture and led to the identification of a unique zinc site that is proposed to be important for ligand binding and in triggering downstream communication. On the basis of the structure, the key sensitivity determinants were investigated, and initial attempts to shift the specificity profile of MopR were undertaken with the broader aim of designing logic based specific/broad based biosensors. As a future direction, using the structural information of MopR^{AB}, attempts would now be made to design and engineer protein based biosensors encompassing the ATPase readout domain, which can be used to detect and quantify phenol-based as well as other aromatic pollutants in real time environmental samples.

METHODS

The MopR^{AB} gene was cloned from genomic DNA into a modified pET28a expression vector. The following site-specific mutations H106A, W134A, H106N, W134N, W134A−H106A, I191V−F132M, C189A, and E178A were performed on MopR^{AB}, and the native and mutated proteins were overexpressed and purified using standard protocols (cloning and protein purification detailed in the Supporting Information). The purified MopR^{AB} was screened using commercially available screens in the presence of phenol and other effectors at the CrystalMation (Rigaku) integrated robotic workstation of the high-throughput crystallization facility at Monash University, Australia, and

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357−2365

the crystals obtained were optimized, cryo-protected, and diffracted at the microfocus beamline (MX2) of the Australian Synchrotron. The crystal structure of the MopR^{AB}−3-chlorophenol complex was solved using the 3W-MAD protocol of the Auto-Rickshaw software,[42−44] and model building and refinement were carried out using COOT[45] and REFMAC5,[46] respectively (see Supporting Information text and Table S1 for details of crystallization, data collection, and structure determination). All figures were made in PyMOL. Detailed methods for ligand-binding experiments using ITC and Tm shift assay and circular dichroism (CD) studies are provided in the Supporting Information.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschembio.6b00020.

Supporting results, discussion, and methods; seven figures (Figures S1−S7); two tables (Tables S1 and S2); and references (PDF)

Accession Codes

Atomic coordinates have been deposited at the PDB with accession code 5KBE (MopR^{AB}-phenol), 5KBG (MopR^{AB}-o-cresol), 5KBH (MopR^{AB}-3-cp), and 5KBI (MopR^{AB}-catechol).

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: santosh.panjikar@synchrotron.org.au.

*E-mail: ruchi@chem.iitb.ac.in.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by an Australia−India Council (AIC) grant, Wadhwani Research Center for Bioengineering (WRCB), IIT Bombay, India, and DST, Government of India (Grant Number EMR/2015/002121).

■ REFERENCES

(1) Tropel, D., and van der Meer, J. R. (2004) Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol. Mol. Biol. R. 68*, 474−500.

(2) Shingler, V. (2003) Integrated regulation in response to aromatic compounds: from signal sensing to attractive behaviour. *Environ. Microbiol. 5*, 1226−1241.

(3) Timmis, K. N., and Pieper, D. H. (1999) Bacteria designed for bioremediation. *Trends Biotechnol. 17*, 201−204.

(4) Kim, M. N., Park, H. H., Lim, W. K., and Shin, H. J. (2005) Construction and comparison of Escherichia coli whole-cell biosensors capable of detecting aromatic compounds. *J. Microbiol. Methods 60*, 235−245.

(5) Gupta, S., Saxena, M., Saini, N., Mahmooduzzafar, Kumar, R., and Kumar, A. (2012) An effective strategy for a whole-cell biosensor based on putative effector interaction site of the regulatory DmpR protein. *PLoS One 7*, e43527.

(6) Xue, H., Shi, H., Yu, Z., He, S., Liu, S., Hou, Y., Pan, X., Wang, H., Zheng, P., Cui, C., Viets, H., Liang, J., Zhang, Y., Chen, S., Zhang, H. M., and Ouyang, Q. (2014) Design, construction, and characterization of a set of biosensors for aromatic compounds. *ACS Synth. Biol. 3*, 1011−1014.

(7) Skärfstad, E., O'Neill, E., Garmendia, J., and Shingler, V. (2000) Identification of an Effector Specificity Subregion within the Aromatic-Responsive Regulators DmpR and XylR by DNA Shuffling. *J. Bacteriol. 182*, 3008−3016.

(8) Sarand, I., Skärfstad, E., Forsman, M., Romantschuk, M., and Shingler, V. (2001) Role of the DmpR-Mediated Regulatory Circuit in Bacterial Biodegradation Properties in Methylphenol-Amended Soils. *Appl. Environ. Microb. 67*, 162−171.

