STRUCTURAL CHARACTERIZATION AND BIOPHYSICAL STUDIES OF BENM, A LYSR-TYPE TRANSCRIPTIONAL REGULATOR IN *ACINETOBACTER BAYLYI* ADP1

by

AJCHAREEYA RUANGPRASERT

(Under the Direction of Cory A. Momany)

ABSTRACT

Bacterial transcriptional regulators have been identified as potential targets for the development of novel antibiotics. LysR-type transcriptional regulators (LTTRs), which comprise the largest group of transcriptional regulators in proteobacteria, regulate diverse metabolic functions including antibiotic resistance and virulence factor synthesis rendering them potential drug targets. Molecular characterization of LTTRs is crucial for the understanding of LTTRs' mechanism of regulation. LTTRs consist of a DNA-binding domain (DBD), a linker helix (LH), and an effector binding domain (EBD) that is composed of two subdomains (EBD-I and EBD-II). BenM, an LTTR found in *Acinetobacter baylyi* ADP1 provides a well-characterized model for structural studies of LTTRs. The structures of full-length BenM and two constitutively active mutants, BenM(E226K) and BenM(R156H), determined by X-ray crystallography illustrate an unexpected "infinite" oligomer of the regulators in their crystal lattices. In addition, all three of the structures are nearly identical suggesting an undefined mechanism that triggers the different *in vivo* activity of these proteins rather than simple conformational

changes as a result of amino acid differences. Analysis of BenM with other LTTRs revealed that there are at least three major oligomerization schemes used by LTTRs. One mode represented by CbnR, DntR and BenM utilizes EBD-II/EBD-II and DBD contacts, a second mode represented by 2ESN, ArgP and TsaR utilizes only DBD contacts for tetramerization, and the third mode utilizes both DBD and EBD-I/EBD-II contacts and is represented by 3FZV. Effector-binding and DNA-binding studies of BenM and its variants confirmed the functionality of the purified proteins used in the crystallization and provided additional insight into the biophysical properties of BenM. The dissociation constants (K_d) of BenM and BenM variants for the effectors, benzoate and *cis,cis*-muconate were derived by tryptophan fluorescence spectroscopy. Affinities of BenM, BenM(E226K) and BenM(R156H) for their target promoters were investigated by electrophoretic mobility shift assay (EMSA).

INDEX WORDS: LysR-type transcriptional regulator, BenM, *Acinetobacter*

STRUCTURAL CHARACTERIZATION AND BIOPHYSICAL STUDIES OF BENM, A LYSR-TYPE TRANSCRIPTIONAL REGULATOR IN *ACINETOBACTER BAYLYI* ADP1

by

AJCHAREEYA RUANGPRASERT

B.S., Chulalongkorn University, Thailand, 2003

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2010

STRUCTURAL CHARACTERIZATION AND BIOPHYSICAL STUDIES OF BENM, A LYSR-TYPE TRANSCRIPTIONAL REGULATOR IN *ACINETOBACTER BAYLYI ADP1*

by

AJCHAREEYA RUANGPRASERT

Major Professor: Cory Momany

Committee: Ellen Neidle
Robert Phillips
Shelley Hooks
Anthony Capomacchia

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2010

DEDICATION

This dissertation is dedicated to with love and gratitude my family. Thank you for your unconditional love, support, and sacrifice throughout my life.

ACKNOWLEDGEMENTS

It is a pleasure to thank those who made this dissertation possible. My deepest gratitude is to my major advisor, Dr. Cory Momany. I am deeply grateful for his supervision and constant encouragement that has helped me in finishing my dissertation. Dr. Ellen Neidle's insightful comments and advices she had offered has allowed me to overcome many crisis throughout different stages of my research. I would also like to acknowledge to Dr. Robert Phillips, Dr. Tony Capomacchia, and Dr. Shelley Hooks who have all been supportive along the way, this dissertation would have impossible without their help and guidance.

Lastly, I offer my regards to all of those who supported me in any respect during the completion of my research.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Gene expression is the process of decoding genetic information stored in the form of DNA into diverse gene products including proteins and other non-protein products such as rRNA and tRNA. For protein-encoding DNA, genetic code is initially decrypted through a process called transcription, from which an intermediary information carrier called mRNA is generated. The information reposited in mRNA transcript is subsequently decoded into protein via a process called translation. In the realm of prokaryotic gene expression, the process of transcription is essentially carried out by the global transcription machinery, DNA-dependent RNA polymerase, which is composed of five subunits, $\alpha_2\beta\beta'\omega$[1; 2]. In most cases, transcription initiation occurs once the multi-subunit core enzyme forms a complex with σ subunit enabling the resulting RNA polymerase holoenzyme to recognize specific promoter sequences[1; 3; 4]. The holoenzyme also gains the ability to position itself at a target promoter correctly and to unwind the double-stranded DNA to start transcription[1]. Several σ factors have been identified in bacteria allowing the holoenzyme to specifically recognize different promoters[4]. At a given promoter, the RNA polymerase holoenzyme initially interacts with two major DNA motifs, the -10 element and the -35 element[5]. A consensus sequence TATAAT of the -10 hexamer is recognized by domain 2 of the σ subunit while the -35 hexamer harboring a consensus sequence TTGACA is recognized by domain 4 of the σ subunit[1; 4; 5; 6]. The extended -10 element and the UP element are two other

crucial sequence elements within promoter regions that are also recognized by RNA polymerase and contribute to initial binding of RNA polymerase to a promoter[1].

As alteration of transcription directly influences patterns of gene expression or, in other words, phenotypes of bacteria, therefore, regulation of gene transcription is considered the key adaptation mechanism of prokaryotes[1]. It is critical that gene transcription be tightly regulated in order for bacteria to undergo different metabolic conditions and growth status or thrive in varying environments. A myriad of factors are involved in prokaryotic transcriptional regulation including regulatory proteins which are proteins that recognize specific DNA sequence and up-regulate (activators) or down-regulate (repressors) transcription of target genes via different mechanisms[1]. Most transcriptional activators increase gene transcription by improving recruitment or affinity of RNA polymerase to a promoter while many repressors primarily generate steric hindrance that prevents RNA polymerase from binding to promoters[1].

LysR-type transcriptional regulators (LTTRs)

To date, a large number of transcriptional regulators in prokaryotes have been identified and categorized into different families[1; 7]. Among over fifteen families of regulatory proteins, the family of LysR-type transcriptional regulators (LTTRs) represents the largest group of transcriptional regulators in prokaryotes[8; 9]. Members of the LTTR family are present in several genera of prokaryotes and regulate transcriptions of genes involved in diverse metabolic functions ranging from amino acid biosynthesis, antibiotic resistance, aromatic compound degradation, response to oxidative stress, response to nitrogen limitation, to synthesis of virulence factors[9; 10]. Interestingly, despite the diversity of target genes controlled by LTTRs, most regulators share at least 20% amino acid sequences identity over the entire protein length

with another family member[9]. The primary structure conservation is especially high in the central region of the 66 N-terminal residues where the sequence is approximately 40% identical in all LTTRs while residues in the C-terminal portions lack significant sequence conservation[9]. Beyond similarity at the sequence level, LTTR family members also display commonality of other characteristics. The N-terminal region of LTTRs assumes a winged helix-turn-helix (wHTH) DNA binding motif which is structurally conserved throughout the family. On average, LTTR monomer consists of 300-350 amino acids and forms biologically active unit by assembling into multimer, most likely tetramer or dimer[9]. Most LTTRs function by sensing the presence of small molecule effectors and convey the ligand-binding signals to transcriptional activation or repression[9]. However, in general, the presence of effectors is dispensable for binding of regulators to their regulated DNA promoters, and binding affinity of proteins to DNA may or may not be altered upon binding of inducers[9; 11; 12; 13]. LTTR genes are commonly found divergently transcribed from a promoter that overlaps a promoter of a regulated gene. This overlapping promoter organization allows many LTTRs to both control the transcription of target genes and negatively autoregulate their own transcriptions simultaneously. Nevertheless, promoter structure of several LTTRs lacks divergent nature such as those of CysB, MleR, , Nac, NahR, OxyR, PhcA, and SpvR[9]. Despite the absent of divergent promoter organization, CysB, Nac, and OxyR still exhibit negative autoregulation[9].

DNA-binding characteristics of LTTRs

DNA binding properties of LTTRs have been widely investigated giving insight into how the family members interact with promoter DNA. Several footprinting experiments revealing large regions of DNA protected by LTTRs regardless the presence of effectors verified that

inducer binding was not a prerequisite for protein-DNA interactions[12; 13; 14; 15]. Nevertheless, the presence of cognate ligands usually affects patterns of DNA protection[13; 16; 17; 18]. The intergenic region between an LTTR gene and its target gene, where the regulators generally bind, harbors multiple sites with modular DNA sequence that the proteins recognize[10; 13]. Several of these intergenic regions have been characterized and two modular binding sites, repression binding site (RBS) and activation binding site (ABS), have been identified[17; 18; 19; 20; 21]. The RBS is commonly overlapping the promoter for the gene encoding itself, and binding of LTTRs to the RBS thus allows autorepression[13]. The RBS for LTTRs displays an imperfect palindromic repeat ATAC-N_7-GTAT containing a consensus T-N_{11}-A element which is generally known as the "LTTR motif" or "LTTR box" which was suggested to interact with one regulator dimer [10; 22; 23]. This modular inverted repeat was identified in numerous promoters regulated by LTTRs, but the base pair composition and length can be varied[10]. The RBS is considered the primary binding site of LTTRs[24]. Unlike the RBS, the ABS lacks a conserved sequence motif, but, in certain cases, the ABS shows some similarity to the RBS sequence[10; 24]. Binding of regulators to the ABS which is located between the RBS and the regulated promoter is suggested to be necessary for mediating contacts with RNA polymerase and transcription activation[13; 15; 22]. Affinity of LTTRs for each of these binding sites is proposed to be subject to inducer binding[9; 12; 17; 24]. In the absence of inducer, CatR, a LysR family member from *Pseudomonas putida*, for instance, only binds to the RBS, while the protein interaction with the ABS occurs once inducer is bound[10; 15; 17; 20; 22; 25].

In general, the presence of cognate effectors has very little effect on DNA-binding affinity of LTTRs for their target DNA[26]. For instance, in the case of AtzR, an LTTR of *Pseudomonas* sp. strain ADP, its cognate inducer, cyanuric acid, did not appear to alter the

overall binding affinity of the regulator to the *atzR-atzDEF* promoter region[13]. However, in certain cases, the overall affinity of the regulators for the target promoter DNA has been suggested to increase once the inducers are bound. Binding constants of some LTTRs to their target promoters in the absence and presence of inducers have been investigated via gel retardation assays. The equilibrium dissociation constants (K_d) of CatR from *Pseudomonas putida* for the *catBC* promoter was decreased from 7.0 x 10^{-11} M in the absence of its inducer, *cis,cis*-muconate to 3.1 x 10^{-11} M when 100 μM muconate was present[17; 25]. Similar results were observed when NADPH was added to the binding assay of CbbR, an LTTR identified in *Xanthobacter flavus*, and *cbbR-cbbL* intergenic region harboring its cognate binding sites. NADPH increased affinity of CbbR for the DNA and also relaxed CbbR-induced DNA bending[12; 27].

Binding of many LTTRs to promoters reportedly cause DNA bending[9; 11; 12; 28; 29; 30]. Hypersensitive cleavage sites were commonly identified in promoters regions between two LTTR binding motifs upon binding of LysR family members[12; 13; 14; 22; 31]. At certain promoters, the presence of co-inducers appears to relieve DNA bends suggesting that binding of co-inducers to LTTRs result in alterations in protein conformations that affect nature of protein-DNA interaction and, consequently, create effective contact surface with RNA polymerase leading to transcriptional activation[11; 12; 14; 21; 26; 28; 30; 32; 33]. These alterations in DNA bending and relaxation correspond to a shift in DNA binding sites of LTTRs upon ligand binding.

Binding of inducers to LTTRs has been proposed to provoke protein conformational changes which account for altered DNA binding patterns and transcriptional regulation[9; 10]. Conformational differences between free and effector-bound LTTRs have been first demonstrated in the crystal structures of BenM EBDs[34]. The ability of different LTTRs to

recognize various inducers is conferred from low sequence conservation in the C-terminal EBD. Mostly, the inducers are intermediates of metabolic pathways they regulate[22]. Inducer binding properties of some LTTRs have been characterized. DntR and Cbl both display low micromolar dissociation constants for their cognate effectors, sodium salicylate and adenosine phosphosulfate, respectively while BenM was found to bind to benzoate and *cis,cis*-muconate with millimolar and high micromolar affinities, respectively[35; 36; 37].

LTTR-RNA polymerase interactions

Transcriptional activators are generally categorized into three major classed based on mechanisms of activation, Class I, II and III[1]. Class I activators up-regulate transcription by binding to a region upstream of the -35 element and recruits RNA polymerase to the promoter by interacting with C-terminal domain of α-subunit (αCTD) of the holoenzyme[1; 38]. The αCTD of RNA polymerase was demonstrated as a crucial factor for transcriptional activation by some LTTRs including CatR, TrpI, OxyR, and CysB, and direct contact between αCTD and the regulators was identified suggesting that LTTRs basically belong to Class I transcriptional regulators[39; 40; 41; 42; 43; 44]. Mutations of certain regions of αCTD especially the contact site I region that could disrupt contact between the regulators and RNA polymerase were demonstrated to affect activation of transcription by LTTRs. Residues in the turn of the wHTH motif of LTTRs, as well as some residues lying in the C-terminal part of the proteins, mapped by mutation analysis were proposed to comprise surface involved in interactions with αCTD, also known as activating region[44; 45].

Structural Studies of LTTRs

To date, an increasing number of LTTRs have been discovered and characterized. Nonetheless, a detailed mechanism of transcriptional regulation by LTTRs remains a mystery partly due to lack of crystal structures of full-length proteins with promoter DNA. In order to achieve better understanding of how proteins in this family regulate gene transcriptions, extensive efforts to characterize the regulators at a molecular level have been made. Despite the growing number of family members reported, very few structures of full-length LTTRs have been determined successfully[46]. Further, there is only one full-length LTTR whose structure is determined in the presence of its cognate effector, and none of the full-length regulator has been crystallized with its target promoter or RNA polymerase. This limited availability of LTTR crystal structures is, to a certain extent, due to the notoriously low solubility of LTTRs, which impedes the process of achieving crystals[31; 47]. Further, requirement of high salt concentrations in buffer also hinder structural determination of LTTRs in complex with their natural ligands[48].

CbnR, the first family member whose structure was characterized as a full-length regulator, regulates the expression of genes responsible for chlorocatechol degradation in *Ralstonia eutropha* NH9[48; 49]. The crystal structure of full-length CbnR revealed classic LTTR domain organization that is composed of N-terminal DNA-binding domain (DBD) and C-terminal effector-binding domain (EBD) connected by a long linker helix (LH) (Figure 1.1)[48]. The presence of a wHTH motif was confirmed in the tri-helical structures of CbnR[48].

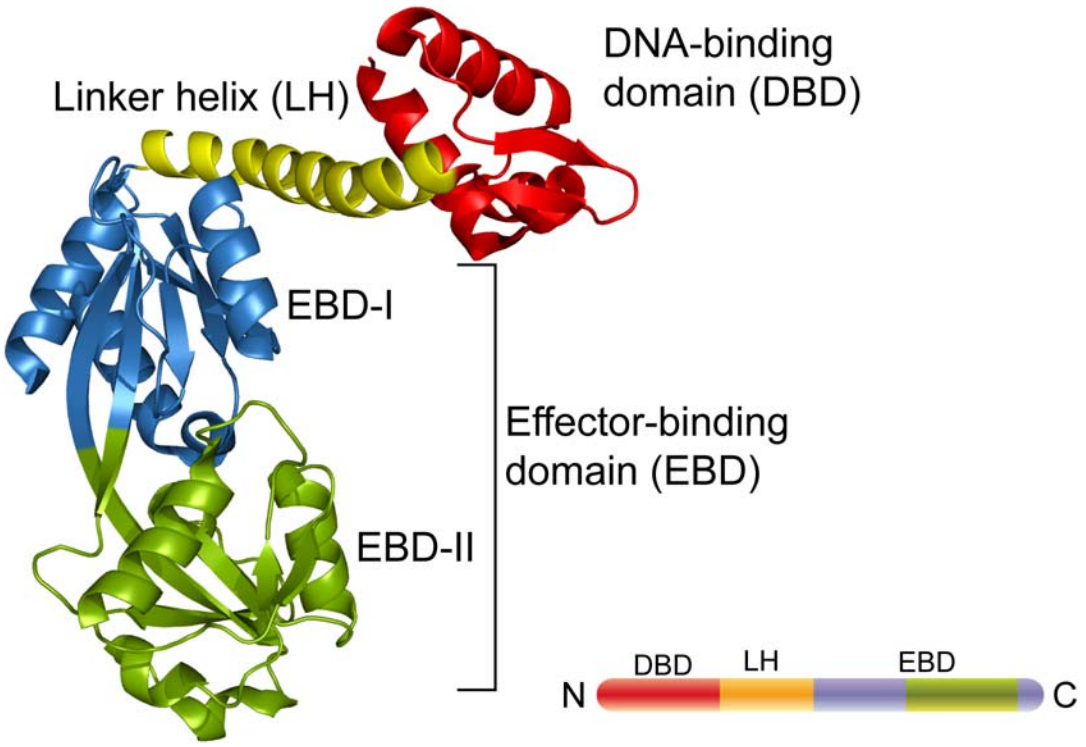

Figure 1.1. Domain organization of LTTRs based on the crystal structure of CbnR (PDB ID 1IZ1). DBD and LH are shown in red and yellow, respectively. Two subdomains of EBD, EBD-I and EBD-II, are shown in blue and green, respectively.

The wHTH is one permutation of the best characterized class of DNA binding motifs, the HTH class[50]. The fundamental fold of the HTH is composed of three helices forming a right-handed bundle. The second and the third helices of the tri-helical bundle, together with a sharp turn in between, constitute a characteristic feature that corresponds to the name of the motif, the HTH[48; 50]. The three helices of the DBD are arranged in the fashion that they create conserved hydrophobic core that helps stabilize the fold and present polar residues on the surface that typically interact with the DNA[48; 50]. In general, the third helix of the domain, widely referred to as the recognition helix, possesses shape complementary to the DNA and provides base-specific

interactions with the DNA by inserting itself into the DNA major groove[50; 51]. The β-strand hairpin located C-terminal to the third helix defines the wing part of the wHTH domain. This wing potentially provides an interacting surface with the phosphate backbone or the minor groove of DNA and helps position regulators to target promoters[50]. However, residues involved in LTTR-DNA interactions are not confined solely within the wHTH motif but are distributed across the entire DBD region including the first helix of the tri-helical bundle[52; 53].

The long LH serves as a connection between the N-terminal DBD and the inducer-binding domain. The crystal structures of C-terminal EBDs of all the LTTRs structurally characterized to date possess the conserved Rossmann-like folds which resembles the fold of periplasmic binding proteins[34; 46; 48; 54; 55; 56; 57; 58]. Low sequence similarity of C-terminal parts of LTTRs makes this structural conservation highly striking[9]. A typical EBD of LTTRs is composed of two subdomains, EBD-I and EBD-II, which are connected by two cross over β-strands[48; 54; 55; 56]. Located between the two EBD subdomains is a predicted effector binding pocket where binding of natural effectors was observed in the crystal structures of BenM, CatM, and TsaR (Figure 1.2)[34; 46]. Additionally, an unexpected secondary effector binding cavity was discovered in the crystal structures of BenM-EBD[34]. This secondary site is located in the hydrophobic core of EBD-I and is able to accommodate interactions with the effector[34]. Other than playing a major role in inducer recognition, EBDs were proposed to provide a critical surface for oligomerization and interactions with αCTD of RNA polymerase[52; 59].

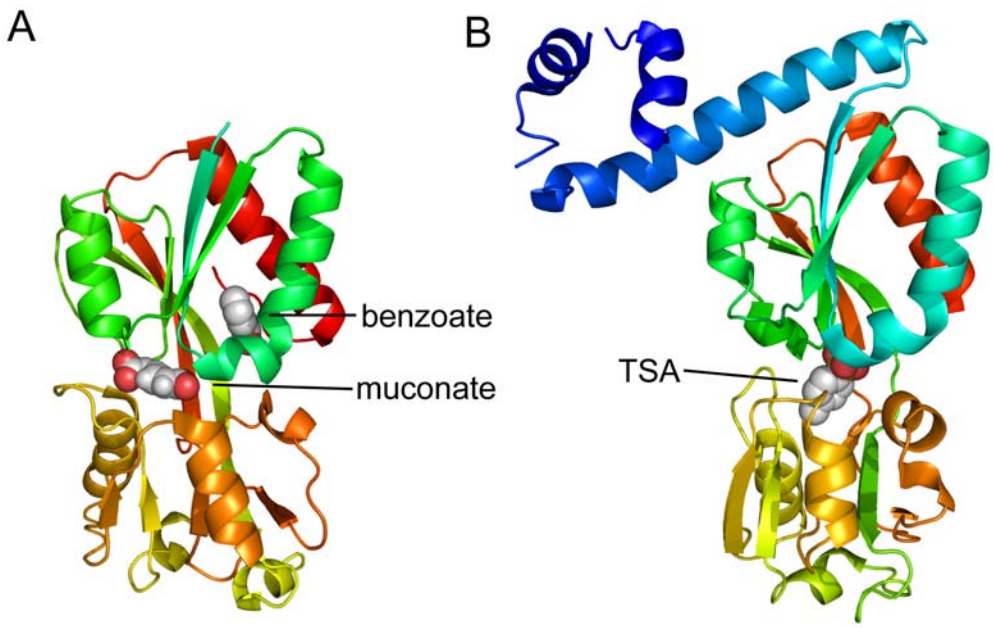

Figure 1.2. Ribbon representations of BenM-EBD and TsaR with their natural effectors.(A) BenM-EBD houses its natural effectors, *cis*,*cis*-muconate (muconate) and benzoate, in the primary binding site and secondary binding site, respectively, and (B) full-length TsaR binds to its cognate effector, *p*-toluenesulfonate (TSA).

All of the crystal structures of full-length LTTRs available to date reveal similar domain organizations[31; 46; 48; 57; 58]. Nevertheless, despite this structural resemblance, the oligomerization modes of the LTTRs are highly diverse. CbnR, the first family member to be crystallized as a full-length protein, was found to form a tetrameric unit in the crystals that was suggested to represent the biologically active form of the regulator[48]. The CbnR tetramer is composed of a dimer of dimers. Each CbnR dimer is formed by two monomers with two distinct conformations, compact and extended (Figure 1.3)[48]. The ability of two CbnR monomers with identical primary structure to adopt different conformations is highly intriguing. This alteration of conformation is

possibly due to high of flexibility at the hinge region connecting the LH to the EBD. This flexibility results in different projections of LHs and DBDs and, consequently, confers the two discrete forms of the monomer. Likewise, two alternative conformations of LTTR monomer are also observed in the crystal structures of TsaR, ArgP, and two other uncharacterized LTTRs, 2ESN and 3FZV, suggesting that this feature is probably common among LTTRs[46; 57].

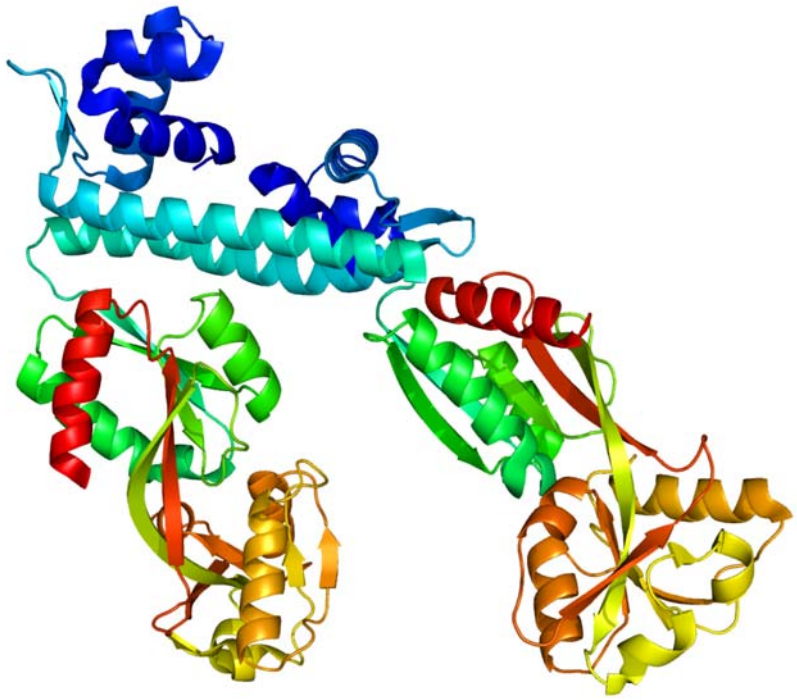

Figure 1.3. Compact and extended conformations of CbnR (PDB ID 1IZ1).

As observed in CbnR, TsaR, ArgP, 2ESN, and 3FZV dimers, the EBD of the two subunits are related by a local 2-fold axis and interact with each other through a large interface burying an area over 1000 Å (Figure 1.4)[248]. This dimerization interface is also commonly present in truncated structures of C-terminal EBDs of all other LTTRs.

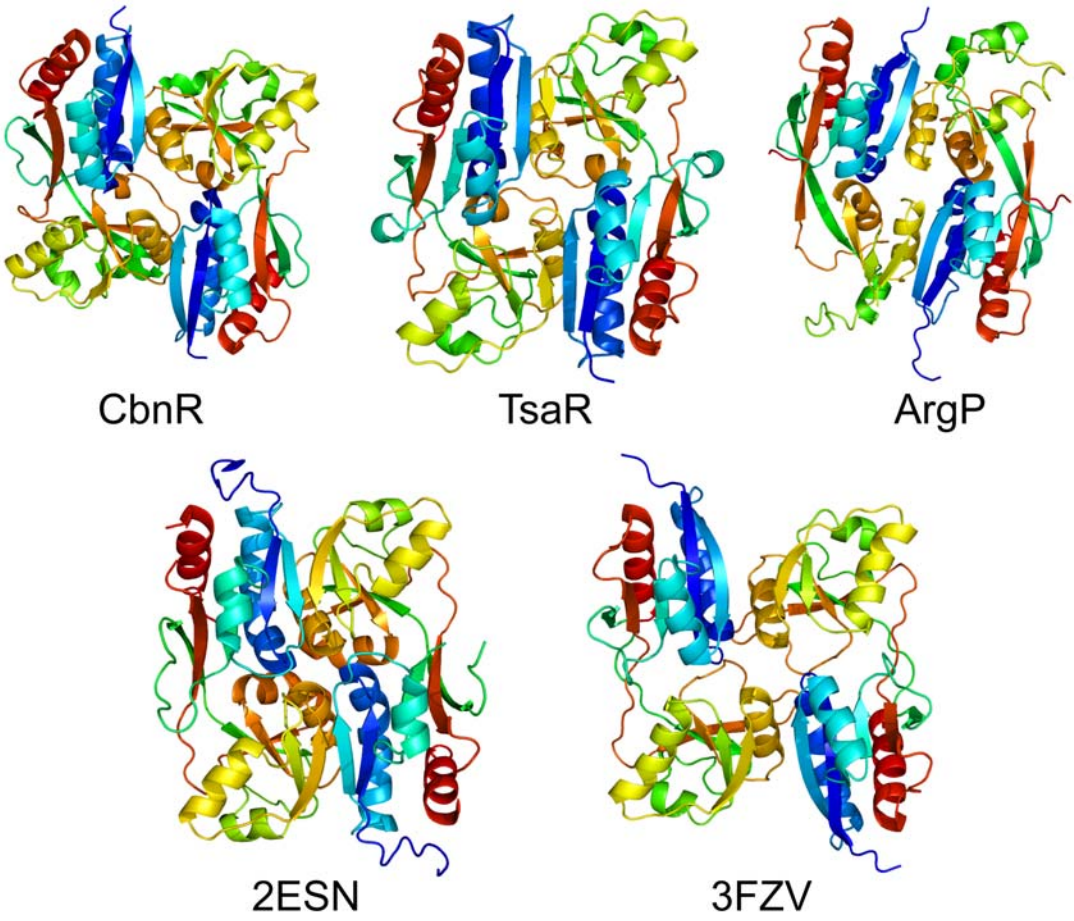

Figure 1.4. Ribbon representations of dimers of CbnR, TsaR, ArgP, 2ESN, and 3FZV. Only EBDs are shown.

Despite the similarity in the overall features of these LTTR dimers, modes of assembly into larger oligomers of the regulators are vastly diverse (Figure 1.5). Two dimers of CbnR form a tetramer via interactions in both LH and EBD regions. LHs from one dimer interact with those from the other dimer in an anti-parallel fashion. A tetramerization interface located between α10 helices of two EBDs from different dimers is also utilized to help holding the two dimers together. Similar interfaces comparable to the tetramerization interface of CbnR were also

identified in the structures of DntR EBD, CysB EBD, and BenM EBD suggesting that these LTTRs tetramerize in the same manner. However, in the crystal structures of TsaR and ArgP, although the interactions in LH parts are similar to those found in the CbnR structures, no contact is observed between the EBDs of the two dimers. The structure of one of the two uncharacterized LTTRs, 2ESN, also shares this feature, while the other, 3FZV, adopts another distinct mode of oligomerization. Unlike CbnR or TsaR, 3FZV subunits form a tetramer by clustering EBDs of all the four monomers together rendering a large contact interface on each EBD.

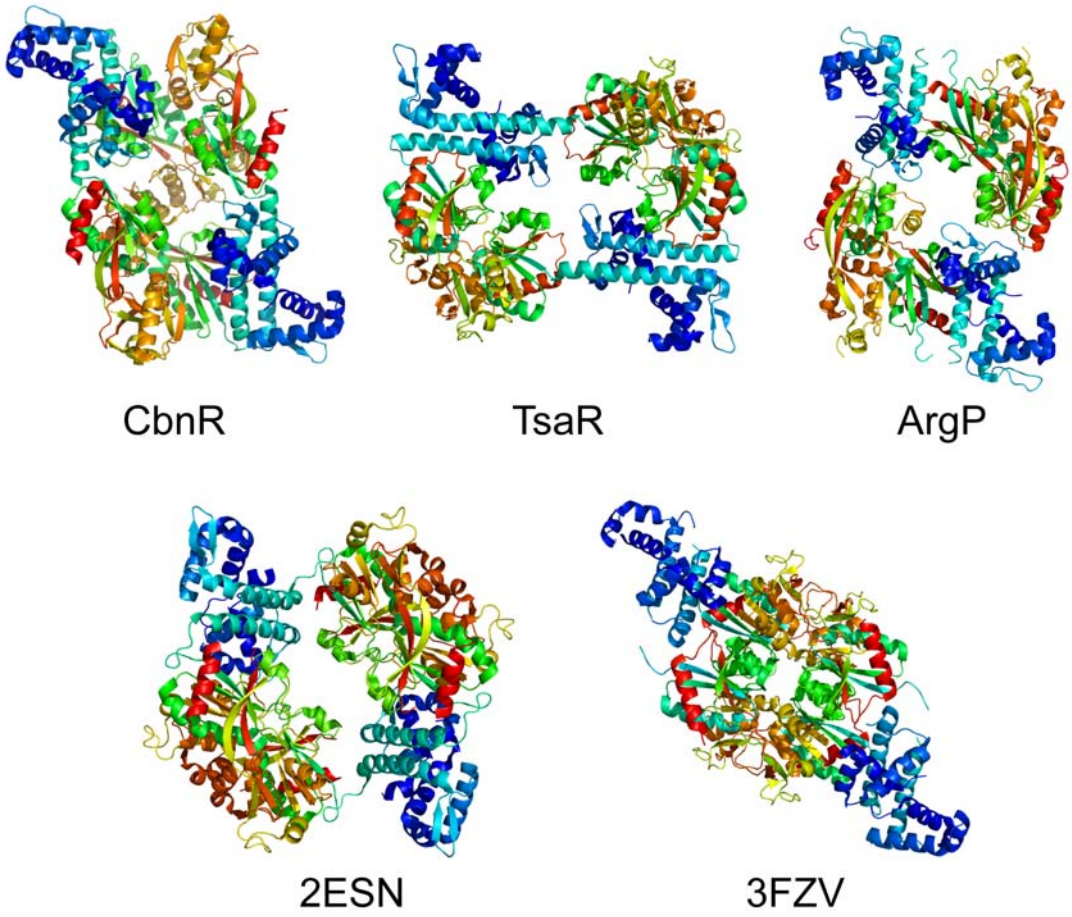

Figure 1.5. Tetramers of CbnR, TsaR, ArgP, 2ESN, and 3FZV.

The ability of LTTRs to adopt two distinct conformations is seen in all the structures of full-length regulators except that of CrgA. Instead, CrgA assembles into octamers using only one conformation in the crystals that is comparable to the compact forms indentified in other LTTRs (Figure 1.6A)[58]. Eight CrgA subunits arrange into a square-shaped ring structure where each CrgA subunit makes contacts with two adjacent subunits, one via EBD interfaces and the other using anti-parallel LH interfaces (Figure 1.6B)[58]. The EBD interfaces used in CrgA octamerization is essentially similar to those involved in dimer formation of all other LTTRs.

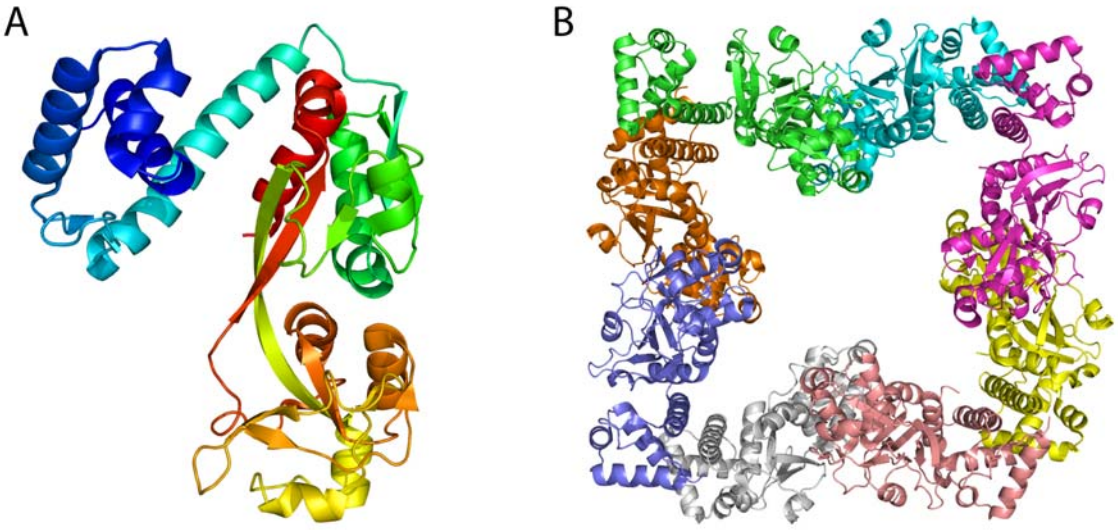

Figure 1.6. (A) Ribbon representation of CrgA monomer. (B) The crystal structure of octameric ring-like CrgA.

BenM as a model to understand the mechanism of transcriptional regulation by LTTRs

BenM is an LTTR that controls expressions of multiple genes involved in the degradation of benzoate via the β-ketoadipate pathway in the soil bacterium *Acinetobacter baylyi* ADP1

(Figure 1.7A)[60]. The pathway of benzoate degradation in this bacterium is under the regulation of two LTTR paralogs with overlapping functions, BenM and CatM, whose primary structures are 59% identical (54% in the EBD) and 75% similar[60]. The sequence identity is especially high in the N-terminal 58 amino acids of BenM and CatM (98% similarity)[34]. Additionally, both regulators respond to the same metabolite formed in the benzoate degradation pathway, *cis*, *cis*-muconate (hereafter referred to as muconate)[60; 61]. However, unlike CatM, BenM is also able to recognize benzoate as its sole effector[60; 61].

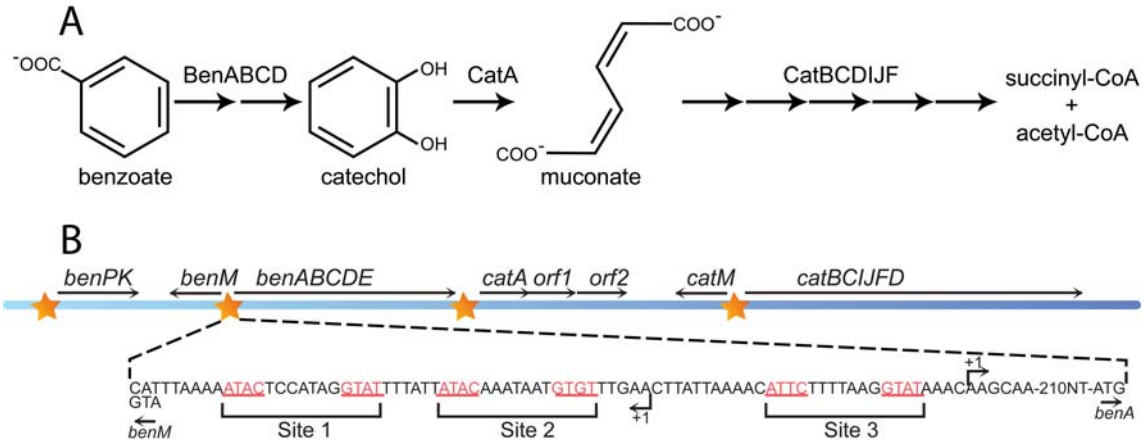

Figure 1.7. (A) Catechol branch of β-ketoadipate pathway. (B) Organization of *ben* and *cat* genes. Three potential binding sites for BenM and CatM in the intergenic region between *benM* and *benA* genes are indicated as Site 1, Site 2 and Site 3.

Several enzymes required for conversion of benzoate to tricarboxylic acid intermediates are encoded by chromosomal *ben* and *cat* genes[62]. The *ben* genes are necessary for converting benzoate to catechol, and the *cat* genes are indispensible for further catechol catabolism[63]. Although BenM and CatM both are capable of activating the transcriptions from *benA*, *catA*, and

catB promoters, distinct effects on transcriptional activations are exerted at different loci. BenM plays a major role in activation of the transcription from the *benA* promoter whereas CatM more efficiently activates the transcription from the *catB* promoter[60; 61; 64]. At the *catA* promoter, however, both BenM and CatM exhibit essentially comparable degrees of transcriptional activation but only in response to muconate [60; 61]. Both *benM* and *catM* demonstrate a divergent overlapping promoter arrangement with their regulated targets, *benA* and *catB* respectively, which is a common feature shared by many LTTRs (Figure 1.7B)[9]. This promoter structure is suggested to allow BenM and CatM to repress their own transcriptions.

Detailed investigations on BenM and *benA* interactions revealed that the *benA*-*benM* intergenic region harbors three sites, Site 1, Site2 and Site 3, with a LTTR consensus binding sequence for BenM (Figure 1.7B). Each binding site displays dyad symmetry, and one BenM-DBD is suggested to bind one half-site[14]. Site 1 has a perfect consensus sequence, ATAC-N_7-GTAT, while the sequences of Site 2, ATAC-N_7-GT<u>G</u>T, and Site 3, AT<u>T</u>C-N_7-GTAT, both differ from the consensus binding sequence at one position (underlined)[14]. Binding patterns of BenM to these three binding sites in the intergenic region were found to be effector-dependent. In the absence of effectors, BenM, putatively as a tetramer, binds simultaneously to Site 1 and Site 3, whose centers are 52 bp apart resulting in formation of a predicted DNA loop[14]. Binding of the regulator to Site 3 prevents the access of the transcription machinery to the promoter and represses basal *benA* gene expression[14]. In the presence of benzoate and/or muconate, BenM interacts with Site 1 and Site 2, whose centers are separated by 21 bp, causing the DNA bending to relax[14]. This effector-induced pattern of DNA interactions renders accessible the *benA* promoter and most likely allows contacts between the regulator and RNA polymerase resulting in transcriptional activation.

Effects of benzoate and muconate on transcription activation of *benA* by BenM have been extensively studied both *in vivo* and *in vitro*. Either benzoate or muconate individually can act as a coactivator allowing activation of transcription from the *benA* promoter by BenM, with muconate being more effective as the sole inducer[14]. Intriguingly, synergistic transcriptional activation of *benA* by BenM was demonstrated when both benzoate and muconate were present[14].

In order to fully understand the mechanism of regulation by BenM, structural studies of BenM using X-ray crystallography have been conducted. However, due to the notoriously low solubility character of LTTRs, only truncated BenM was successfully crystallized until recently[34; 65]. BenM-EBD has the characteristic periplasmic-binding protein fold consisting of two subdomains connected by a hinge formed by two antiparallel β-strands as observed in previously available structures of other LTTR EBDs[34]. EBD-I of BenM encompasses residues 87 to 161 and residues 268 to C-terminus, and residues 162 to 267 constitute EBD-II[34]. Crystals of BenM-EBD were also soaked with its natural inducers yielding broken but diffractable crystals of ligand-bound BenM[34]. Muconate-bound BenM-EBD structures and Benzoate-bound BenM-EBD structures confirmed that the primary effector binding site lies at the cavity between EBD-I and EBD-II. An unexpected second effector binding site that accommodates benzoate molecule was also identified (Figure 1.2B)[34]. The discovery of this secondary binding site that allows benzoate to bind to BenM simultaneously with muconate is proposed to produce the reported synergistic effects of benzoate and muconate on transcriptional activation of *benA* by BenM[34]. Furthermore, the fact that benzoate was not introduced during the purification or crystallization of the muconate-bound BenM crystals implied that benzoate might have been present in very

small amounts as a contaminant, and, thus, it might have had relatively high affinity for the secondary binding site.

Specific interactions between muconate and protein residues in the primary binding site are abundant while benzoate only utilizes a subset of these interactions and occupies the space in the cavity loosely (Figure 1.8)[34]. Domain movement upon inducer binding which potentially plays a critical part in its transcriptional activation mechanism is evident in the crystal structures. Binding of muconate triggers EBD-I and EBD-II to move closer together and clamp down on the effector molecule itself, while binding of benzoate creates conformational changes to a lesser extent[34]. Movement of helices at the ligand binding site were suggested to generate more pronounced global conformational changes once four subunits assemble into a tetramer, and these changes potentially confer effector-mediated shifts in DNA binding[34].

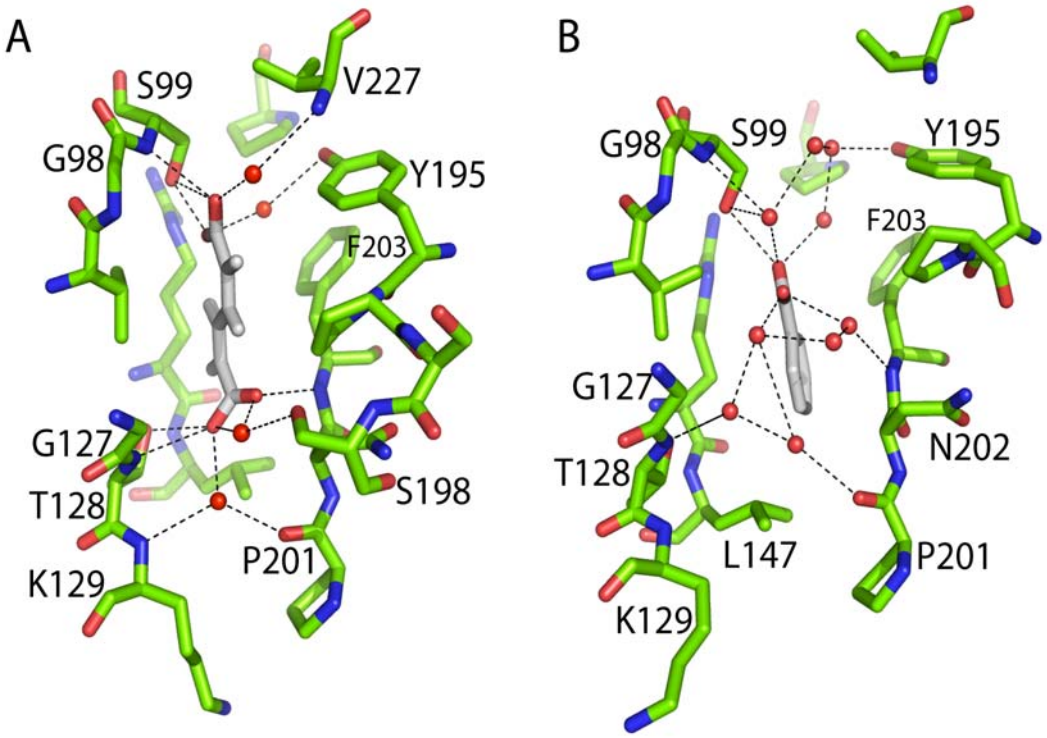

Figure 1.8. Interactions in the primary binding site of BenM-EBD with (A) muconate and (B) benzoate. Muconate binding involves higher number of specific interactions with BenM residues than binding of benzoate while more water molecules are present in the benzoate-bound cavity.