(9) Pavel, H., Forsman, M., and Shingler, V. (1994) An aromatic effector specificity mutant of the transcriptional regulator DmpR overcomes the growth constraints of Pseudomonas sp. strain CF600 on para-substituted methylphenols. *J. Bacteriol. 176*, 7550−7557.

(10) Ramos, J. L., Marqués, S., and Timmis, K. N. (1997) Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. *Annu. Rev. Microbiol. 51*, 341−373.

(11) Devos, D., Garmendia, J., Lorenzo, V. d., and Valencia, A. (2002) Deciphering the action of aromatic effectors on the prokaryotic enhancer-binding protein XylR: a structural model of its N-terminal domain. *Environ. Microbiol. 4*, 29−41.

(12) Suresh, P. S., Kumar, R., and Kumar, A. (2010) Three Dimensional Model for N-Terminal A Domain of DmpR (2-Dimethylphenol) Protein Based on Secondary Structure Prediction and Fold Recognition. *In Silico Biol. 10*, 223−233.

(13) Wang, Y., Zhang, Y.-n., Zhao, G., Tian, H., Shi, H., and Zhou, T. (2012) Design of a Novel Cu2O/TiO2/Carbon Aerogel Electrode and Its Efficient Electrosorption-Assisted Visible Light Photocatalytic Degradation of 2,4,6-Trichlorophenol. *ACS Appl. Mater. Interfaces 4*, 3965−3972.

(14) Long, M., Cai, W., Cai, J., Zhou, B., Chai, X., and Wu, Y. (2006) Efficient Photocatalytic Degradation of Phenol over Co3O4/BiVO4 Composite under Visible Light Irradiation. *J. Phys. Chem. B 110*, 20211−20216.

(15) Belhadj Tahar, N., and Savall, A. (1999) Electrochemical degradation of phenol in aqueous solution on bismuth doped lead dioxide: a comparison of the activities of various electrode formulations. *J. Appl. Electrochem. 29*, 277−283.

(16) Li, M., Feng, C., Hu, W., Zhang, Z., and Sugiura, N. (2009) Electrochemical degradation of phenol using electrodes of Ti/RuO2−Pt and Ti/IrO2−Pt. *J. Hazard. Mater. 162*, 455−462.

(17) Wang, W., Jing, L., Qu, Y., Luan, Y., Fu, H., and Xiao, Y. (2012) Facile fabrication of efficient AgBr−TiO2 nanoheterostructured photocatalyst for degrading pollutants and its photogenerated charge transfer mechanism. *J. Hazard. Mater. 243*, 169−178.

(18) Cantalapiedra, A., Gismera, M. J., Sevilla, M. T., and Procopio, J. R. (2014) Sensitive and Selective Determination of Phenolic Compounds from Aromatic Plants Using an Electrochemical Detection Coupled with HPLC Method. *Phytochem. Anal. 25*, 247−254.

(19) Li, Z., Chen, J., Yang, J., Su, Y., Fan, X., Wu, Y., Yu, C., and Wang, Z. L. (2015) β-cyclodextrin enhanced triboelectrification for self-powered phenol detection and electrochemical degradation. *Energy Environ. Sci. 8*, 887−896.

(20) Belhadj Tahar, N., and Savall, A. (1998) Mechanistic Aspects of Phenol Electrochemical Degradation by Oxidation on a Ta/PbO2 Anode. *J. Electrochem. Soc. 145*, 3427−3434.

(21) Kustu, S., North, A. K., and Weiss, D. S. (1991) Prokaryotic transcriptional enhancers and enhancer-binding proteins. *Trends Biochem. Sci. 16*, 397−402.

(22) Morett, E., and Segovia, L. (1993) The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol. 175*, 6067−6074.

(23) Bush, M., and Dixon, R. (2012) The Role of Bacterial Enhancer Binding Proteins as Specialized Activators of σ(54)-Dependent Transcription. *Microbiol. Mol. Biol. R. 76*, 497−529.

(24) Buck, M., and Cannon, W. (1992) Specific binding of the transcription factor sigma-54 to promoter DNA. *Nature 358*, 422−422.