The secondary binding site of BenM is located in the core of EBD-I, adjacent to the primary binding site (Figure 1.2A). Although residues in this hydrophobic pocket are highly conserved in BenM and CatM, the ability of the secondary site to bind benzoate is only exclusive to BenM, possibly due to two key residues, Arg160 and Tyr293, that are only present in the secondary site of BenM but not in that of CatM[34; 66]. These two residues provide salt bridges and hydrogen bond with carboxyl group of benzoate molecule, and other hydrophobic side chains in the pocket form hydrophobic interactions with the benzoate ring (Figure 1.9). Certain mutations of Arg160 and Tyr293 not only abolish the ability of benzoate to bind at the secondary site, but are also indispensable in BenM's response to benzoate[66]. Studies indicated that both Arg160 and Tyr293 are essential for benzoate-activated transcription and synergistic effect in BenM[66]. BenM variants that carry individual or double replacements of Arg160 and Tyr293 lost the ability to grow on benzoate and benzoate-activated *benA* expression[66]. However, two BenM variants that retained double replacements resumed the ability to grow on benzoate by spontaneously obtaining a third mutation at Arg225 to His or Glu226 to Lys[66]. These extra mutations allow the double mutants to achieve high levels of *benA* expression even in the absence of muconate or benzoate, and these levels of expression are higher than muconate-induced *benA* expression by wild-type BenM[66]. Furthermore, BenM variants with single mutations, R225H or E226K, without any other replacements or added inducers, were enough to enable high level *benA* expression[66].

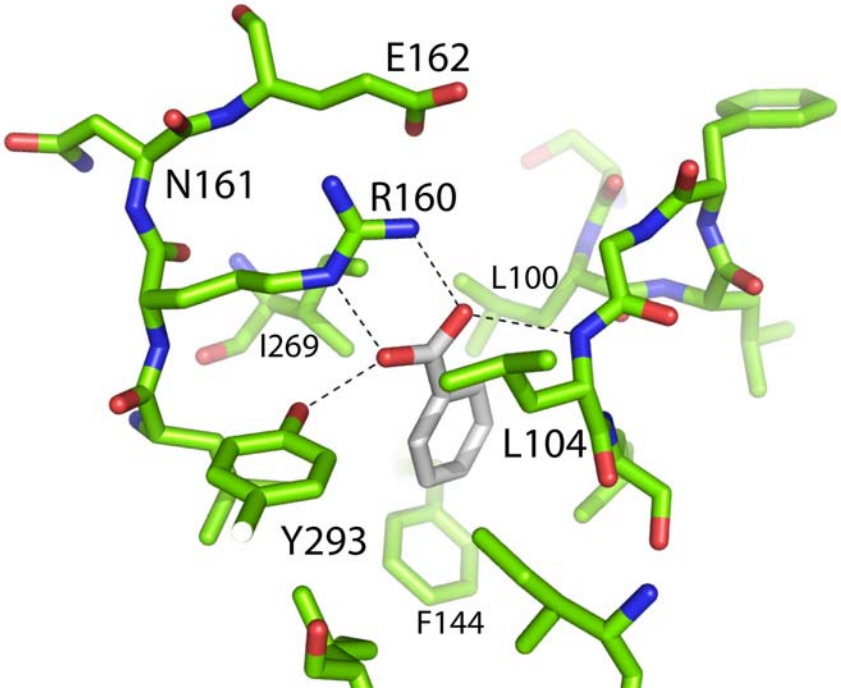

Figure 1.9. Binding of benzoate at the secondary site of BenM. The effector forms salt-bridges with Arg160 and hydrogen bonds with Tyr293. Hydrophobic side chains of residues in the vicinity also help accommodate benzoate ring.

Binding affinities of benzoate and muconate to BenM were determined by fluorescence emission spectroscopy revealing that BenM has higher affinity for muconate than for benzoate[35]. Full-length BenM binds benzoate with a dissociation constant of 1.2 mM while it binds muconate with dissociation constant of 0.28 mM[35]. These values are consistent with findings from effector-bound BenM structures where several specific interactions occur in muconate binding while only a subset of these interactions were utilized in binding of benzoate[34]. Moreover, in the presence of 1 mM benzoate, the binding affinity of the regulator for muconate

is decreased suggesting that at this concentration of benzoates, the two effectors compete for the same primary binding site[35]. Nevertheless, the binding affinity of BenM for benzoate at the secondary pocket has never been studied by biophysical techniques.

Comprehensive knowledge of gene organization, transcriptional activation and inducer response of BenM and its variants, together with the availability of a wide variety of different BenM-EBD crystal structures make BenM an excellent model for investigating mechanism of transcriptional regulation by LTTRs. In an attempt to achieve a complete understanding of how BenM functions, detailed investigation of full-length BenM structures is required. In Chapter 2, the crystal structures of full-length BenM and BenM mutants are elaborated. Comparison of BenM with other full-length LTTR structures and analysis of DBD orientations are described in detail. The chapter also summarizes schemes of oligomerization observed in all LTTR structures available to date. Chapter 3 focuses on effector binding properties of BenM. Binding affinities for benzoate and muconate at distinct binding sites of BenM and BenM mutants determined by fluorescence spectroscopy are presented and compared with previously determined values. Certain complications in determining dissociation constants are discussed. In addition, DNA-binding studies of BenM and BenM variants by electrophoretic mobility shift assay (EMSA) are included in Appendix A. The final chapter provides conclusions of the work present in this dissertation as well as future research directions.

References

1. Browning, D. F. & Busby, S. J. W. (2004). The regulation of bacterial transcription initiation. *Nature Reviews Microbiology* **2**, 57-65.

2. Darst, A. (2001). Bacterial RNA polymerase. *Current Opinion in Structural Biology* **11**, 155-162.

3. Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L. & Darst, S. A. (2002). Structure of the bacterial RNA polymerase promoter specificity [sigma] subunit. *Molecular cell* **9**, 527-539.

4. Helmann, J. D. & Chamberlin, M. J. (1988). Structure and function of bacterial sigma factors. *Annual review of biochemistry* **57**, 839-872.

5. Busby, S. & Ebright, R. H. (1994). Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell* **79**, 743-746.

6. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285.

7. Pareja, E., Pareja-Tobes, P., Manrique, M., Pareja-Tobes, E., Bonal, J. & Tobes, R. (2006). ExtraTrain: a database of Extragenic regions and Transcriptional information in prokaryotic organisms. *BMC microbiology* **6**, 29.

8. Henikoff, S., Haughn, G. W., Calvo, J. M. & Wallace, J. C. (1988). A large family of bacterial activator proteins. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 6602.

9. Schell, M. A. (1993). Molecular biology of the LysR family of transcriptional regulators. *Annual Reviews in Microbiology* **47**, 597-626.

10. Maddocks, S. E. & Oyston, P. C. F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology-Sgm* **154**, 3609-3623.

11. Hryniewicz, M. M. & Kredich, N. M. (1994). Stoichiometry of binding of CysB to the cysJIH, cysK, and cysP promoter regions of Salmonella typhimurium. *Journal of bacteriology* **176**, 3673.

12. Van Keulen, G., Girbal, L., van den Bergh, E. R., Dijkhuizen, L. & Meijer, W. G. (1998). The LysR-type transcriptional regulator CbbR controlling autotrophic CO2 fixation by Xanthobacter flavus is an NADPH sensor. *Journal of bacteriology* **180**, 1411.

13. Porrúa, O., García-Jaramillo, M., Santero, E. & Govantes, F. (2007). The LysR-type regulator AtzR binding site: DNA sequences involved in activation, repression and cyanuric acid-dependent repositioning. *Molecular microbiology* **66**, 410-427.

14. Bundy, B. M., Collier, L. S., Hoover, T. R. & Neidle, E. L. (2002). Synergistic transcriptional activation by one regulatory protein in response to two metabolites. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 7693-7698.

15. Chugani, S. A., Parsek, M. R. & Chakrabarty, A. M. (1998). Transcriptional repression mediated by LysR-type regulator CatR bound at multiple binding sites. *Journal of bacteriology* **180**, 2367.

16. Chang, M. & Crawfordt, I. P. (1990). The roles of indoleglycerol phosphate and the Trpl protein in the expression of trpBA from Pseudomonas aeruginosa. *Nucleic acids research* **18**, 979.

17. Parsek, M. R., Shinabarger, D. L., Rothmel, R. K. & Chakrabarty, A. M. (1992). Roles of CatR and cis, cis-muconate in activation of the catBC operon, which is involved in benzoate degradation in Pseudomonas putida. *Journal of bacteriology* **174**, 7798.

18. Wek, R. C. & Hatfield, G. (1988). Transcriptional activation at adjacent operators in the divergent-overlapping ilvY and ilvC promoters of Escherichia coli* 1. *Journal of molecular biology* **203**, 643-663.

19. Coco, W. M., Parsek, M. R. & Chakrabarty, A. M. (1994). Purification of the LysR family regulator, ClcR, and its interaction with the Pseudomonas putida clcABD chlorocatechol operon promoter. *Journal of bacteriology* **176**, 5530.

20. Rothmel, R. K., Shinabarger, D. L., Parsek, M. R., Aldrich, T. L. & Chakrabarty, A. M. (1991). Functional analysis of the Pseudomonas putida regulatory protein CatR: transcriptional studies and determination of the CatR DNA-binding site by hydroxyl-radical footprinting. *Journal of bacteriology* **173**, 4717.

21. Díaz, E. & Prieto, M. A. (2000). Bacterial promoters triggering biodegradation of aromatic pollutants. *Current Opinion in Biotechnology* **11**, 467-475.

22. Tropel, D. & Van Der Meer, J. R. (2004). Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiology and Molecular Biology Reviews* **68**, 474.

23. Goethals, K., Van Montagu, M. & Holsters, M. (1992). Conserved motifs in a divergent nod box of Azorhizobium caulinodans ORS571 reveal a common structure in promoters regulated by LysR-type proteins. *Proceedings of the National Academy of Sciences* **89**, 1646.

24. Porrúa, O., Platero, A. I., Santero, E., del Solar, G. & Govantes, F. (2010). Complex interplay between the LysR-type regulator AtzR and its binding site mediates atzDEF activation in response to two distinct signalsmmi_7100 331.. 347. *Molecular Microbiology* **76**, 331-347.

25. Parsek, M. R., Ye, R. W., Pun, P. & Chakrabarty, A. M. (1994). Critical nucleotides in the interaction of a LysR-type regulator with its target promoter region. catBC promoter activation by CatR. *Journal of Biological Chemistry* **269**, 11279.

26. Perez-Martin, J., Rojo, F. & De Lorenzo, V. (1994). Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiology and Molecular Biology Reviews* **58**, 268.

27. Van Den Bergh, E. R., Dijkhuizen, L. & Meijer, W. G. (1993). CbbR, a LysR-type transcriptional activator, is required for expression of the autotrophic CO2 fixation enzymes of Xanthobacter flavus. *Journal of bacteriology* **175**, 6097.

28. Parsek, M. R., Kivisaar, M. & Chakrabarty, A. M. (1995). Differential DNA bending introduced by the Pseudomonas putida LysR-type regulator, CatR, at the plasmid-borne pheBA and chromosomal catBC promoters. *Molecular Microbiology* **15**, 819-828.

29. Fisher, R. F. & Long, S. R. (1993). Interactions of NodD at the nod box: NodD binds to two distinct sites on the same face of the helix and induces a bend in the DNA. *Journal of molecular biology* **233**, 336-348.

30. Hryniewicz, M. M. & Kredich, N. M. (1991). The cysP promoter of Salmonella typhimurium: characterization of two binding sites for CysB protein, studies of in vivo transcription initiation, and demonstration of the anti-inducer effects of thiosulfate. *Journal of bacteriology* **173**, 5876.

31. Sainsbury, S., Ren, J., Saunders, N. J., Stuart, D. I. & Owens, R. J. (2008). Crystallization and preliminary X-ray analysis of CrgA, a LysR-type transcriptional regulator from pathogenic Neisseria meningitidis MC58. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **64**, 797-801.

32. Storz, G., Tartaglia, L. A. & Ames, B. N. (1990). Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. *Science* **248**, 189.

33. Huang, J. Z. & Schell, M. A. (1991). In vivo interactions of the NahR transcriptional activator with its target sequences. Inducer-mediated changes resulting in transcription activation. *Journal of Biological Chemistry* **266**, 10830.

34. Ezezika, O. C., Haddad, S., Clark, T. J., Neidle, E. L. & Momany, C. (2007). Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. *Journal of molecular biology* **367**, 616-629.

35. Clark, T. J., Phillips, R. S., Bundy, B. M., Momany, C. & Neidle, E. L. (2004). Benzoate decreases the binding of cis, cis-muconate to the BenM regulator despite the synergistic effect of both compounds on transcriptional activation. *Journal of bacteriology* **186**, 1200.

36. Lönneborg, R., Smirnova, I., Dian, C., Leonard, G. A. & Brzezinski, P. (2007). In Vivo and in Vitro Investigation of Transcriptional Regulation by DntR. *Journal of Molecular Biology* **372**, 571-582.

37. Stec, E., Witkowska-Zimny, M., Hryniewicz, M. M., Neumann, P., Wilkinson, A. J., Brzozowski, A. M., Verma, C. S., Zaim, J., Wysocki, S. & D Bujacz, G. (2006). Structural basis of the sulphate starvation response in E. coli: crystal structure and mutational analysis of the cofactor-binding domain of the Cbl transcriptional regulator. *Journal of molecular biology* **364**, 309-322.

38. Ishihama, A. (1993). Protein-protein communication within the transcription apparatus. *Journal of bacteriology* **175**, 2483.

39. McFall, S. M., Chugani, S. A. & Chakrabarty, A. M. (1998). Transcriptional activation of the catechol and chlorocatechol operons: variations on a theme1. *Gene* **223**, 257-267.

40. Tao, K., Zou, C., Fujita, N. & Ishihama, A. (1995). Mapping of the OxyR protein contact site in the C-terminal region of RNA polymerase alpha subunit. *Journal of bacteriology* **177**, 6740.

41. Tao, K., Fujita, N. & Ishihama, A. (1993). Involvement of the RNA polymerase subunit C-terminal region in co-operative interaction and transcriptional activation with OxyR protein. *MOLECULAR MICROBIOLOGY* **7**, 859-859.

42. Gussin, G. N., Olson, C., Igarashi, K. & Ishihama, A. (1992). Activation defects caused by mutations in Escherichia coli rpoA are promoter specific. *Journal of bacteriology* **174**, 5156.

43. Jafri, S., Urbanowski, M. L. & Stauffer, G. V. (1995). A mutation in the rpoA gene encoding the alpha subunit of RNA polymerase that affects metE-metR transcription in Escherichia coli. *Journal of bacteriology* **177**, 524.

44. Lochowska, A., Iwanicka-Nowicka, R., Zaim, J., Witkowska-Zimny, M., Bolewska, K. & Hryniewicz, M. M. (2004). Identification of activating region (AR) of Escherichia coli LysR-type transcription factor CysB and CysB contact site on RNA polymerase alpha subunit at the cysP promoter. *Molecular microbiology* **53**, 791-806.

45. Wang, X., Mukhopadhyay, P., Wood, M. J., Outten, F. W., Opdyke, J. A. & Storz, G. (2006). Mutational analysis to define an activating region on the redox-sensitive transcriptional regulator OxyR. *Journal of bacteriology* **188**, 8335.

46. Monferrer, D., Tralau, T., Kertesz, M. A., Dix, I., Solà, M. & Usón, I. (2010). Structural studies on the full-length LysR-type regulator TsaR from Comamonas testosteroni T-2

reveal a novel open conformation of the tetrameric LTTR fold. *Molecular Microbiology* **9999**.

47. Monferrer, D., Tralau, T., Kertesz, M. A., Panjikar, S. & Uson, I. (2008). High crystallizability under air-exclusion conditions of the full-length LysR-type transcriptional regulator TsaR from Comamonas testosteroni T-2 and data-set analysis for a MIRAS structure-solution approach. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **64**, 764-769.

48. Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K. & Senda, T. (2003). Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. *Journal of molecular biology* **328**, 555-566.

49. Muraoka, S., Okumura, R., Uragami, Y., Nonaka, T., Ogawa, N., Miyashita, K. & Senda, T. (2003). Purification and crystallization of a LysR-type transcriptional regulator CbnR from Ralstonia eutropha NH9. *Protein and Peptide Letters* **10**, 325-329.

50. Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS microbiology reviews* **29**, 231-262.

51. Brennan, R. G. & Matthews, B. W. (1989). The helix-turn-helix DNA binding motif. *J. Biol. Chem* **264**, 1903-1906.

52. Lochowska, A., Iwanicka-Nowicka, R., Plochocka, D. & Hryniewicz, M. M. (2001). Functional Dissection of the LysR-type CysB Transcriptional Regulator. *Journal of Biological Chemistry* **276**, 2098-2107.

53. Kullik, I., Stevens, J., Toledano, M. B. & Storz, G. (1995). Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for DNA binding and multimerization. *J. Bacteriol.* **177**, 1285-1291.

54. Tyrrell, R., Verschueren, K. H. G., Dodson, E. J., Murshudov, G. N., Addy, C. & Wilkinson, A. J. (1997). The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. *Structure* **5**, 1017-1032.

55. Smirnova, I. A., Dian, C., Leonard, G. A., McSweeney, S., Birse, D. & Brzezinski, P. (2004). Development of a bacterial biosensor for nitrotoluenes: the crystal structure of the transcriptional regulator DntR. *Journal of molecular biology* **340**, 405-418.

56. Choi, H. J., Kim, S. J., Mukhopadhyay, P., Cho, S., Woo, J. R., Storz, G. & Ryu, S. E. (2001). Structural basis of the redox switch in the OxyR transcription factor. *Cell* **105**, 103-113.

57. Zhou, X., Lou, Z., Fu, S., Yang, A., Shen, H., Li, Z., Feng, Y., Bartlam, M., Wang, H. & Rao, Z. (2009). Crystal Structure of ArgP from Mycobacterium tuberculosis Confirms Two Distinct Conformations of Full-length LysR Transcriptional Regulators and Reveals Its Function in DNA Binding and Transcriptional Regulation. *Journal of Molecular Biology*.

58. Sainsbury, S., Lane, L. A., Ren, J., Gilbert, R. J., Saunders, N. J., Robinson, C. V., Stuart, D. I. & Owens, R. J. (2009). The structure of CrgA from Neisseria meningitidis reveals a new octameric assembly state for LysR transcriptional regulators. *Nucleic Acids Research*.

59. Colyer, T. E. & Kredich, N. M. (1996). In vitro characterization of constitutive CysB proteins from Salmonella typhimurium. *Molecular microbiology* **21**, 247-256.

60. Collier, L. S., Gaines, G. L., III & Neidle, E. L. (1998). Regulation of Benzoate Degradation in Acinetobacter sp. Strain ADP1 by BenM, a LysR-Type Transcriptional Activator. *J. Bacteriol.* **180**, 2493-2501.

61. Romero-Arroyo, C. E., Schell, M. A., Gaines, G. L., 3rd & Neidle, E. L. (1995). catM encodes a LysR-type transcriptional activator regulating catechol degradation in Acinetobacter calcoaceticus. *J. Bacteriol.* **177**, 5891-5898.

62. Harwood, C. S. & Parales, R. E. (1996). The -ketoadipate pathway and the biology of self-identity. *Annual Reviews in Microbiology* **50**, 553-590.

63. Ornston, L. N. & Neidle, E. L. (1991). Evolution of genes for the-ketoadipate pathway in Acinetobacter calcoaceticus. *The Biology of Acinetobacter. Plenum Press. New York, NY*, 201–237.

64. Ezezika, O. C., Collier-Hyams, L. S., Dale, H. A., Burk, A. C. & Neidle, E. L. (2006). CatM Regulation of the benABCDE Operon: Functional Divergence of Two LysR-Type Paralogs in Acinetobacter baylyi ADP1. *Appl. Environ. Microbiol.* **72**, 1749-1758.

65. Clark, T., Haddad, S., Neidle, E. & Momany, C. (2003). Crystallization of the effector-binding domains of BenM and CatM, LysR-type transcriptional regulators from Acinetobacter sp. ADP1. *Acta Crystallographica Section D: Biological Crystallography* **60**, 105-108.

66. Craven, S. H., Ezezika, O. C., Haddad, S., Hall, R. A., Momany, C. & Neidle, E. L. (2009). Inducer responses of BenM, a LysR-type transcriptional regulator from Acinetobacter baylyi ADP1. *Molecular Microbiology* **72**, 881-894.

CHAPTER 2

FULL-LENGTH STRUCTURES OF BENM AND TWO VARIANTS REVEAL DIFFERENT OLIGOMERIZATION SCHEMES FOR LYSR-TYPE TRANSCRIPTIONAL REGULATORS

[1]Ruangprasert, A., S.H. Craven, E.L. Neidle, and C.A. Momany. Submitted to *Journal of Molecular Biology*

Introduction

LysR-type transcriptional regulators (LTTRs) form the largest family of homologous regulators in prokaryotes.[1; 2; 3] Bioinformatic analysis reveals individual genomes that may encode more than 100 such regulators with distinct functions.[1] Considering their abundance and widespread distribution, it is not surprising that LTTRs regulate all types of biological processes. This investigation focuses on BenM, which regulates aromatic compound catabolism by *Acinetobacter baylyi*, a soil bacterium that encodes 44 LTTRs.[4; 5] Previous comparisons of BenM to CatM, a closely related paralog with overlapping function, lay the foundation for these structural studies of the full-length BenM protein.[5; 6; 7; 8; 9]

BenM and CatM have the same organization as other LTTRs in which sequence conservation is greatest in the DNA-binding domain (DBD). This region consists of approximately 66 N-terminal residues that bind DNA with a winged helix-turn-helix (wHTH) motif.[3; 10; 11] A linker helix (LH) joins the DBD to the effector-binding domain (EBD), which can be further divided into two regions (EBD-I and EBD-II) that are held together by hinge-like β-strands.[3; 7; 11; 12] Each LTTR contains circa 300 amino acid residues and usually functions as a tetramer. Typically, the binding of a small molecule effector to the EBD causes a conformational change that alters the position of the LTTR on promoter-region DNA and enables transcriptional activation.[13] The *benM* and *catM* genes, like those encoding many LTTRs, are negatively autoregulated and located divergently to genes whose expression they activate.[4]

Despite ongoing efforts, it has been difficult to determine how effector-binding signals are transmitted in LTTRs to alter the interactions between DNA, the subunits of the regulator, and RNA polymerase. To date, no structures of LTTR-DNA complexes have been published, and most crystal structures of LTTRs lack the DBD and LH regions. Among the EBDs,

structural conservation is high, even between proteins that share little sequence similarity.[7; 11; 12; 14; 15] The EBD resembles a periplasmic-binding protein in which a small molecule effector can bind between EBD-I and EBD-II.[16] We showed that this cleft is used by BenM and CatM to bind *cis,cis*-muconate (muconate), a metabolite that induces benzoate degradation.[7] Genetic evidence suggests that an effector-binding site is similarly located in many family members.[3; 4; 7]

BenM has a secondary binding site for benzoate in its EBD-II.[7] This hydrophobic pocket enables BenM to activate very high levels of transcription in response to the combination of benzoate and muconate. Thus, a single regulator integrates signals from different metabolic steps in a pathway.[6; 13] BenM and CatM regulate many genes, including the *benABCDE* operon. At this locus, BenM represses expression when effectors are absent.[5; 13] In response to benzoate, muconate, or both, BenM shifts on the DNA, and transcription can be activated. At the same *benA* promoter, CatM regulates differently from BenM despite their similar amino acid sequences.[9] Like BenM, CatM can activate *benA* expression in response to muconate, albeit at relatively low levels. However, CatM fails to respond to benzoate.[5; 9]

Additional studies are needed to understand how effector binding impacts protein conformation and transcription. Four full-length LTTR structures have been reported, CbnR, CrgA, ArgP, and TsaR.[11; 17; 18; 19] However, their interactions with effectors remain incompletely understood. In protein crystals, CbnR uses a homo-tetrameric oligomerization scheme whereas CrgA exists as a homo-octamer. Moreover, CrgA subunits display a single conformation while two such conformations are observed for CbnR, ArgP, and TsaR. [11; 17; 18; 19] These conformations differ in the way the LH orients the DBD relative to the EBD.

To achieve a more thorough understanding of LTTRs, we studied full-length BenM and two variants, BenM(E226K) and BenM(R156H), that each activate *benA* expression without

benzoate or muconate.[6] Here, the structures of the BenM proteins are reported and compared to other LTTRs, including family members in the Protein Data Bank (PDB) that have not been published. LTTR oligomerization schemes are discussed.

Results

Crystallization of histidine-tagged BenM, BenM(E226K), and BenM(R156H)

C-terminally hexahistidine-tagged full-length BenM (BenM-His) was expressed in *Escherichia coli* and purified as previously described.[13] Proteins were maintained as dilute solutions (0.3-0.5 mg ml^{-1}) until just before use. Immediately after concentration (2-5 mg ml^{-1}), microbatch crystallization trials at 295 K were done with commercial crystallization kits. Two conditions quickly yielded microcrystals, condition 19 (CSI-19) and condition 31 (CSII-31) from Hampton Research Crystal Screens 1 and 2, respectively. The incubation temperature, precipitant composition, and protein concentrations were then optimized. Crystals reached a final size of approximately 0.05 x 0.05 x 0.07 mm^3 within 7 days. When these crystals were used in "microseed matrix" techniques, crystals were generated from additional cocktails but no obvious improvements justified using alternative conditions.[20]

Prior to the in-house manual screens, high-throughput screens of 1536 conditions were conducted with BenM-His using microbatch under oil methods at the Hauptman-Woodward Medical Institute.[21; 22] While this high-throughput approach worked well for the effector-binding domains of BenM and CatM,[23] the concentrated samples of full-length protein that were shipped off-site did not crystallize. Nevertheless, an identical batch of purified protein that failed in high-throughput screens did form crystals when freshly concentrated and prepared in house as described above.

The manual crystallization approach used for BenM-His was repeated with two variants, each having a single amino acid replacement. The E226K variant crystallized with only one condition in microbatch setups (CSII-31). With this cocktail, the hanging drop vapor diffusion method for crystallization yielded much larger crystals.[24] BenM(E226K)-His crystals were obtained with an average dimension of 0.05 x 0.07 x 0.1 mm^3. For the R156H variant, crystals did not form using commercial kits. However, the addition of small amounts of Crystal Screen condition 6 or Crystal Screen 2 condition 26, which introduced small amounts of ammonium sulfate to the CSII-31 condition, triggered protein crystallization. Optimization of the conditions yielded BenM(R156H)-His crystals with the average dimension of 0.07 x 0.1 x 0.1 mm^3.

Structure determination

We report the refinement statistics for data with the best resolution and lowest R_{merge} values after several data sets were collected (Table 2.1). All sets had relatively large $\chi 2$ values (~2-4) during data processing (HKL2000 in house and HKL3000 beamline). Moreover, split diffraction spots indicated that the E226K variant crystal was broken or a mosaic.

Space group choices were ambiguous. Data could be processed in an orthorhombic space group (*mmm* Laue symmetry, R_{merge} 0.076) with nearly identical *a* and *b* cell constants or a tetragonal space group (R_{merge} 0.114 for 4/*m* or 0.096 4/*mmm*) regardless of the crystallization condition. Based on previous LTTR studies[11; 19], it seemed likely that there would be two conformations of BenM subunits. Thus, at least two molecules were expected per asymmetric unit. The orthorhombic space groups were favored since, unlike the tetragonal ones, they could accommodate two BenM-His molecules in the asymmetric unit.[25]

When molecular replacement methods were used with various search models, including CbnR, the best choices were two BenM-EBD structures (PDB IDs: 2F97 and 2F8D).[26] The solutions in space group $P222_1$ positioned two BenM-EBD subunits in the asymmetric unit related by non-crystallographic two-fold symmetry. Single EBD molecules could also be readily placed in space group $P4_222$ with a low refined R_{factor}. The contacts between the two EBDs were the same as those previously observed for BenM at an interface resembling that of CbnR.[11; 26]

To generate phases to orient the DBD and LH regions, two search models were used simultaneously, a BenM-EBD structure (PDB ID: 2F97) and a recently determined BenM-DBD dimer structure (unpublished data). A full-length BenM-His model was obtained that revealed two BenM-His molecules in the asymmetric unit possessing two distinct arrangements of the DBD with respect to the EBD. These compact and extended conformations of the full-length protein resembled the two conformations of full-length CbnR.[11] With additional refinement and the introduction of water molecules, which helped with problematic electron density issues at the connections between the EBDs and LH regions, a complete model was generated.

Despite completion of a full model, spurious density features and relatively high $R_{factors}$ suggested twinning issues in the crystals. Tests with the $P222_1$ data (see Materials and Methods) supported the inference of twinning in the crystals of the native and variant BenM proteins. A pseudo-merohedral twin element applied to nearly identical cell constants created the symmetry features that allowed the data to be processed in the tetragonal space groups, which have higher symmetry than the appropriate orthorhombic space group. The twin operation (*k, h, -l*), when applied to an asymmetric unit, would superimpose the EBDs while swapping the orientation of the DBDs (Figure 2.1A). The addition of the DBDs to the EBD model and introduction of twin refinement resolved the electron density problems and vastly improved the features and statistics

of the BenM-His and variant models (Table 2.1, Figure 2.1). Careful evaluation of the local structures around the replacement sites of BenM(E226K)-His and BenM(R156H)-His revealed density maps and features consistent with the expected residues of each protein (Figure S1).

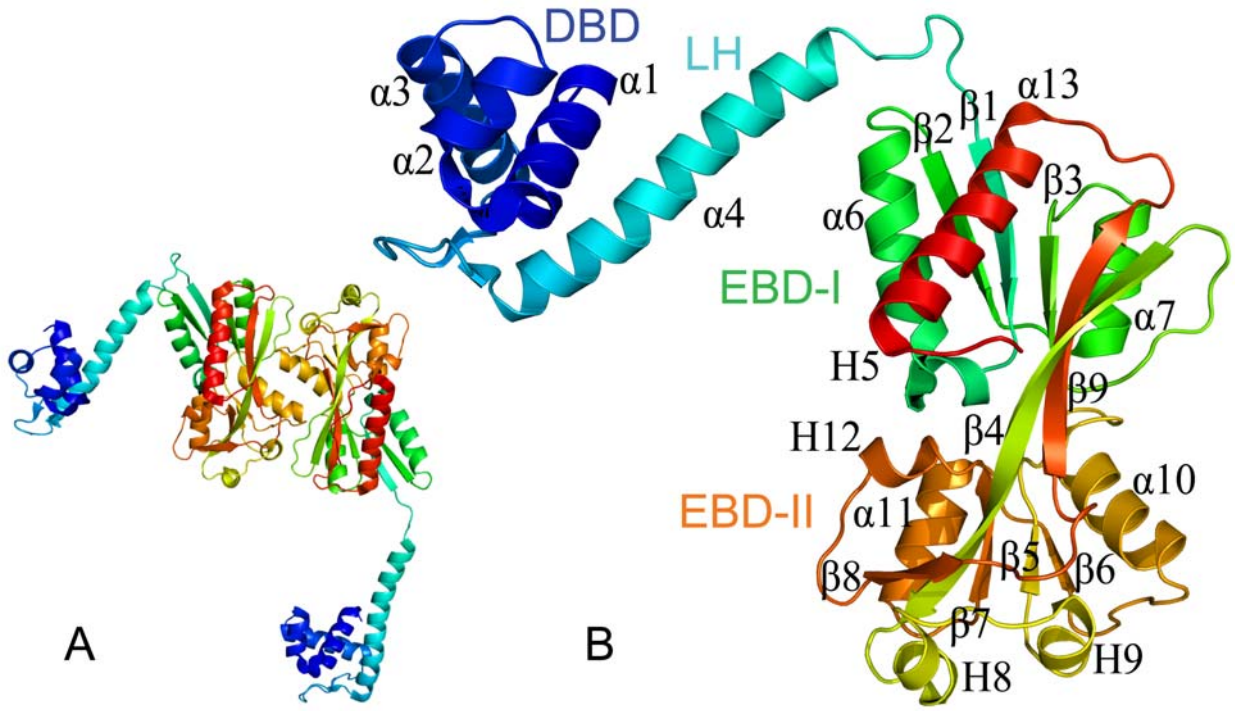

Figure 2.1. Ribbon representations of the full-length BenM-His structure. Panel A shows the two molecules of the BenM-His asymmetric unit colored blue to red, N-terminus to C-terminus. The left-most subunit is in the compact conformation of BenM-His, while the right subunit is in the extended conformation. Panel B shows the domain organization of a full-length BenM-His subunit with the secondary structure features numbered consecutively from N- to C-terminus of the full-length protein. Abbreviations: alpha helix (α), 3_{-10} helix (H), and beta strand (β).

Table 2.1. Crystallographic data collection and refinement statistics

	BenM-His (PDB ID: 3K1N)	BenM(R156H)-His (PDB ID: 3K1M)	BenM(E226K)-His (PDB ID: 3K1P)
Beamline[a]	19-ID	19-ID	22-ID
Wavelength (Å)	1.00800	0.97934	0.99999
Cell constants (Å)[b]	a = 70.24, b = 70.10, c = 187.87	a = 70.00, b = 70.79, c = 186.28	a = 70.60, b = 70.70, c = 187.28
Space group	$P222_1$	$P222_1$	$P222_1$
Resolution range (Å)	39.04-2.99 (3.1-2.9)	49.8-2.3 (2.34-2.30)	48.3-3.0 (3.1-3.0)
Completeness (%)	86.8 (82.8)	97.8 (69.6)	90.3 (70.2)
R_{merge} (%)[c]	7.6 (30.4)	7.1 (50.7)	8.70 (29.6)
Average I/σI	15.4 (4.1)	13.7 (3.8)	18.9 (4.7)
R_{value} (%)[d]	15.6 (20.0)	14.8 (19.6)	19.9 (26.5)
R_{free} (%)[e]	18.0 (16.8)	17.9 (25.1)	22.6 (43.2)
Number of solvent atoms	263	520	373
Other molecules	2 Cl^-, 2 imidazole	2 Cl^-, 2 imidazole, 2 glycerol	None
r.m.s.d. bond lengths (Å)	0.009	0.008	0.008
r.m.s.d. bond angles (°)	1.221	1.037	1.125
Average B-factors ($Å^2$)	32.7	26.0	38.1
Ramachandran distribution			
Favored (%)	94.5	98.2	90.4
Allowed (%)	4.3	1.8	6.8
Outlier (%)	1.2	0.0	2.8
Refined twin fraction (*k*, *h*, *-l*)	0.425	0.386	0.402

[a] At the Advanced Photon Source, Argonne, IL, USA, beamline 19-BM was operated by Structural Biology Consortium Collaboratory Access Team (SBC-CAT) and 22-ID by the SouthEast Regional Collaboratory Access Team (SER-CAT).

[b] Cell constants for *a* and *b* were iteratively optimized for native BenM-His (from *a* = 70.943, *b* = 70.993), BenM(E226K)-His (from *a* = 70.803, *b* = 70.648), and BenM(R156H)-His (from *a* = 70.302, *b* = 70.485)

[c] R_{merge} is the unweighted R value on I between symmetry related reflections.

[d] $R_{value} = \sum_{hkl} | Fobs(hkl) - Fcalc(hkl) | / \sum_{hkl} Fobs(hkl)$ for reflections in the working data set.

[e] 5% of the reflections were used in the cross-validation data set.

Structural Features and Differences Between The BenM Structures

Characteristic of LTTRs, the BenM-His and variant structures had an N-terminal DBD with a wHTH motif (residues 1-59) connected via the LH (residues 60-86) to a C-terminal EBD (87-300). The EBD is often referred to as the regulatory domain (RD), but we avoid such terminology since DNA binding is also part of the regulatory response. In the periplasmic-binding-protein fold of BenM-EBD, a hinge separates two subdomains, domain I (residues 87-161 and 268-300) and domain II (residues 162-267).[6; 7; 26] While amino acid residues between 1 and 302 had visible electron density in at least one subunit, several regions were poorly defined. The disordered regions included the loop between the DBD and LH (residues 48-58), the hinge connecting LH to EBD I (residues 86-89), and several C-terminal residues (300-306).

These new full-length structures were compared to the EBD structures of wild-type and variant BenM crystallized under different conditions with and without effectors.[6; 7; 26] Domain I regions were aligned to that of the wild-type BenM-EBD with invariant core analysis from

Bio3D[27] as previously described.[6] The full structures were then superimposed to assess the positions of EBD-II relative to EBD-I. By this analysis, the full-length BenM-His structure most closely resembled wild-type BenM-EBD crystallized under a high pH condition without effectors (PDB ID: 2F97), as indicated by a root mean square deviation (RMSD) of 0.364 Å.[26]

The EBDs of the full-length variants adopted similar conformations to the wild-type EBD (PDB ID: 2F97) with RMSDs of 0.428 Å and 0.323 Å. When these variant regions were compared to that of the full-length wild-type BenM-His, the RMSDs were also low, 0.330 Å for BenM(E226K)-His and 0.356 Å for BenM(R156H)-His. Hereafter, structural analysis of full-length BenM-His will be based mainly on the higher resolution structure of BenM(R156H)-His, since both had the same crystal packing characteristics and nearly identical structures.

Oligomerization of BenM

BenM(R156H)-His crystallized as a dimer in the asymmetric unit (Figure 2.1A). While the complete subunits displayed two conformations, the EBDs were related by noncrystallographic two-fold symmetry. In this dimer, the EBD-II/EBD-II interface closely resembled one observed previously in studies of BenM-EBD that likely represents the interface between EBD dimers in a functional tetramer.[11; 26] The following description of this interface in the BenM(R156H)-His structure refers to the secondary structure features shown in Figure 2.1. The interface was located mostly on the surface of H9, α10, and the loop region between β3 and β4. It buried an area of 866.73 $Å^2$. Residues contributing hydrogen bond interactions between the EBD-II subunits included Lys148, Ser150, Asn185, Asn209, Ser212, Asp213, and Asp262. We will refer to this as the α10-α10 tetramer interface. This interface was only observed previously in the EBD structures of proteins that crystallized in higher pH conditions.

In the crystal lattice, another interface between EBD subunits corresponded to one consistently observed in all known LTTR structures, both full-length and EBDs.[26] It has been termed the dimer- or RD dimer-interface. In BenM(R156H)-His, this interface spanned a contact area of 1321.1 $Å^2$, and it involved interactions between α6, β2, α11, and H12 of each subunit. This interface is organized by subunits arranged in an antiparallel fashion. For clarity, we will refer to this as the α6/β2-α11 dimer interface.

Gel filtration suggested that BenM-His forms tetramers in solution[13], and thus, closed BenM-His tetramers were expected in the crystal lattice. Attempts to identify such a "closed" tetramer were unsuccessful after examining all symmetry-related options (and twinning). Two DBDs of a protein dimer projected in diverse directions and interacted with DBDs from distant dimers (Figure 2.2). In the structure of CbnR, two conformations of the protein are observed at the α6/β2-α11 dimer interface, while the α10-α10 tetramer interface had two-fold symmetry. Thus, CbnR associates as a "dimer of dimers" in a closed tetramer. In constrast, a crystallographic two-fold at the BenM-His α6/β2-α11 dimer interface prevents the same packing in the BenM structures and creates an oligomerization scheme that does not close. We did not expect to see this interlocking pattern in crystals as the phenomenon seemed unlikely to support crystal growth. However, within the crystallization cocktails, gentle dissociation of the tetrameric species may have occurred with reassociation taking place to form the crystal lattice.

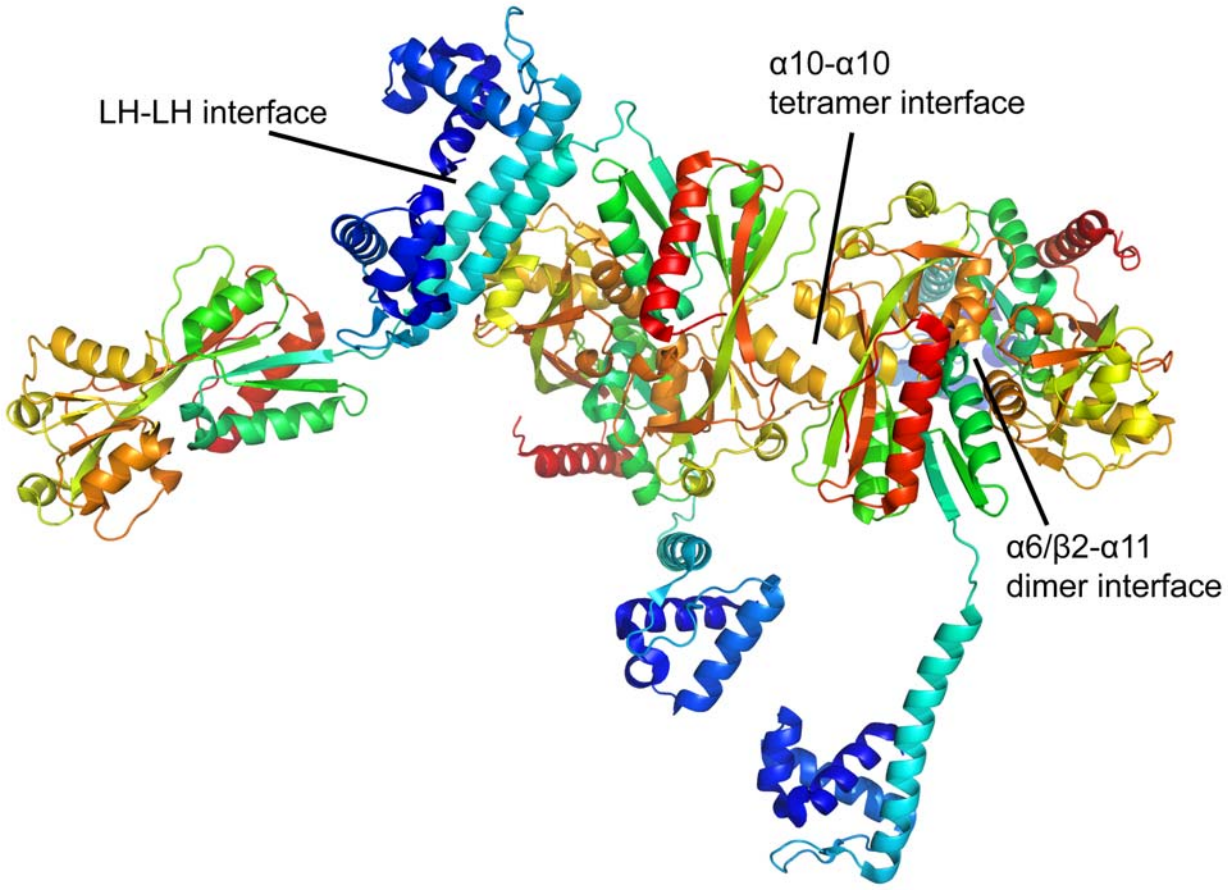

Figure 2.2. Interlocking BenM-His molecules form the crystal lattice via common LTTR interfaces. A BenM-His dimer in one asymmetric unit interacts with neighboring BenM-His molecules through the α6/β2-α11 dimer interfaces of the EBDs. The α6/β2-α11 interface is observed in all known LTTR crystal structures. The two chains of the asymmetric unit interact with one another at the α10-α10 tetramer interface. Whereas the EBDs are related by two-fold symmetry, the LH/DBDs assume two conformations.