(25) North, A. K., Klose, K. E., Stedman, K. M., and Kustu, S. (1993) Prokaryotic enhancer-binding proteins reflect eukaryote-like modularity: the puzzle of nitrogen regulatory protein C. *J. Bacteriol. 175*, 4267−4273.

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357−2365

(26) Wikström, P., O'Neill, E., Ng, L. C., and Shingler, V. (2001) The regulatory N-terminal region of the aromatic-responsive transcriptional activator DmpR constrains nucleotide-triggered multimerisation1. *J. Mol. Biol. 314*, 971−984.
(27) Wootton, J. C., and Drummond, M. H. (1989) The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. *Protein Eng., Des. Sel. 2*, 535−543.
(28) O'Neill, E., Ng, L. C., Sze, C. C., and Shingler, V. (1998) Aromatic ligand binding and intramolecular signalling of the phenol-responsive σ54-dependent regulator DmpR. *Mol. Microbiol. 28*, 131−141.
(29) Ng, L. C., O'Neill, E., and Shingler, V. (1996) Genetic Evidence for Interdomain Regulation of the Phenol-responsive 54-dependent Activator DmpR. *J. Biol. Chem. 271*, 17281−17286.
(30) Shingler, V. (1996) Signal sensing by σ54-dependent regulators: derepression as a control mechanism. *Mol. Microbiol. 19*, 409−416.
(31) Garmendia, J., Devos, D., Valencia, A., and De Lorenzo, V. (2001) À la carte transcriptional regulators: unlocking responses of the prokaryotic enhancer-binding protein XylR to non-natural effectors. *Mol. Microbiol. 42*, 47−59.
(32) Delgado, A., Salto, R., Marqués, S., and Ramos, J. L. (1995) Single Amino Acids Changes in the Signal Receptor Domain of XylR Resulted in Mutants That Stimulate Transcription in the Absence of Effectors. *J. Biol. Chem. 270*, 5144−5150.
(33) Shingler, V., and Pavel, H. (1995) Direct regulation of the ATPase activity of the transcriptional activator DmpR by aromatic compounds. *Mol. Microbiol. 17*, 505−513.
(34) Delgado, A., and Ramos, J. L. (1994) Genetic evidence for activation of the positive transcriptional regulator Xy1R, a member of the NtrC family of regulators, by effector binding. *J. Biol. Chem. 269*, 8059−8062.
(35) Turnbull, A. P., Kümmel, D., Prinz, B., Holz, C., Schultchen, J., Lang, C., Niesen, F. H., Hofmann, K.-P., Delbrück, H., Behlke, J., Müller, E.-C., Jarosch, E., Sommer, T., and Heinemann, U. (2005) Structure of palmitoylated BET3: insights into TRAPP complex assembly and membrane localization. *EMBO J. 24*, 875−884.
(36) Batchelor, J. D., Doucleff, M., Lee, C.-J., Matsubara, K., De Carlo, S., Heideker, J., Lamers, M. H., Pelton, J. G., and Wemmer, D. E. (2008) Structure and Regulatory Mechanism of Aquifex aeolicus NtrC4: Variability and Evolution in Bacterial Transcriptional Regulation. *J. Mol. Biol. 384*, 1058−1075.
(37) Lee, S.-Y., De La Torre, A., Yan, D., Kustu, S., Nixon, B. T., and Wemmer, D. E. (2003) Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA(+) ATPase domains. *Genes Dev. 17*, 2552−2563.
(38) Christianson, D. W. (1991) Structural biology of zinc. *Adv. Protein Chem. 42*, 281−355.
(39) Flores, S. C., and Gerstein, M. B. (2007) FlexOracle: predicting flexible hinges by identification of stable domains. *BMC Bioinf. 8*, 215.
(40) O'Neill, E., Sze, C. C., and Shingler, V. (1999) Novel effector control through modulation of a preexisting binding site of the aromatic-responsive sigma(54)-dependent regulator DmpR. *J. Biol. Chem. 274*, 32425−32432.
(41) Abril, M. A., Michan, C., Timmis, K. N., and Ramos, J. L. (1989) Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. *J. Bacteriol. 171*, 6782−6790.
(42) Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2005) Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 61*, 449−457.
(43) Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2009) On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 65*, 1089−1097.
(44) Collaborative (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 50*, 760−763.
(45) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126−2132.
(46) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 53*, 240−255.

DOI: 10.1021/acschembio.6b00020
ACS Chem. Biol. 2016, 11, 2357−2365