Structural comparisons with other full-length LTTRs

We sought full-length LTTR structures to compare with those of BenM and its variants. In addition to published reports of several complete structures (CbnR, CrgA, TsaR, and ArgP), a

full-length DntR structure was mostly characterized despite an inability to refine its DBD.[11; 15; 17; 18; 19; 28; 29] Further searches of the PDB using key words and DALI methods[30] revealed additional structures of proteins likely to be LTTRs that had not been reported elsewhere at the time of our analysis. Two proteins from *Pseudomonas aeruginosa*, with PDB IDs 2ESN and 3FZV, were found to have features of LTTRs, including more than 20% sequence identity to another LysR member and a conserved N-terminal DBD.[3; 31] One structure has been in the protein databank for a number of years without due recognition (2ESN), while the other (3FZV) was separately identified as a LTTR in studies of TsaR, where it was referred to as PA01-PR.[18] 2ESN, which is 34% identical in sequence to DntR, and 3FZV, which is 33% identical in sequence to CbnR, have typical LTTR domain organizations, but are not characterized with regard to function. In each structure, the N-terminal DBD is connected via a LH to a C-terminal domain with a periplasmic-binding protein fold. The 2ESN structure is unique among known LTTR structures in having an extra helix at the N-terminus of its wHTH DNA-binding motif (residues 1-10). However, such insertions are evident in the sequences of other family members (data not shown).

To evaluate structural differences among the full-length proteins, the EBD-I portion of CbnR was set as a fixed reference structure, and the comparable regions of 2ESN, 3FZV, CrgA, ArgP, TsaR, and BenM were aligned to it using invariant core analysis.[27] As discussed below, there were notable differences in domain arrangements. We also compared the interfaces between subunits of different LTTRs to assess variation in the overall conformation of the regulatory oligomers. Protein interfaces were analyzed for CbnR, DntR, 2ESN, 3FZV, CrgA, ArgP, BenM, OxyR, and CysB using the PROTORP and PISA v1.18 servers.[32; 33] Many LTTRs, such as CbnR, appear to function as a tetramer formed by association of two dimers, each composed of one subunit in an extended conformation and the other in a compact formation.[11]

Others utilize DBD-DBD interfaces and separately, α6/β2-α11 dimer interfaces. We evaluated the contacts between different subunits used to form dimers (the α6/β2-α11 dimer interface) and tetramers (the α10-α10 tetramer interface) (Tables 2.2, 2.3 and 2.4, supplementary materials).

Discussion

The crystal structures of full-length BenM-His and its variants display the domain organization and protein folds characteristic of LTTRs.[4] While DNase I footprint experiments and gel filtration suggest that BenM functions as a tetramer and exists as a tetramer in solution,[13] the protein molecules in crystals formed high-order oligomers. Interlocking arrays of BenM-His appear to result from a high degree of rotation at the hinge between the LH and the EBD. The flexibility of this region is biologically relevant, because it allows the DNA-bound tetramer to assume distinct conformations to repress or activate transcription.[11; 18] Yet, the flexible nature of the LH combined with high protein concentrations in heterologous protein expression and crystallization may enhance opportunities for adjacent BenM-His molecules to become entangled. The observed interlinking of subunits supports our model that DBD swapping among neighboring tetrameric units contributes to the problematically low solubility of many LTTRs.[26] While such oligomers do not represent biologically functional units, this type of interconnected array may be seen in other structures due to the innate flexibility of the EBD-LH region.

Different arrangements of the LH-DBD regions in various LTTRs

To compare the relative positions of the LH and DBD regions in different LTTRs, the EBD-I regions of the full-length forms of BenM, TsaR, ArgP, 3FZV, 2ESN, and CrgA were aligned with that of a CbnR subunit. With the exception of CrgA, each LTTR has two

conformations of subunits, an extended and a compact form, as seen in CbnR.[11] The single conformation of the CrgA crystal structure corresponds to the compact form.[17] These comparisons revealed variation in the LH arrangements of different LTTRs in both conformations (Figure 2.3). The positions the LHs started to diverge from each other at the hinges near the N-termini of the EBDs (around residue 89 in BenM). In the compact forms, the LHs of 2ESN, 3FZV, and TsaR orient their DBDs in an overall similar direction, albeit with different angles relative to the LH of CbnR (Figure 2.3A). However, the LHs of BenM and CrgA project in completely different directions from each other and from the other LHs depicted. In the case of CrgA, the observed angle allows formation of an octameric oligomer.

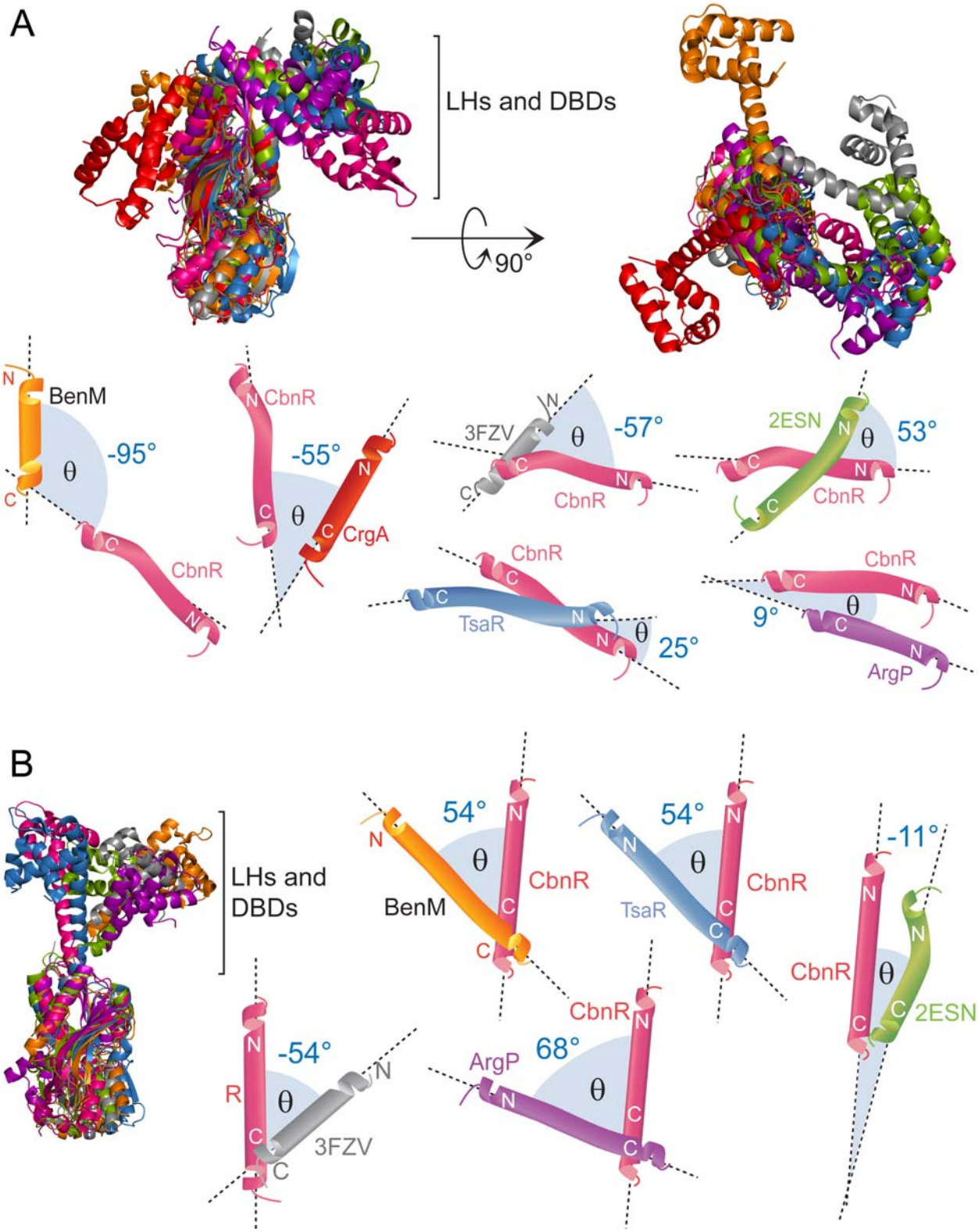
A
LHs and DBDs
90°
BenM
-95°
CbnR
CbnR
-55°
CrgA
3FZV
-57°
CbnR
2ESN
53°
CbnR
CbnR
TsaR
25°
CbnR
9°
ArgP
B
LHs and
DBDs
54°
CbnR
BenM
54°
CbnR
TsaR
-11°
CbnR
2ESN
-54°
3FZV
CbnR
ArgP
68°

Figure 2.3. Alignment of different LTTRs using invariant core analysis with the EBD-I region of CbnR as a fixed reference. Compact conformations are shown in panel A and extended conformations in panel B. The position of the CbnR LH is compared to that of the corresponding helix of other LTTRs, with angles calculated as described in materials and methods. In the crystal lattice, CrgA assumes a single conformation resembling a compact subunit (panel A). In both conformations, additional rotation of the DBDs can be seen with respect to the fixed EBD such that the structures ultimately differ substantially in the final oligomer.

In the extended protein conformations, the angles between each LH and that of CbnR differed in a less extreme fashion than observed for the compact forms of the subunits (Figure 2.3B). Nevertheless, it is evident that this LTTR hinge region is highly flexible. When the amino acid sequences of the different LTTRs were aligned, we could find no evidence of a consensus sequence for the residues at the hinge regions or residues involved in the interactions between the LHs and EBDs (data not shown).

Interfaces between the subunits of LTTRs

To determine the effect of subunit differences on oligomeric conformations, we assessed the formation of dimers and tetramers. As described earlier, interactions between a compact-form subunit and an extended-form subunit can occur in a region designated the α6/β2-α11 dimer interface. All full-length LTTR structures available to date use a similar α6/β2-α11 dimer interface that buries more than 1000 $Å^2$ of surface area. Despite this similarity, which is evident in the appearance of different dimers, the exact nature of this interface is variable. Different secondary structures may contribute to interactions, and variations are seen in the calculated gap volume. Yet in all structures, two subunits align at the dimer interface in an anti-parallel manner

with two-fold rotational symmetry of the EBDs. While the DBDs of CrgA share the same rotational axis as the two-fold related EBDs, variations were observed for other LTTRs in the relative positions of the DBDs with respect to the EBD. For example, in CbnR, there are two orientations of the DBD with respect to the EBD. Further, even the EBDs are not always arranged with perfect two-fold crystallographic symmetry. In some structures, such as the reduced form of a variant OxyR[12] and BenM-EBD bound with muconate,[7] the ligand binding and overall conformation is subtly different between the two subunits of the dimer.

When dimers combine to form a tetramer or larger oligomer, they can interact in another region, designated the α10-α10 tetramer interface. This interface is similar in the structures of CbnR, DntR and BenM, and it lies on the surface of the EBDs burying an area of over 800 $Å^2$. Nevertheless, the angle at which the subunits of CbnR, DntR, or BenM cross each other at the α10-α10 tetramer interface varies substantially (Figure 2.4). Since a complete tetramer has not yet been characterized for DntR or BenM, it is unclear how these differences affect the shapes or characteristics of the entire oligomer. Furthermore, this interface is likely to be influenced by effector binding due to the close proximity of the effector-binding sites to the interface. While the significance of this α10-α10 tetramer interface has not been demonstrated for OxyR, the possibility of a functional role is suggested by the use of this interface to pack the EBDs of a reduced variant OxyR in the crystal lattice.[12]

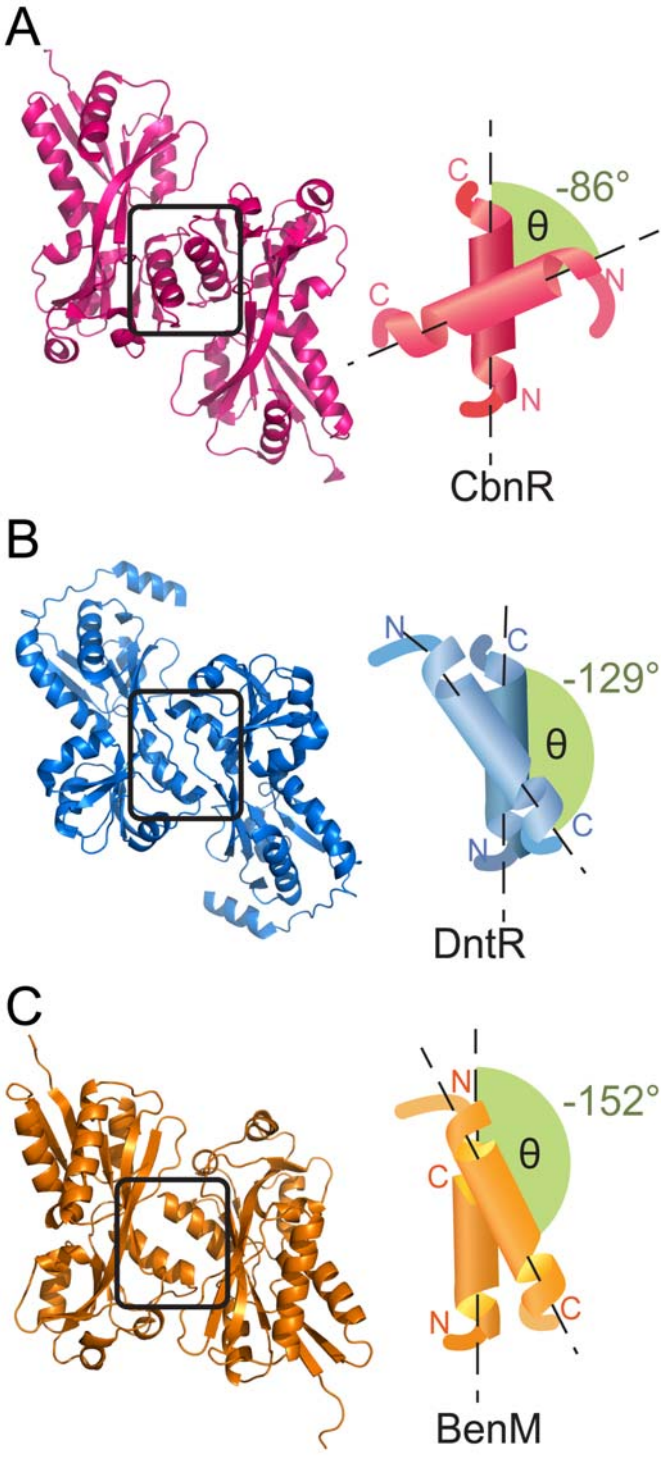
A
C
-86°
θ
C
N
N
CbnR
B
N
C
-129°
θ
N
C
DntR
C
N
-152°
θ
C
N
C
BenM

Figure 2.4. Comparison of the α10-α10 tetramer interfaces of CbnR, DntR, and BenM. For each LTTR, the α10 helix from the EBD-II region of two subunits forms an interface that can help join two heterodimers into a tetramer. The center of the interface is boxed (at the left), and the angle at which the two α10 helices cross each other is shown (at the right).

Oligomerization schemes

The α10-α10 tetramer interface for CbnR has been suggested to be weak,[18] and additional DBD interactions appear to be important for tetramer stability. We introduce here a new designation, Scheme I, to indicate the method by which CbnR forms a tetramer via interactions between the EBDs and DBDs of two heterodimers (Figure 2.5A, 2.5D). Yellow coloration in the image highlights the structurally conserved α10-α10 tetramer interface that marks the contact area between the EBDs of CbnR, DntR, and BenM (Figure 2.5D). However, the remaining known LTTR structures do not use Scheme I to form oligomers. In our analyses, we identified two other, previously uncharacterized, schemes for LTTR assembly.

In the structure of 2ESN, which is an apparent LTTR from *P. aeruginosa* that crystallized as a tetramer, two heterodimers are held together by interactions between the DBD and LH regions without using the α10-α10 tetramer interface (Figure 2.5B). This organization, also used by ArgP and TsaR, was designated Scheme II (Figure 2.5D). While the α6/β2-α11 dimer interfaces are used for assembly in combination with the DBD/LH interactions, the EBDs of the two heterodimers are not in contact with one another. In ArgP, contacts between the EBDs and DBDs were to shown by mutational analysis to contribute to dimer formation,[19] and similar interactions could be important for other LTTRs that employ Scheme II oligomerization. Despite its propensity to form an octamer rather than a tetramer, CrgA also assembles using

Scheme II. CrgA differs from the other LTTRs in the angular relationship of the LH to the EBD (Figure 2.5E). Additional variations among the Scheme II proteins are notable in the orientation of the DBDs. While 2ESN and ArgP both create tetramers with large voids in the middle, the DBDs are oriented differently. The DBDs of ArgP reside on one surface in an almost linear fashion, while 2ESN presents its DBDs in a sharp curve that could only accommodate bent DNA.

In contrast to Scheme II, the EBDs of two heterodimers interact extensively in Scheme III, an arrangement illustrated by the structure of 3FZV, another putative LTTR of *P. aeruginosa* (Figure 2.5C, 2.5D). In the 3FZV tetramer, all EBDs pack closely against one another to create a large interacting surface. 3FZV was reported to have a weak EBD-II/EBD-II interface, indicated by a complexation significance score (CSS) of 0.049.[18] However, this calculation applies only to the EBD-II/EBD-II contacts of the α10-α10 interface, while there are many additional EBD-I/EBD-II contacts in this tetramer. By our calculations, the non-DBD interfaces (of the A-B and C-D subunits) have an average surface area of 875.455 Å^2. Because these interfaces are related by two-fold symmetry, the total interaction surface of 3FZV is substantial.

Despite the similarity of global characteristics and shared features of dimer interfaces, the known structures of full-length LTTRs display three different schemes for forming oligomers (Figure 2.5). While the dimer and tetramer interfaces appear labile in some structures, the LH/DBD interfaces are consistently strong. In all the full-length structures, interface surface areas are in the range of 1200 Å^2 (Table 2.3, supplementary materials). Thus, the LH/DBD interface may be the most significant contributor to oligomer assembly. This N-terminal protein region clearly has the highest level of sequence conservation among diverse LTTRs.

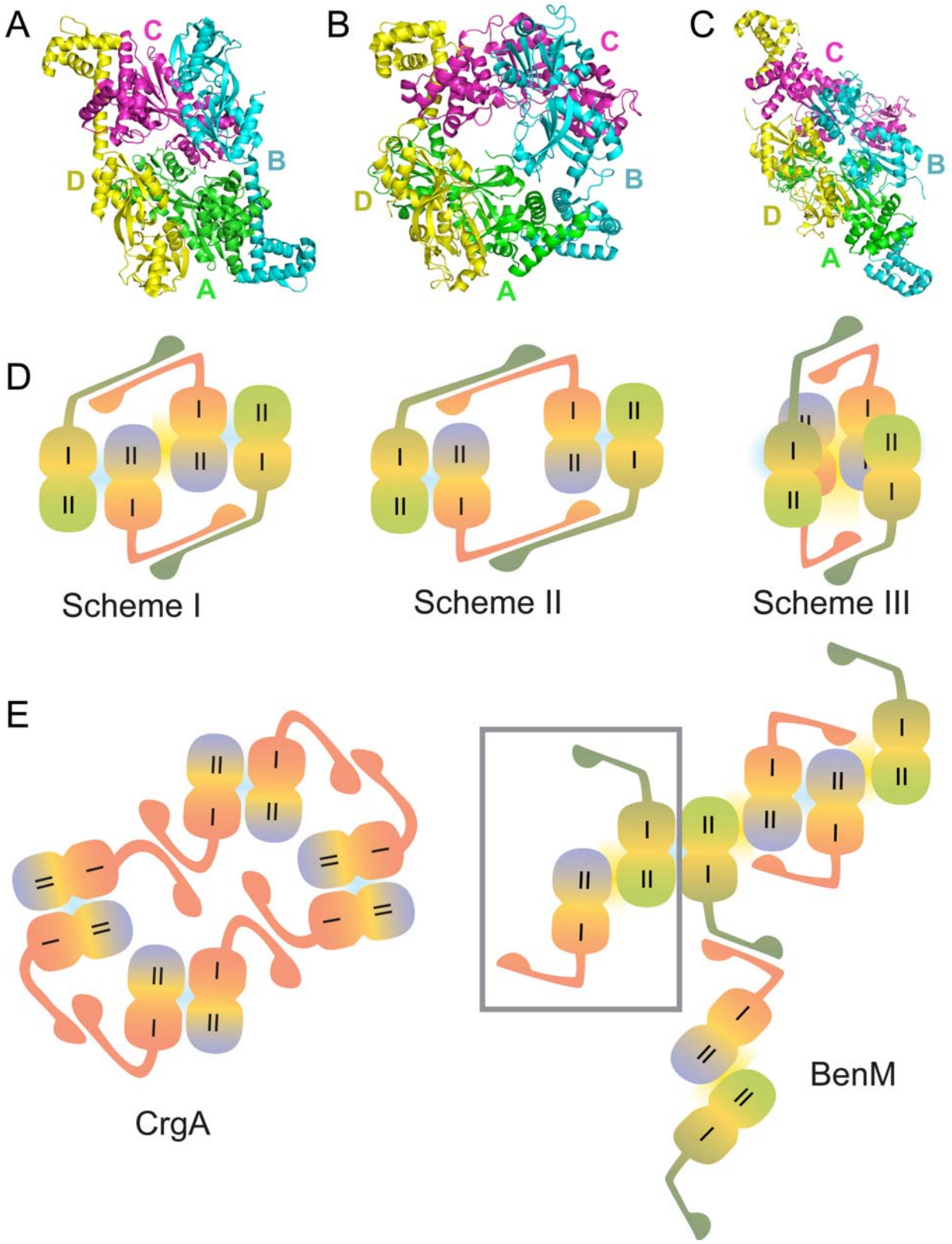
A
C
B
D
A
B
C
B
D
A
C
B
D
A
D
Scheme I
Scheme II
Scheme III
E
CrgA
BenM

Figure 2.5. Three oligomerization schemes used by representative LTTR structures. The CbnR tetramer (panel A) is composed of two dimers (A-D and B-C subunits) interacting through the α10-α10 tetramer interface. The 2ESN tetramer (panel B) is formed by two dimers (A-D and B-C subunits) interacting through their LH/DBDs. The 3FZV tetramer (panel C) forms via a large interacting area of four centrally located EBDs, while the DBDs project outward at the ends. In schematic representations of the oligomerization schemes (panels D and E), each subunit is comprised of two EBD subdomains (labeled I and II) connected by a LH (thin line) to a DBD (small semi-oval). Blue shading between the EBDs of different subunits represents the α6/β2-α11 dimer interface. Yellow shading between dimeric units represents additional contact sites that include the α10-α10 tetramer interface and also the more complex interface observed in the 3FZV structure. CrgA (panel E, left) forms an octomeric structure using a single subunit conformation and an assembly that follows scheme II. BenM (panel E, right) appears to use Scheme I, but by swapping DBDs in the crystal lattice and using altered DBD/LH/EBD relationships, it was unable to create a closed oligomer in the crystal structure.

Comparisons between native and variant forms of BenM

One goal of this investigation was to identify structural differences between LTTRs in active and inactive conformations. Without benzoate or muconate, BenM binds a site overlapping the -10 region of the *benA* promoter, thereby repressing transcription.[5; 13] This inactive conformation of BenM protects a large operator region from DNase I digestion. However, effector-induced changes yield a smaller footprint and cause the DBD regions of two subunits to relocate from a site centered at -12 to one centered at -42 relative to the *benA*

transcriptional start site. In this effector-bound conformation, BenM activates transcription through a mechanism likely to involve direct contact with RNA polymerase.

Investigations of the variant proteins sought to identify features of an activated conformation since BenM(E226K) and BenM(R156H) enhance *benA* transcription without the addition of benzoate or muconate.[6] Yet, all structures of the full-length proteins in these studies were essentially indistinguishable. This result was surprising, because earlier investigations of the isolated EBDs revealed clear structural differences between the regulatory domains of the variant and wild-type proteins.[7] A variant EBD with the R156H replacement, without benzoate or muconate, had a structure more similar to the wild-type BenM-EBD bound to effectors than to the unliganded form of the wild-type protein.[7]

In CatM and BenM, the primary effector-binding site lies between the EBD domains, which draw together as they clamp down on an inducer. Our regulatory model proposes that the additive effect of each subunit contraction results in an altered conformation of the tetramer that enables transcriptional activation. Several factors might explain why none of the full-length BenM structures assumed conformations resembling the benzoate- or muconate-bound EBD structures reported previously, such as PDB IDs 2F78 or 2F7A. One possibility is that DNA helps the full-length protein assume the correct "activated" conformation. BenM binds DNA in the presence or absence of effectors, and DBD-DNA interactions may be important in the proper formation of a tetramer. In the full-length protein studies, the proclivity of the DBDs to become entangled in high-order oligomers may affect EBD positioning. The BenM-EBD conformations are known to vary considerably depending on the crystallization conditions and the ions that localize to the effector-binding site. Unlike the structures of the variant EBD with the R156H replacement, the full-length structures of BenM(E226K)-His and BenM(R156H)-His contained

no obvious anions in their effector-binding sites. Thus, the variant regulators may activate transcription *in vivo* without benzoate or muconate via the assembly of oligomeric structures that were not observed here. Another intriguing possibility that requires further study is that the amino acid replacements alter effector recognition such that atypical anions act as inducers.

LTTR generalizations

Because LTTRs comprise the largest family of transcriptional regulators in bacteria, many efforts center on developing a general model for the molecular basis of their regulation. However, because they have been difficult to characterize structurally, LysR-type proteins are not as well understood as other types of regulators. As more full-length structures are determined, it becomes evident that regulatory models must account for substantial differences as well as common characteristics in LTTR-mediated gene expression. Our studies suggest that no single tetrameric structure can be used exclusively to describe the whole family. To date, the group of well-characterized full-length LTTRs contains fewer than ten different structures, yet three different schemes are needed to describe how they form tetramers. Moreover, tetramers are not the only oligomeric forms of LTTRs that appear to be functionally significant. An improved understanding of transcriptional regulation requires further investigation of protein-DNA-effector interactions that will build on structural studies of full-length LTTRs such as BenM.

Materials and Methods

Chemicals

Reagent grade chemicals and 18 M Ohm cm^{-1} Millipore-filtered water were used. ReagentPlus (99%) imidazole from Aldrich was used for protein purification by metal chelating

column. Ultra grade (≥ 99.5%) imidazole from Fluka was used for protein purification by anion exchange column.

Purification of BenM

Previously described plasmids pBAC433 and pBAC778 were used to express full-length proteins, BenM-His and BenM(E226K)-His respectively, both with C-terminal hexa-histidine tags.[6; 13] Plasmid pBAC933, which encodes BenM(R156H)-His, was constructed from pBAC433 by site-directed mutagenesis with a method (QuikChange, Stratagene) and mutagenic primers (SC1-BenMR156H and SC2-BenMR156H) that were previously described.[6] The sequences of these plasmids were confirmed by DNA sequencing. Plasmids were transformed into *E. coli* BL21(DE3) cells, which were subsequently grown in autoinduction medium as detailed by Studier[34]. Approximately 4-5 g of cell pellets were obtained per 0.5 L of culture. Cell pellets were harvested from cultures and suspended in 40 ml of binding buffer (30 mM Tris, 500 mM NaCl, 30% glycerol, 5 mM imidazole, 10 mM BME, pH 7.9) on ice. Cells were lysed with a chilled (4 °C) French pressure cell at 15,000 psi. The cell lysates were centrifuged at 39,000*g* for 30 min at 4 °C. The clarified supernatants were applied to a HiTrap 5 ml metal chelating column (GE Healthcare Life Sciences) charged with nickel.

Purification was done at room temperature with an ÄKTA (Pharmacia) system. Each protein was eluted with a linear buffer gradient at 2 ml min^{-1} with elution buffer (30 mM Tris, 500 mM NaCl, 30% glycerol, 500 mM imidazole, 10 mM BME, pH 7.9). Purified protein fractions were pooled and passed through a HiTrap 5 ml desalting column (GE Healthcare Life Sciences) equilibrated with Q start buffer (30 mM Tris, 50 mM NaCl, 10% glycerol, 250 mM imidazole, 10 mM BME, pH 9.0). Desalted protein fractions were subsequently applied to a

HiTrap 5 ml Q FF anion exchange column (GE Healthcare Life Sciences) and eluted with Q elution buffer (30 mM Tris, 1 M NaCl, 10% glycerol, 250 mM imidazole, 10 mM BME, pH 9.0). Protein fractions were pooled and stored at 4 °C until use.

The relative molecular mass of BenM-His and variants was determined by gel-exclusion chromatography using a Superdex 200 10/300 GL column (GE Biosciences) in protein buffer with the following calibration standards: blue dextran (2000 kDa) , apoferritin (440 kDa) , amyloglucosidase (97 kDa) , bovine serum albumin (68 kDa) , ovalbumin (43 kDa) , carbonic anhydrase inhibitor II (29 kDa) , trypsin inhibitor (20 kDa), and lentil lectin (14.5 kDa). The apparent relative molecular masses in each case were consistent with tetrameric species of the proteins: BenM-His M_r 132000, BenM(E226K)-His M_r 139000, BenM(R156H)-His M_r 151000.

Crystallization

Initial screens for BenM-His, BenM(E226K)-His, and BenM(R156H)-His crystals were performed in-house by the microbatch-under-oil method using kits from Hampton Research (Crystal Screen, Crystal Screen 2, Index, Natrix™, PEG/Ion™, Grid Screen Ammonium Sulfate, Grid Screen Sodium Malonate, Grid Screen Sodium Chloride, Grid Screen PEG 6000, Grid Screen PEG/LiCl). Each of over 300 conditions was set up by combining 2 µl of protein solution (2 mg ml^{-1}) with 2 µl precipitating solution per well under a layer of paraffin oil. Two crystallization cocktails that generated BenM microcrystals at 295 K are here referred to as solution A (Crystal Screen I solution 19: 35% v/v iso-propanol, 0.1 M Tris HCl pH 8.5, 0.2 M Ammonium acetate) and solution B (Crystal Screen II solution 31: 20% v/v Jeffamine M-600, 0.1 M HEPES, pH 7.5). BenM(E226K)-His only crystallized upon mixing with solution B.

Crystals of BenM-His and BenM(E226K)-His were optimized using the microbatch-under-oil method and the hanging-drop vapor diffusion method. Immediately prior to crystallization, the protein solution was concentrated from a dilute solution (0.3-0.5 mg ml^{-1}) with an Ultrafree centrifuge concentrator (Millipore) to 5 mg ml^{-1} and incubated with equal volume of crystallization solution A or B for 5 min at room temperature. Following incubation, the mixture was centrifuged at 14,000 rpm for 5 min using a benchtop microcentrifuge. The supernatant was incubated in a well under paraffin oil or in a hanging drop setup at 295 K. The optimal reservoir for vapor diffusion experiments contained 18 mM Tris , 36 mM NaCl, 6% glycerol, 150 mM imidazole, 6 mM BME, 8% v/v Jeffamine M-600, and 0.04 M HEPES pH 7.5.

Microbatch-under-oil crystallization screens of BenM(R156H)-His set up using crystallization screen kits (Hampton Research) were initially unsuccessful. Mixing of different precipitating solutions from crystallization screen kits was required to generate BenM(R156H)-His crystals. Small amounts of Crystal Screen I6 (0.2 M $MgCl_2$ hexahydrate, 0.1 M Tris HCl pH 8.5, and 30% w/v PEG 4000) or Crystal Screen II26 (0.1 M MES pH 6.5, 0.2 M Ammonium sulfate, 30% w/v PEG MME 5000) solutions were used as additives in the crystallization of BenM(R156H)-His with solution B. To obtain BenM(R156H)-His crystals, 4 µl protein solution was combined with 5 µl of optimized precipitating solution containing 16% v/v Jeffamine M-600, 0.08 M HEPES pH 7.5, 0.04 M Ammonium sulfate, 6% w/v PEG MME 5000, and 0.02 M MES pH 6.5 or 16% v/v Jeffamine M-600, 0.08 M HEPES pH 7.5, 0.04 M $MgCl_2$ hexahydrate, and 6% w/v PEG 4000. Single crystals were transferred into cryoprotectant solutions containing crystallization growth conditions and 30% glycerol or 30% Jeffamine M-600 and flash frozen in liquid nitrogen.

Freshly purified BenM-His samples (2 mg ml^{-1}) were shipped overnight on ice and submitted to the Center for High-Throughput Structural Biology (CHTSB) at the Hauptman-Woodward Medical Institute for high-throughput crystallization screens using the microbatch under oil method.[21; 22] Over a period of a few years, multiple samples from independently purified protein preparations were subjected to several rounds of crystallization trials with 1536 different cocktails, but no conditions were identified that produced crystals. Initial trials were performed using only metal-chelate purified protein. Later, an additional step of anion-exchange chromatography at pH was used to generate more highly purified proteins.

Data collection, data processing, structure determination, and analysis

Samples were shipped to the Advanced Photon Source, Argonne, IL in a dry-shipping dewar. Datasets were collected under cryogenic conditions (100K) at the SBC-CAT 19-ID beamline fo BenM-His and BenM(R156H)-His) or SER-CAT 22-ID beamline for BenM(E226K)-His using MarCCD detectors. The detector distances, total oscillation range and oscillation widths were optimized to reduce overlaps at high resolution and obtain complete data sets using the HKL3000 prediction routines. Data were processed and scaled with HKL 3000 at the beamline or reprocessed with HKL2000 on home workstations.[35]

The structure determinations were performed by molecular replacement using the programs PHASER and MolREP in the CCP4 suite [36] using PDB ID: 2F97[7] and the unpublished structure of the BenM-DBD (Alanazi, Neidle and Momany, manuscript in preparation). Despite many rounds of modeling using COOT [37]and refinement, the electron density map in the DBD regions was poorly-defined relative to the EBD density (***e.g.*** R_{factor} 0.243, R_{free} 0.321). Moreover, numerous positive density features in difference Fourier maps were observed, indicating the

presence of un-modeled protein components in the structure. These features could be modeled by adding a second DNA binding domain wherein the coordinates were swapped between the two conformations of the BenM molecules. Refinement of the DBD domains with partial occupancies reduced the R_{factor} marginally (R_{factor} 0.222, R_{free} 0.296). Twinning tests (*CTRUNCATE* and *DETWIN* from the *CCP4* suite) suggested that twinning was present. Final refinement of the structures was performed using REFMAC5 [38] version 5.5.0072 with the twinning mode activated and TLS domains identified by the TLS server (http://skuld.bmsc.washington.edu/~tlsmd/)[39] included in the refinement. Cell constants were iteratively improved by varying *a* and *b* incrementally (0.01 Å steps, maximum 0.7 Å change) before REFMAC5 refinement. Cell constants that demonstrated the best geometric statistics (lowest RMSD in bonds and angles) correspondingly had the lowest R_{factor} and R_{free} values relative to the *Scalepack* refined cell constants.

Protein interface parameters (particularly the interface accessible surface areas) for different LTTR contact surfaces were calculated using the PROTORP server (http://www.bioinformatics.sussex.ac.uk/protorp/).[32] Structural superpositions of EBD-I domains from all known full-length LTTR structures were performed via invariant core analysis using the program Bio3d.[27] Theta angles defining the orientation of each linker helix relative to the orientation of CbnR's linker helix were calculated using *HELIXANG* from the *CCP4* suite.[36]

Accession codes

The atomic coordinates (PDB ID codes: 3K1N, 3K1P, and 3K1M) have been deposited in the Protein Data Bank (PDB), Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ.

Acknowledgements

We are grateful to the staff at the Hauptman-Woodward Institute for performing high-throughput crystallization screens. We also thank the staff at the Southeast Regional Collaborative Access team (SER-CAT) 22-ID and Structural Biology Center Collaborative Access Team (SBC-CAT) 19-ID beam lines at the Advanced Photon Source (APS), Argonne National Laboratory for excellent assistance as well as Dr. Lirong Chen and Dr. John Rose for local support. We thank Amer Alanazi for the prepublication use of his structure of BenM-DBD and his assistance in data collection. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, and Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. The research was funded by National Science Foundation grants MCB-0346422 (to C.M.) and MCB-0516914 (to E.L.N.).

References

1. Pareja, E., Pareja-Tobes, P., Manrique, M., Pareja-Tobes, E., Bonal, J. & Tobes, R. (2006). ExtraTrain: a database of Extragenic regions and Transcriptional information in prokaryotic organisms. *BMC Microbiol.* **6**, 29.

2. Henikoff, S., Haughn, G. W., Calvo, J. M. & Wallace, J. C. (1988). A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. USA* **85**, 6602-6606.

3. Schell, M. A. (1993). Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* **47**, 597-626.

4. Craven, S. H., Ezezika, O. C., Momany, C. & Neidle, E. L. (2008). LysR homologs in *Acinetobacter*: Insights into a diverse and prevalent family of transcriptional regulators. In *Acinetobacter Molecular Biology* (Gerischer, U., ed.), pp. 163-202. Caister Academic Press, Norfolk (UK).

5. Collier, L. S., Gaines, G. L. & Neidle, E. L. (1998). Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *J. Bacteriol.* **180**, 2493-2501.

6. Craven, S. H., Ezezika, O. C., Haddad, S., Hall, R. A., Momany, C. & Neidle, E. L. (2009). Inducer responses of BenM, a LysR-type transcriptional regulator from *Acinetobacter baylyi* ADP1. *Mol. Microbiol.* **72**, 881-894.

7. Ezezika, O. C., Haddad, S., Clark, T. J., Neidle, E. L. & Momany, C. (2007). Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. *J. Mol. Biol.* **367**, 616-629.

8. Romero-Arroyo, C. E., Schell, M. A., Gaines, G. L., 3rd & Neidle, E. L. (1995). *catM* encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. *J. Bacteriol.* **177**, 5891-8.

9. Ezezika, O. C., Collier-Hyams, L. S., Dale, H. A., Burk, A. C. & Neidle, E. L. (2006). CatM regulation of the *benABCDE* operon: functional divergence of two LysR-type paralogs in *Acinetobacter baylyi* ADP1. *Appl. Environ. Microbiol.* **72**, 1749-58.

10. Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol. Rev.* **29**, 231-62.

11. Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K. & Senda, T. (2003). Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. *J. Mol. Biol.* **328**, 555-566.

12. Choi, H., Kim, S., Mukhopadhyay, P., Cho, S., Woo, J., Storz, G. & Ryu, S. E. (2001). Structural basis of the redox switch in the OxyR transcription factor. *Cell* **105**, 103-13.

13. Bundy, B. M., Collier, L. S., Hoover, T. R. & Neidle, E. L. (2002). Synergistic transcriptional activation by one regulatory protein in response to two metabolites. *Proc. Natl. Acad. Sci. USA* **99**, 7693.

14. Tyrrell, R., Verschueren, K. H. G., Dodson, E. J., Murshudov, G. N., Addy, C. & Wilkinson, A. J. (1997). The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. *Structure* **5**, 1017-1032.

15. Smirnova, I. A., Dian, C., Leonard, G. A., McSweeney, S., Birse, D. & Brzezinski, P. (2004). Development of a bacterial biosensor for nitrotoluenes: the crystal structure of the transcriptional regulator DntR. *J. Mol. Biol.* **340**, 405-418.

16. Quiocho, F. A. & Ledvina, P. S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* **20**, 17-25.

17. Sainsbury, S., Lane, L. A., Ren, J., Gilbert, R. J., Saunders, N. J., Robinson, C. V., Stuart, D. I. & Owens, R. J. (2009). The structure of CrgA from *Neisseria meningitidis* reveals a new octameric assembly state for LysR transcriptional regulators. *Nucl. Acids Res.*

18. Monferrer, D., Tralau, T., Kertesz, M. A., Dix, I., Sola, M. & Uson, I. (2010). Structural studies on the full-length LysR-type regulator TsaR from *Comamonas testosteroni* T-2 reveal a novel open conformation of the tetrameric LTTR fold. *Mol.. Microbiol.*

19. Zhou, X., Lou, Z., Fu, S., Yang, A., Shen, H., Li, Z., Feng, Y., Bartlam, M., Wang, H. & Rao, Z. (2010). Crystal structure of ArgP from *Mycobacterium tuberculosis* confirms two distinct conformations of full-length LysR transcriptional regulators and reveals its function in DNA binding and transcriptional regulation. *J. Mol. Biol.* **396**, 1012-24.

20. Ireton, G. C. & Stoddard, B. L. (2004). Microseed matrix screening to improve crystals of yeast cytosine deaminase. *Acta crystallographica* **60**, 601-5.

21. Chayen, N. E. (1997). The role of oil in macromolecular crystallization. *Structure* **5**, 1269-74.

22. Luft, J. R., Wofley, J., Jusrisica, I., Glasgow, J., Fortier, S. & DeTitta, G. T. (2001). Macromolecular crystallization in a high throughput laboratory- the search phase. *J. Cryst. Growth* **232**, 591-595.

23. Clark, T., Haddad, S., Neidle, E. & Momany, C. (2003). Crystallization of the effector-binding domains of BenM and CatM, LysR-type transcriptional regulators from Acinetobacter sp. ADP1. *Acta Crystallog. sect. D* **60**, 105-108.

24. McPherson, A. (1976). Crystallization of proteins from polyethylene glycol. *J. Biol. Chem.* **251**, 6300-6303.

25. Matthews, B. W. (1968). Solvent content of protein crystals. *J Mol Biol* **33**, 491-7.

26. Ezezika, O. C., Haddad, S., Neidle, E. L. & Momany, C. (2007). Oligomerization of BenM, a LysR-type transcriptional regulator: structural basis for the aggregation of proteins in this family. *Acta Crystallog. sect. F* **63**, 361-368.

27. Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A. & Caves, L. S. D. (2006). Bio 3 d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695.

28. Sainsbury, S., Ren, J., Saunders, N. J., Stuart, D. I. & Owens, R. J. (2008). Crystallization and preliminary X-ray analysis of CrgA, a LysR-type transcriptional regulator from pathogenic *Neisseria meningitidis* MC58. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **64**, 797-801.

29. Muraoka, S., Okumura, R., Uragami, Y., Nonaka, T., Ogawa, N., Miyashita, K. & Senda, T. (2003). Purification and crystallization of a LysR-type transcriptional regulator CBNR from Ralstonia eutropha NH9. *Protein and peptide letters* **10**, 325.

30. Holm, L. & Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends in biochemical sciences* **20**, 478-80.

31. Maddocks, S. E. & Oyston, P. C. F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**, 3609.

32. Reynolds, C., Damerell, D. & Jones, S. (2009). ProtorP: a protein-protein interaction analysis server. *Bioinformatics* **25**, 413-4.

33. Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97.

34. Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif* **41**, 207–234.

35. Otwinowski, Z. & Minor, W. (1997). Processing of X-Ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* **276**, 307-26.

36. (1994). CCP4, Collaborative Computational Project, Number 4. *Acta Crystallogr* **D50**, 760-763.

37. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta crystallographica* **60**, 2126-32.

38. Lamzin, V. S. & Wilson, K. S. (1993). Automated refinement of protein models. *Acta Crystallogr.* **D49**, 129-147.

39. Painter, J. & Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta crystallographica* **62**, 439-50.

Supplementary materials

Figure 2.6. Arrangements of residues and electron density at the sites of BenM-His amino acid replacements. The structures of BenM(E226K)-His (green, panel A) and BenM(R156H)-His (blue, panel B) are compared with the native BenM-His (orange) structure near the position of the site-directed alteration. The features of the electron density match the shapes and sizes of corresponding side chains in all the structures. The only significant difference observed among these structures is the unique orientation of residue Asp151 in chain A of the native BenM-His structure. Residues 151 in chain B of native BenM-His and in both chains of E226K all assume similar orientations as that of BenM(R156H)-His.

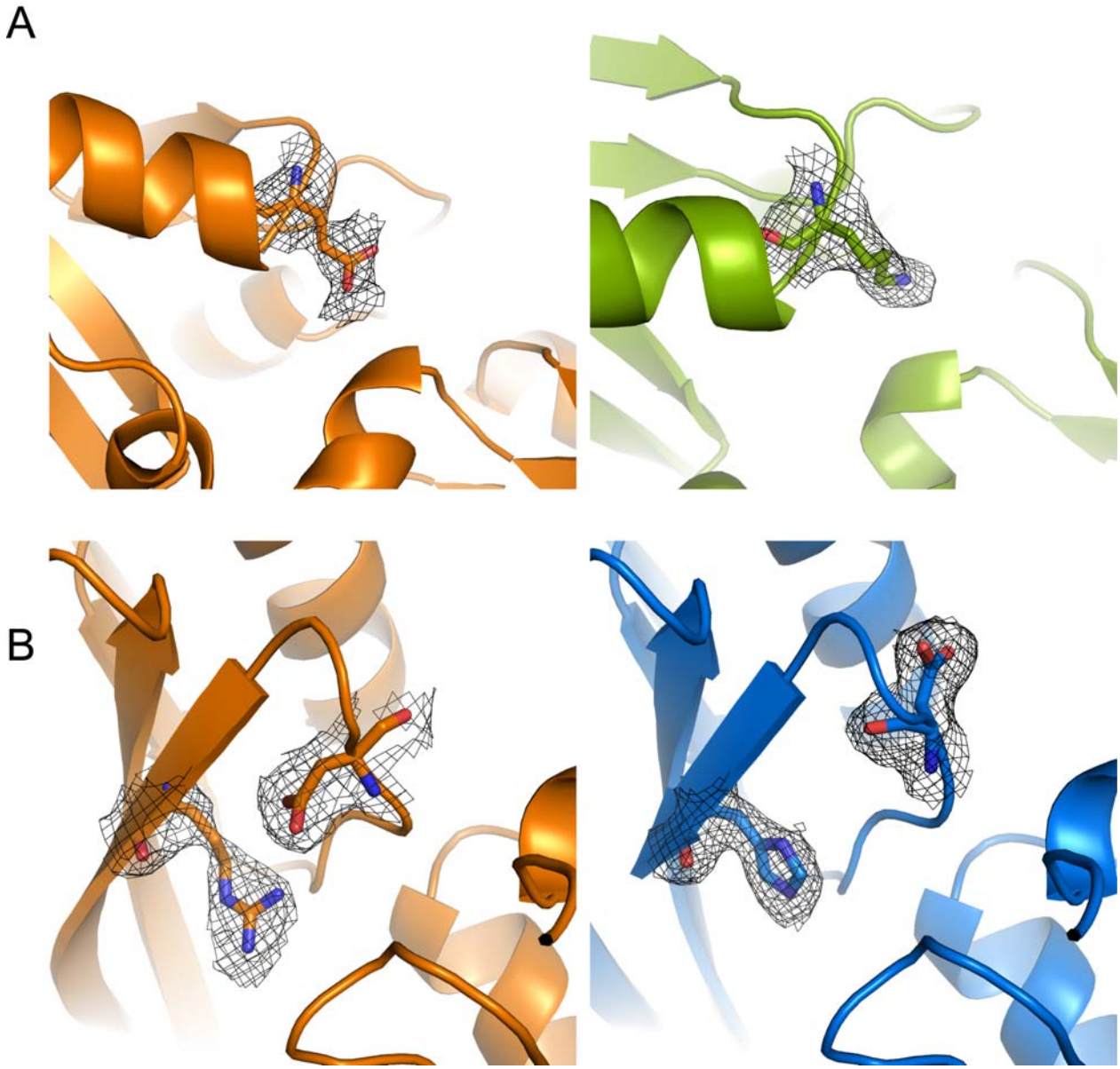
A
B

Table 2.2. Dimer interfaces

	CbnR (AQ/BP)	DntR (AB)	BenM (AA'/BB') [1]
Interface Residue Segments	11/11	7	11/13
Interface Accessible Surface Area (Å^2)	1241.29/1292.12	1531.93	1321.10/1325.73
% Interface Accessible Surface Area	12.47/12.62	14.03	12.36/12.33
Atoms in Interface	108/105	132	114/115
% Polar Atoms Contribution to Interface	26.64/26.86	26.39	29.13/28.51
% Non-Polar Atoms Contribution to Interface	55.78/57.72	57.07	47.82/47.91
% Neutral Atoms Contribution to Interface	16.89/14.68	15.66	22.27/22.77
Residues in Interface	38/37	40	33/34
% Polar Residues in Interface	26.32/27.03	17.50	24.24/26.47
% Non-Polar Residues in Interface	47.37/48.65	62.50	57.58/55.88
% Charged Residues in Interface	26.32/24.32	20.00	18.18/17.65
Residues on Surface	290/276	308	306/316
% Polar Residues on Surface	21.72/21.38	28.90	32.03/31.65
% Non-Polar Residues on Surface	49.66/48.19	43.18	40.52/41.77
% Charged Residues on Surface	28.62/30.43	27.92	27.45/26.58
Planarity (Å)	2.339/2.500	2.520	2.451/2.461
Eccentricity	0.826/0.824	0.911	0.859/0.859
Secondary Structure in Interface	Alpha-Beta/Alpha	Alpha-Beta	Alpha/Alpha-Beta
% Alpha Character in Interface	47.37/48.65	30.00	45.45/23.53
% Beta Character in Interface	21.05/18.92	25.00	6.06/23.53
Secondary Structure on Surface	Alpha-Beta/Alpha-Beta	Alpha/Beta	Alpha-Beta/Alpha-Beta
% Alpha Character on Surface	27.93/28.62	31.49	24.84/23.42
% Beta Character on Surface	31.03/31.16	25.97	26.80/27.85
Hydrogen Bonds	5/11	13	14/14
Salt Bridges	33/23	9	54/42
Disulphide Bonds	0/0	0	0/0
Bridging Water Molecules	0/0	0	0/0
Gap Volume (Å^3)	5902.88/6172.88	5821.88	6375.38/6081.75
Gap Volume Index (Å)	2.38/2.39	1.90	2.41/2.29
Δ^iG (kcal/mol)	-15.9/-15.9	-19.4	-14.5/-10.3

[1]A' and B' molecules of BenM were generated by applying symmetry operations (x,-y-1,-z and -x,y,-z-1/2) to the original A and B subunits of BenM, respectively.

	2ESN (AC/BD)	ArgP (AB)	TsaR (BA/BA')[2]	3FZV (AC/BD)
Interface Residue Segments	10/8	9	10/10	7/7
Interface Accessible Surface Area (Å^2)	1756.34/1758.13	1179.66	1395.71/1395.72	1159.33/1165.54
% Interface Accessible Surface Area	15.38/15.85	7.80	13.39/13.34	10.93/10.68
Atoms in Interface	158/155	101	122/122	98/97
% Polar Atoms Contribution to Interface	25.53/28.21	22.87	23.23/23.23	31.98/32.08
% Non-Polar Atoms Contribution to Interface	46.86/46.62	61.05	63.47/63.47	52.79/52.54
% Neutral Atoms Contribution to Interface	26.95/24.55	14.93	12.69/12.69	14.76/14.84
Residues in Interface	45/47	34	38/38	31/31
% Polar Residues in Interface	24.44/23.40	17.65	18.42/18.42	35.48/35.48
% Non-Polar Residues in Interface	55.56/55.32	58.82	65.79/65.79	45.16/45.16
% Charged Residues in Interface	20.00/21.28	23.53	15.79/15.79	19.35/19.35
Residues on Surface	303/306	423	299/299	331/334
% Polar Residues on Surface	22.77/23.20	24.59	30.43/30.43	27.19/26.95
% Non-Polar Residues on Surface	48.18/47.71	48.46	44.8244.82/	46.83/47.01
% Charged Residues on Surface	29.04/29.08	26.95	24.75/24.75	25.98/26.05
Planarity (Å)	2.544/2.623	2.711	3.149/3.149	3.253/3.263
Eccentricity	0.858/0.835	0.941	0.812/0.812	0.801/0.796
Secondary Structure in Interface	Alpha-Beta/Alpha	Alpha-Beta	Alpha/ Alpha	Beta/Beta
% Alpha Character in Interface	46.67/48.94	35.29	47.37/47.37	19.35/19.35
% Beta Character in Interface	22.22/19.15	23.53	7.89/7.89	45.16/45.16
Secondary Structure on Surface	Alpha-Beta/Alpha-Beta	Alpha-Beta	Alpha-Beta/Alpha-Beta	Alpha/Alpha
% Alpha Character on Surface	29.04/28.76	37.59	25.75/26.42	25.98/26.05
% Beta Character on Surface	29.04/30.39	21.75	30.43/29.10	18.43/19.16
Hydrogen Bonds	19/17	17	16/16	8/9
Salt Bridges	12/17	16	11/11	18/22
Disulphide Bonds	0/0	0	0/0	0/0
Bridging Water Molecules	6/6	0	0/0	0/0
Gap Volume (Å^3)	6453.00/6567.75	9406.12	6942.38/6993.00	6233.62/6196.50
Gap Volume Index (Å)	1.84/1.87	3.99	2.49/2.51	2.69/2.66
Δ^iG (kcal/mol)	-10.8/-13.0	-13.4	-23.4	-25.3/-25.6

[2]A' molecule of TsaR was generated by applying symmetry operation (-x,y,-z+1) to the original A subunit of TsaR.

	OxyR, oxidized (AA') [3]	OxyR, reduced (AB)	CynR, 2HXR (AB)	3FD3 (AA') [4]
Interface Residue Segments	7	12	7	9
Interface Accessible Surface Area (Å^2)	1130.50	1067.13	1217.33	1130.67
% Interface Accessible Surface Area	10.91	10.37	12.09	10.95
Atoms in Interface	95	101	108	98
% Polar Atoms Contribution to Interface	33.37	34.02	33.29	28.25
% Non-Polar Atoms Contribution to Interface	50.01	52.99	46.77	50.52
% Neutral Atoms Contribution to Interface	15.98	12.10	19.35	20.45
Residues in Interface	31	29	35	35
% Polar Residues in Interface	41.94	34.48	37.14	34.29
% Non-Polar Residues in Interface	38.71	48.28	45.71	48.57
% Charged Residues in Interface	19.35	17.24	17.14	17.14
Residues on Surface	302	289	294	304
% Polar Residues on Surface	22.52	24.57	32.31	26.32
% Non-Polar Residues on Surface	45.70	44.98	40.82	50.66
% Charged Residues on Surface	31.79	30.45	26.87	23.03
Planarity (Å)	2.236	1.735	3.044	2.503
Eccentricity	0.772	0.905	0.911	0.950
Secondary Structure in Interface	Beta	Alpha	Alpha/Beta	Alpha
% Alpha Character in Interface	12.90	37.93	48.57	34.29
% Beta Character in Interface	38.71	3.45	20.00	14.29
Secondary Structure on Surface	Alpha/Beta	Coil	Alpha/Beta	Alpha/Beta
% Alpha Character on Surface	21.85	11.07	33.67	32.24
% Beta Character on Surface	29.14	12.80	33.67	25.66
Hydrogen Bonds	28	5	12	108
Salt Bridges	32	3	2	2
Disulphide Bonds	0	0	0	0
Bridging Water Molecules	0	0	8	0
Gap Volume (Å^3)	6257.25	8295.75	5231.25	8116.88
Gap Volume Index (Å)	2.77	3.89	2.15	3.59
$\Delta^i G$ (kcal/mol)	-10.4	-14.8	-14.6	-13.5

[3] A' molecule of oxidized OxyR was generated by applying symmetry operation (-x,y,-z) to the original A subunit of oxidized OxyR.
[4] A' molecule of 3FD3 was generated by applying symmetry operation (y,x,-z) to the original A subunit of 3FD3.

	2QL3 (AB/CD/EF/GH/IJ/KL)	3HHG (AB/CD/EF/GH)
Interface Residue Segments	8/8/7/7/8/8	9/7/6/8
Interface Accessible Surface Area ($Å^2$)	1455.41/1450.14/1388.27/1424.01/1441.65/1436.44	1235.71/1218.48/1300.79/1188.07
% Interface Accessible Surface Area	13.78/13.51/13.01/13.22/13.28/13.46	11.88/11.49/11.81/10.98
Atoms in Interface	129/123/116/123/128/124	105/98/101/102
% Polar Atoms Contribution to Interface	32.94/33.78/32.62/33.30/31.56/37.10	29.12/29.23/26.63/31.38
% Non-Polar Atoms Contribution to Interface	54.96/56.00/56.28/56.76/56.01/54.70	53.56/52.96/52.84/51.96
% Neutral Atoms Contribution to Interface	11.30/9.47/10.27/9.21/11.72/7.48	16.58/17.32/19.47/15.80
Residues in Interface	39/36/36/36/37/36	36/37/35/38
% Polar Residues in Interface	25.64/27.78/27.78/27.78/27.03/27.78	38.89/35.14/37.14/36.84
% Non-Polar Residues in Interface	51.28/52.78/52.78/52.78/51.35/52.78	50.00/54.05/51.43/50.00
% Charged Residues in Interface	23.08/19.44/19.44/19.44/21.62/19.44	11.11/10.81/11.43/13.16
Residues on Surface	305/306/310/308/302/308	297/300/301/305
% Polar Residues on Surface	22.62/22.22/22.58/22.08/22.19/22.73	31.99/30.67/32.23/31.15
% Non-Polar Residues on Surface	52.13/52.29/52.26/52.60/51.99/51.62	41.08/42.33/40.53/42.62
% Charged Residues on Surface	25.25/25.49/25.16/25.32/25.83/25.65	26.94/27.00/27.24/26.23
Planarity (Å)	2.989/3.025/2.969/3.066/3.015/2.990	1.757/1.690/2.002/1.729
Eccentricity	0.859/0.875/0.874/0.864/0.857/0.865	0.930/0.934/0.937/0.911
Secondary Structure in Interface	Alpha-Beta/Alpha-Beta/Alpha-Beta/Alpha-Beta/ Alpha-Beta/ Alpha-Beta	Alpha-Beta/Alpha-Beta/Alpha-Beta/Alpha-Beta
% Alpha Character in Interface	48.72/52.78/52.78/52.78/51.35/52.78	38.89/ 35.14/37.14/36.84
% Beta Character in Interface	23.08/22.22/22.22/22.22/24.32/22.22	25.00/ 27.03/25.71/26.32
Secondary Structure on Surface	Alpha-Beta/Alpha-Beta/Alpha-Beta/Alpha-Beta/ Alpha-Beta/ Alpha-Beta	Alpha-Beta/Alpha-Beta/Alpha-Beta/Alpha-Beta
% Alpha Character on Surface	25.25/26.14/26.45/26.30/26.82/25.97	28.62/ 28.00/29.24/28.20
% Beta Character on Surface	29.18/30.07/29.68/29.22/29.47/29.87	28.96/ 29.67/27.91//29.84
Hydrogen Bonds	11/11/10/9/10/11	10/13/13/12
Salt Bridges	22/15/11/35/28/41	3/1/7/2
Disulphide Bonds	0/0/0/0/0/0	0/0/0/0
Bridging Water Molecules	11/8/13/11/17/15	0/0/0/0
Gap Volume ($Å^3$)	6864.75/7610.62/7563.38/7803.00/7414.88/7678.12	8623.12/7998.75/8697.38/8197.88
Gap Volume Index (Å)	2.36/2.62/2.72/2.74/2.57/2.67	3.49/3.28/3.34/3.45
Δ^iG (kcal/mol)	-21.5/-20.6/-20.8/-20.9/-22.0/-20.8	-21.3/-19.5/-18.8/-19.2

Table 2.3. Tetramer interfaces

	CbnR	DntR	BenM	OxyR, reduced	3FZV (AB/CD)
Interface Residue Segments	6	9	9	8	8/7
Interface Accessible Surface Area (Å^2)	845.68	812.29	866.73	650.28	795.50/955.41
% Interface Accessible Surface Area	8.49	7.46	8.11	6.32	7.50/8.75
Atoms in Interface	85	75	81	61	64/81
% Polar Atoms Contribution to Interface	36.72	44.68	45.86	36.66	39.40/37.52
% Non-Polar Atoms Contribution to Interface	35.29	35.82	27.56	41.75	34.97/34.69
% Neutral Atoms Contribution to Interface	26.93	19.12	25.87	20.85	24.40/26.87
Residues in Interface	23	23	24	20	25/26
% Polar Residues in Interface	8.70	8.70	37.50	15.00	52.00/46.15
% Non-Polar Residues in Interface	52.17	56.52	33.33	45.00	28.00/30.77
% Charged Residues in Interface	39.13	34.78	29.17	40.00	20.00/23.08
Residues on Surface	295	320	325	303	343/323
% Polar Residues on Surface	21.36	27.50	30.46	24.75	27.41/29.10
% Non-Polar Residues on Surface	51.86	45.00	44.31	46.20	46.94/45.51
% Charged Residues on Surface	26.78	27.50	25.23	29.04	25.66/25.39
Planarity (Å)	2.430	2.409	2.807	2.247	1.288/1.251
Eccentricity	0.814	0.769	0.946	0.883	0.908/0.915
Secondary Structure in Interface	Alpha/Beta	Alpha/Beta	Beta	Beta	Alpha-Beta/Alpha-Beta
% Alpha Character in Interface	30.43	34.78	8.33	5.00	24.00/23.08
% Beta Character in Interface	26.09	21.74	33.33	35.00	32.00/26.92
Secondary Structure on Surface	Alpha/Beta	Alpha/Beta	Alpha/Beta	Coil	Alpha/Alpha-Beta
% Alpha Character on Surface	29.83	29.38	26.15	11.88	25.95/24.15
% Beta Character on Surface	29.83	28.12	24.92	12.21	19.83/20.12
Hydrogen Bonds	3	8	12	6	6/11
Salt Bridges	33	28	12	30	18/23
Disulphide Bonds	0	0	0	0	0/0
Bridging Water Molecules	0	0	0	0	0/0
Gap Volume (Å^3)	5379.75	5575.50	6949.12	8066.25	4397.62/3719.25
Gap Volume Index (Å)	3.18	3.43	4.01	6.20	2.76/1.95
Δ^iG (kcal/mol)	-1.8	-3.6	-2.1	-3.9	-2.5/-4.3

Table 2.4. DBD/LH-DBD/LH interfaces

	CbnR (AB/PQ)	BenM (AB)	2ESN (AD/BC)	ArgP (AB)	TsaR (AB)
Interface Residue Segments	5/5	5	7	5	5
Interface Accessible Surface Area (Å^2)	973.88/1006.61	1197.34	1481.74/1524.87	933.96	1238.56
% Interface Accessible Surface Area	16.50/17.00	18.44	23.90/24.69	16.83	20.24
Atoms in Interface	80/90	98	116/120	69	97
% Polar Atoms Contribution to Interface	29.12/26.02	24.19	20.09/16.26	22.82	30.49
% Non-Polar Atoms Contribution to Interface	58.14/59.40	65.92	66.09/68.11	70.03	53.66
% Neutral Atoms Contribution to Interface	11.72/13.80	9.25	12.86/14.87	6.31	14.94
Residues in Interface	25/27	28	32/34	25	28
% Polar Residues in Interface	8.00/7.41	17.86	9.38/11.76	16.00	21.43
% Non-Polar Residues in Interface	68.00/62.96	57.14	65.62/64.71	68.00	53.57
% Charged Residues in Interface	24.00/29.63	25.00	25.00/23.53	16.00	25.00
Residues on Surface	141/136	140	134/134	134	142
% Polar Residues on Surface	17.73/18.38	34.29	26.87/28.36	25.37	28.87
% Non-Polar Residues on Surface	49.65/47.06	34.29	43.28/41.79	50.75	42.96
% Charged Residues on Surface	32.62/34.56	31.43	29.85/29.85	23.88	28.17
Planarity (Å)	2.227/2.653	2.833	4.577/4.935	3.090	2.943
Eccentricity	0.943/0.928	0.887	0.841/0.784	0.869	0.920
Secondary Structure in Interface	Alpha/Alpha	Alpha	Alpha/Alpha	Alpha	Alpha
% Alpha Character in Interface	76.00/77.78	78.57	46.88/47.06	68.00	82.14
% Beta Character in Interface	0.00/0.00	0.00	3.12/2.94	4.00	3.57
Secondary Structure on Surface	Alpha/Alpha	Alpha	Alpha/Alpha	Alpha	Alpha
% Alpha Character on Surface	61.70/61.03	66.43	58.21/58.21	57.46	70.42
% Beta Character on Surface	11.35/11.03	5.00	7.46/7.46	5.97	6.34
Hydrogen Bonds	6/3	4	8/9	2	6
Salt Bridges	16/18	13	13/21	9	26
Disulphide Bonds	0/0	0	0/0	0	0
Bridging Water Molecules	0/0	0	0/0	0	0
Gap Volume (Å^3)	2764.12/2365.88	2264.62	3358.12/3337.88	3736.12	2521.12
Gap Volume Index (Å)	1.42/1.18	0.95	1.13/1.09	2.00	1.02
Δ^1G (kcal/mol)	-18.7/-18.0	-23.7	-21.3/-25.8	-21.1	-24.9

	3FZV (AD/BC)	CrgA (AH/BC/DE/FG)
Interface Residue Segments	4/5	5/5/5/5
Interface Accessible Surface Area (Å^2)	504.26/819.88	1265.88/1185.74/1244.26/1250.48
% Interface Accessible Surface Area	10.45/16.73	19.77/18.66/18.25/19.97
Atoms in Interface	48/72	106/98/106/104
% Polar Atoms Contribution to Interface	27.91/29.92	37.83/38.98/35.98/39.03
% Non-Polar Atoms Contribution to Interface	63.28/63.67	53.51/46.95/48.31/52.81
% Neutral Atoms Contribution to Interface	8.08/6.01	7.57/13.28/14.85/7.25
Residues in Interface	17/20	29/27/30/30
% Polar Residues in Interface	17.65/20.00	24.14/25.93/23.33/26.67
% Non-Polar Residues in Interface	58.82/55.00	44.83/40.74/46.67/46.67
% Charged Residues in Interface	23.53/25.00	31.03/33.33/30.00/26.67
Residues on Surface	129/133	139/135/145/139
% Polar Residues on Surface	31.78/34.59	33.81/34.81/33.10/34.53
% Non-Polar Residues on Surface	43.41/40.60	33.81/31.85/35.86/33.81
% Charged Residues on Surface	24.81/24.81	32.37/33.33/31.03/31.65
Planarity (Å)	2.328/2.706	2.553/2.565/2.764/3.009
Eccentricity	0.920/0.905	0.915/0.923/0.907/0.922
Secondary Structure in Interface	Alpha/Alpha	Alpha/Alpha/Alpha/Alpha
% Alpha Character in Interface	58.82/65.00	68.97/66.67/66.67/66.67
% Beta Character in Interface	0.00/0.00	0.00/0.00/3.33/3.33
Secondary Structure on Surface	Alpha/Alpha	Alpha/Alpha/Alpha/Alpha
% Alpha Character on Surface	59.69/57.89	63.31/62.22/64.14/63.31
% Beta Character on Surface	0.00/0.00	4.32/4.44/4.14/3.60
Hydrogen Bonds	1/3	10/11/8/8
Salt Bridges	1/2	38/30/29/39
Disulphide Bonds	0/0	0/0/0/0
Bridging Water Molecules	0/0	0/0/0/0
Gap Volume (Å^3)	3169.12/2983.50	4434.75/4185.00/5123.25/4039.88
Gap Volume Index (Å)	3.14/1.82	1.75/1.76/2.06/1.62
Δ^iG (kcal/mol)	-11.3/-13.4	-17.8/-11.8/-13.7/-18.5

CHAPTER 3

INVESTIGATION OF EFFECTOR BINDING AT TWO SEPARATE BINDING SITES OF BENM AND BENM VARIANTS BY FLUORESCENCE SPECTROSCOPY

[1]Ruangprasert, A., S.H. Craven, E.L. Neidle, R.S. Phillips, and C.A. Momany. To be submitted to *Journal of Bacteriology*

Introduction

BenM, a LysR-type transcriptional regulator (LTTR) in the soil bacterium *Acinetobacter baylyi*, plays a major role in controlling the expression of *ben* and *cat* genes, which encode the enzymes required for aromatic compound degradation through β-ketoadipate pathway [1; 2; 3]. Through the catechol branch of this pathway, benzoate is converted into catechol by the *ben*-encoded enzymes, and catechol is further converted into *cis*,*cis*-muconate (hereafter referred to as muconate) by CatA, which is subsequently degraded to tricarboxylic acid cycle intermediates, succinyl-coA and acetyl-coA, by the *cat*-encoded enzymes (Figure 3.1) [1]. Both benzoate and muconate can individually act on BenM to activate *ben* gene transcription while muconate induces *ben* gene expression more effectively as a sole effector[1]. Further, benzoate and muconate can together generate a synergistic effect on the transcriptional activation of the *ben* genes by BenM [4]. The structures determined from crystals of effector binding domain (EBD) of BenM soaked with effectors clearly revealed two distinct binding sites for inducer molecules (Figure 3.2) [5]. The primary effector binding site, located between subdomains I and II of the C-terminal EBD of the regulator, was found to be able to accommodate either benzoate or muconate. Residues in the primary binding site make several specific contacts with muconate while upon benzoate binding at the site, only a subset of these direct interactions are present [5]. The presence of this site is consistent with other family members that have been shown to utilize comparable regions to interact with their ligands [6; 7; 8; 9; 10; 11].Mutational analyses have mapped this region as involved in inducer recognition [12; 13]. Moreover, other LTTR crystal structures had exogenous ions bound at this site [9; 11; 14]. Binding of cognate effectors to the primary site was suggested to induce conformational changes of the LTTRs that allow transcriptional activation of the target genes [8; 15; 16; 17]. This idea is also supported by the investigation of the muconate-bound structure

of BenM EBD. Binding of muconate to the primary binding site of BenM EBD causes noticeable conformational changes of the protein that involve movement of the domains I and II together several Ångstroms when compared with the unliganded structure of BenM EBD [5].

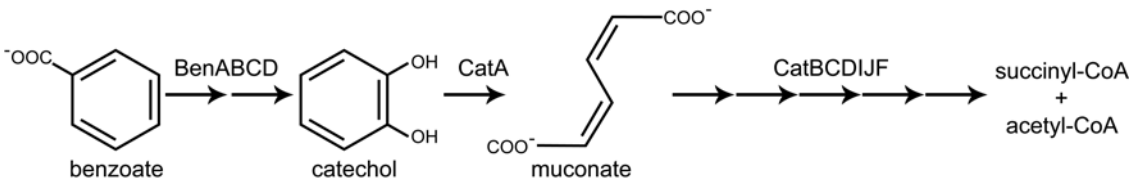

Figure 3.1. Benzoate degradation through the β-ketoadipate pathway involving *ben*- and *cat*-encoded enzymes.

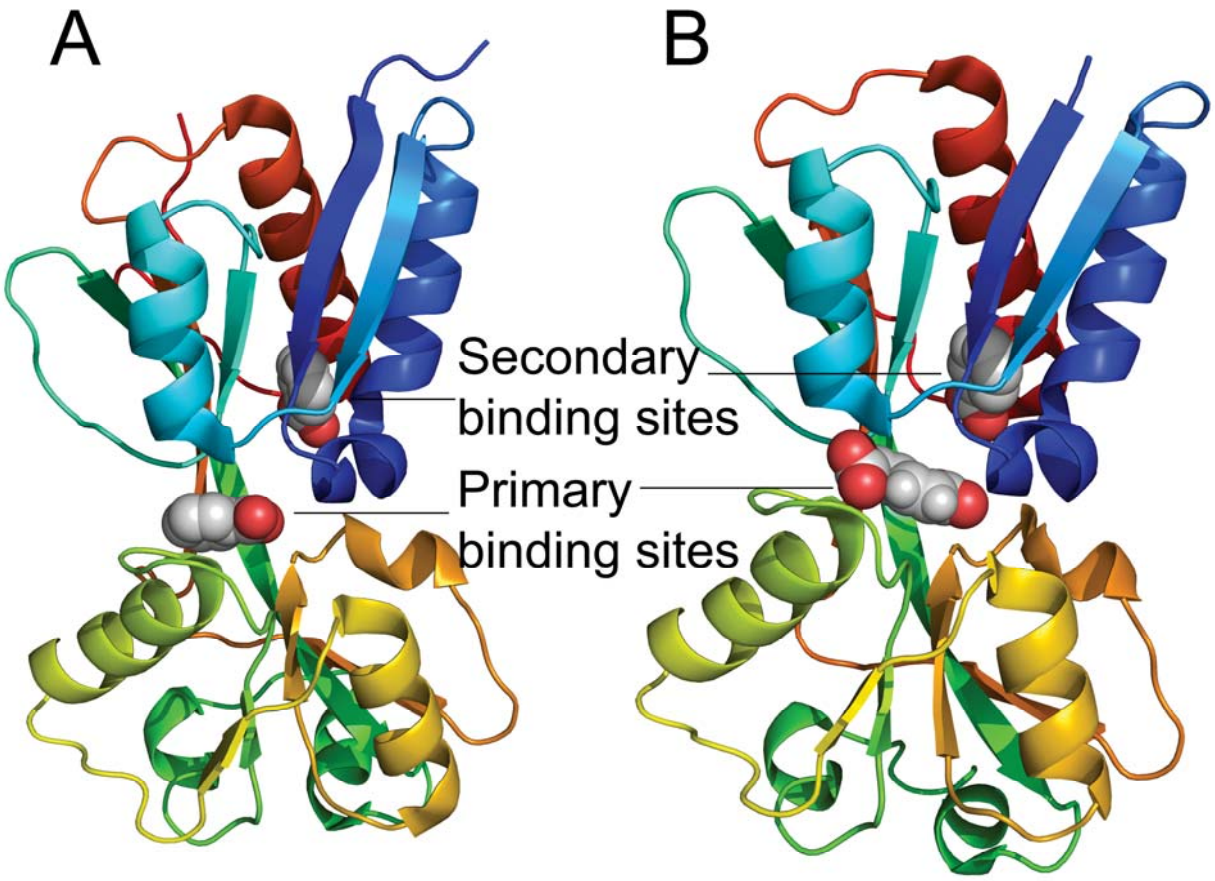

Figure 3.2. Ribbon representations of BenM-EBD with its natural effectors. Positions of the two separate binding sites are indicated. (A) Structure of BenM EBD in complex with benzoate (PDB ID 2F78). Two benzoate molecules were found bound to both primary and secondary binding site of BenM. (B) Structure of BenM-EBD (A subunit) in complex with muconate (PDB ID 2F7A). Muconate molecule was bound to the primary binding site. Benzoate ligand was modeled in the secondary site of the crystal structure. The B subunit did not have significant electron density features in the second site.

Nevertheless, the discovery of second binding site in the structure of BenM-EBD was unforeseen. A benzoate molecule was found enclosed in a hydrophobic pocket adjacent to the primary site in both the crystal structures of BenM EBD soaked with muconate and with benzoate [5]. The benzoate molecule present in this secondary site of the muconate-soaked BenM structure, which had not been intentionally introduced into the crystallization solution implied that the affinity of benzoate for this site should be relatively high [5]. At the second site, specific interactions of BenM with benzoate molecule are provided by two critical residues, Arg160 and Tyr293 [5; 14]. The side chains of these two key residues interact with the carboxyl group of benzoate by forming a salt-bridge and hydrogen bonds to the effector. In addition, several hydrophobic residues lining the binding pocket participate in extensive hydrophobic interactions with the aromatic ring of benzoate. CatM, a paralog of BenM, which has a significantly conserved hydrophobic pocket at the equivalent site but possesses His160 and Phe293 instead of Arg and Tyr, respectively, lacks the ability to respond to benzoate [5]. The requirements of Arg160 and Tyr293 in response of BenM to benzoate were demonstrated by mutational analyses, and the results suggested the importance of both residues in benzoate-dependent *ben* gene expression by BenM[14]. BenM(R160H) and BenM(Y293F) activate transcription of *benA* in response to muconate, but do not respond to benzoate and lose the synergistic transcriptional activation effect possibly because of impaired secondary binding sites [14]. However, further structural and biophysical characterizations of ligand binding at the secondary site of BenM and its variants remain largely unexplored.

Measurement of intrinsic protein fluorescence has been used as a method to observe binding interactions in several proteins[18]. Intrinsic protein fluorescence originates from three aromatic amino acids, tryptophan, tyrosine, and phenylalanine [18]. Tryptophan, which contributes

most to the fluorescence emission of proteins, is highly sensitive to its environment [18]. Conformational changes, binding of ligands, subunit association or protein denaturation can trigger changes in tryptophan environments and, consequently, induce alterations in fluorescence spectra of proteins[18]. In the determination of binding affinity by fluorescence spectroscopy, the extent of quenching of intrinsic protein fluorescence caused by addition of ligand is suggested to be proportional to the concentrations of liganded protein allowing the observation of binding interactions and the calculation of binding constants [18; 19].

BenM possesses a single tryptophan residue (Trp304) at the C-terminal end of the protein. Previous investigations showed that addition of benzoate or muconate induced changes in local environment of Trp304 leading to quenching of intrinsic fluorescence signal of BenM and a red-shift[21]. The estimated dissociation constants of full-length BenM for benzoate and muconate are 1.2 mM and 0.28 mM, respectively[21]. Moreover, benzoate at the concentration of 1 mM appears to compete with muconate for binding site resulting in an apparent increase of the dissociation constant of muconate to 0.49 mM despite the *in vivo* synergistic effect of the two compound on transcriptional activation[21]. However, the binding constants previously estimated predated the availability of the crystal structures of BenM with its effectors, and, thus, were only calculated assuming a single binding site.

DntR is an LTTR family member whose dissociation constants (K_d) for its inducer have been estimated using tryptophan fluorescence spectroscopy [9; 19]. The dissociation constant of DntR for its physiological inducer, sodium salicylate, is approximately 4-5 μM [9; 19]. Cbl, another LTTR whose ligand binding affinity was also determined by fluorescence emission spectroscopy, binds to its effector, adenosine phosphosulphate (APS) with the dissociation constants of 40.4 and 28.3 μM for full-length protein and Cbl-EBD, respectively [20].

Here, we utilized tryptophan fluorescence emission spectroscopy to determine the affinities of benzoate and muconate for both effector binding sites of full-length BenM and two variants, BenM(E226K) and BenM(Y293F). BenM(E226K) is a constitutively active mutant whose activity was found to be independent of benzoate or muconate, while BenM(Y293F) completely loses benzoate-responsive transcriptional activation due to a mutation at one of the key residues in binding of benzoate to the secondary site.

Results

Intrinsic fluorescence of BenM, BenM(E226K), and BenM(Y293F)

Generally, tryptophan residue is excited at 280 nm or at longer wavelengths giving the emission maximum in water near 350 nm. However, different proteins have been reported to display vastly different emission maxima (308-350 nm) and quantum yields[18; 19]. When purified BenM was excited at 280 nm, a single tryptophan residue at the C-terminus of BenM displayed a tryptophan fluorescence peak at approximately 350 nm and two extraneous fluorescence peaks at 364 and 380 nm (Figure 3.3). The two BenM mutants, BenM(E226K) and BenM(Y293F), also showed the same fluorescence profile. These extraneous signals were eliminated by exhaustive dialysis of the protein solutions against several changes of a freshly prepared dialysis buffer. The dialyzed BenM, BenM(E226K), and BenM(R156H) exhibited single intrinsic fluorescence peak at 326 nm (Figure 3.3). Purified BenM WT and BenM(E226K) were confirmed to be functional DNA-binding proteins by EMSA as the regulators were able to bind promoters located within an intergenic region between *benM* and *benA* promoters (Appendix A).

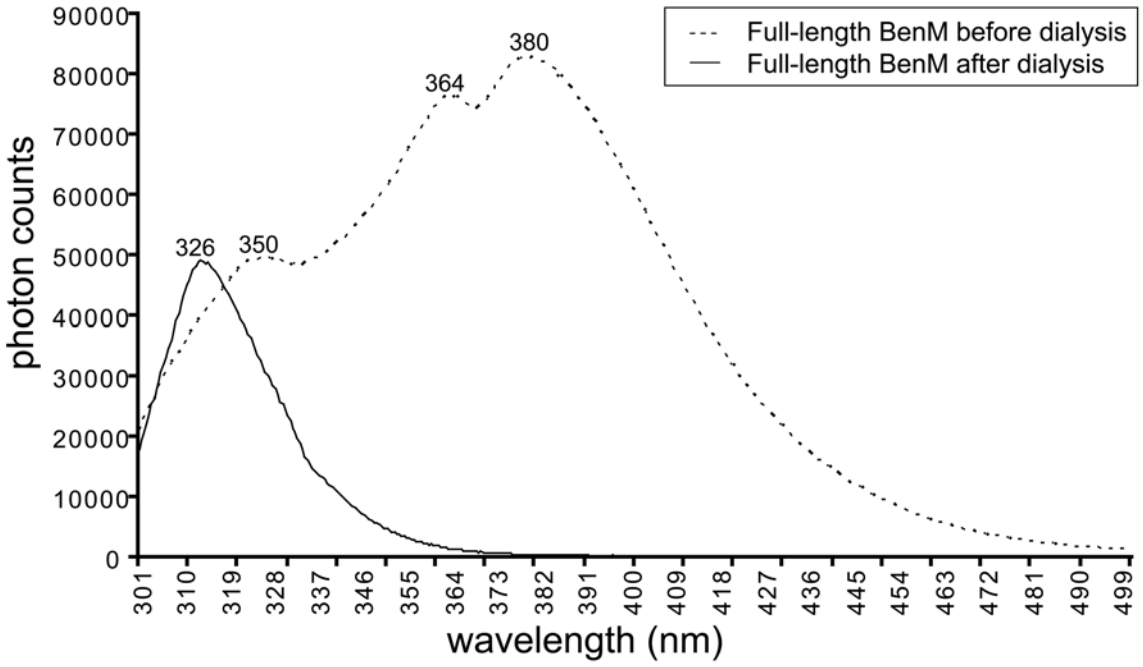

Figure 3.3. Fluorescence emission spectra of purified full-length BenM before (represented by the dashed line) and after (represent by the solid line) dialysis upon excitation with 280 nm wavelength.

Dissociation constants of BenM for benzoate

Initially, emission spectra of 2 μM histidine-tagged full-length BenM in the dialysis buffer displayed the tryptophan peaks at 326 nm. Upon benzoate titration into the protein solution, descending peaks were observed in a benzoate concentration-dependent manner. Previous works had shown a red-shift, but this was not observed in this present study[21]. Normalized values of the quenching of the fluorescence signals at 326 nm were plotted against ligand concentrations, giving a hyperbolic curve that could be fitted well into a one-site ligand

binding model (Figure 3.4A). The calculated dissociation constant of BenM for benzoate was approximately 7.5 μM (Table 3.1).

According to the crystal structures of BenM EBD, two binding sites with dissimilar binding affinity for benzoate are present. The dissociation constant for the high affinity binding site was therefore estimated using a lower protein concentration (0.2 μM) and a lower range of benzoate titration (0.01-10 μM). At the lower benzoate concentrations, the decreases in fluorescence spectra were also hyperbolically related to benzoate concentrations, and the data were also fit into a one-site ligand binding equation giving an estimated dissociation constant of 0.04 μM for the high affinity binding site (Figure 3.4B, Table 3.1).

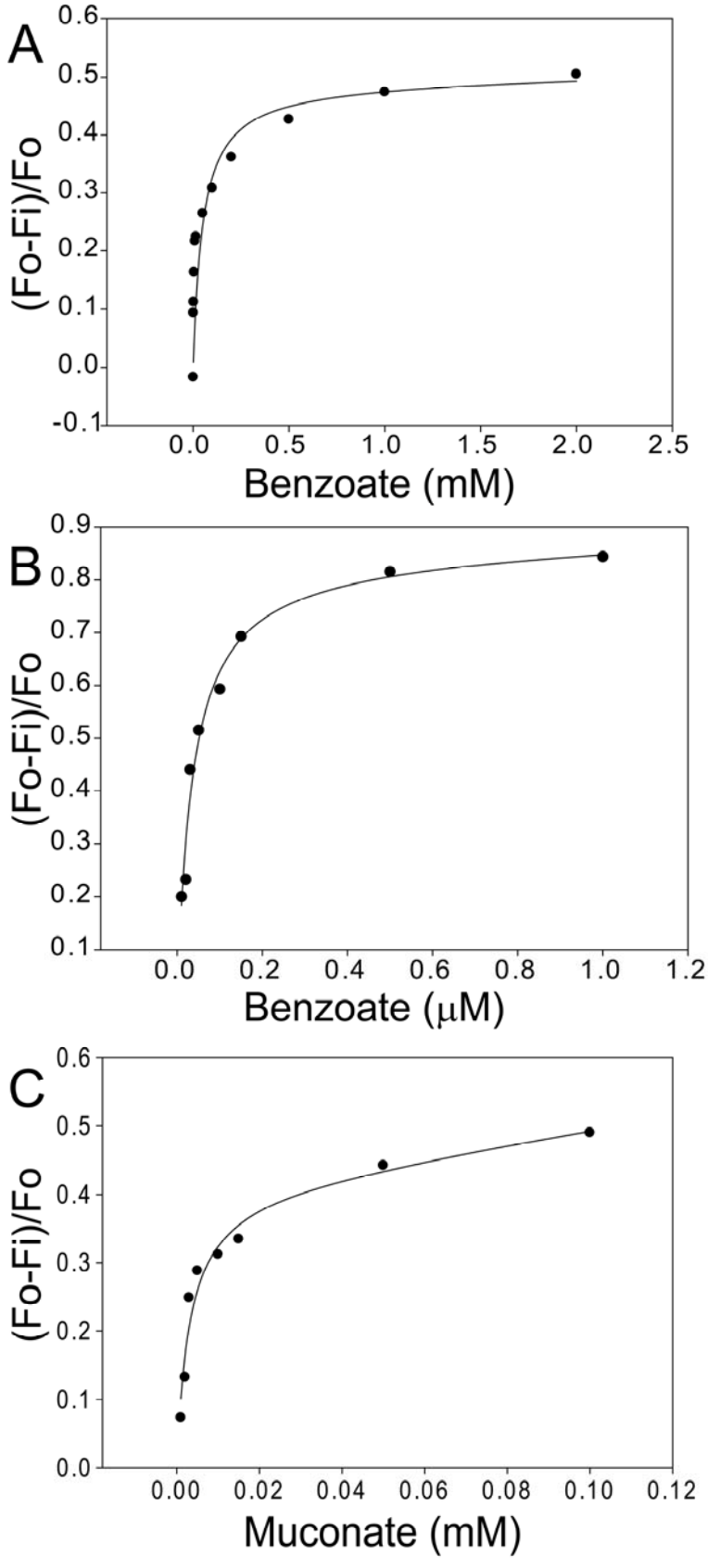
A
(Fo-Fi)/Fo
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
0.0
0.5
1.0
1.5
2.0
2.5
Benzoate (mM)
B
(Fo-Fi)/Fo
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Benzoate (μM)
C
(Fo-Fi)/Fo
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Muconate (mM)

Figure 3.4. Binding of benzoate (A and B) and muconate (C) to wild-type BenM monitored as the changes of relative fluorescence intensity quenching, (Fo-Fi)/Fo, measured at 326 nm. Fo represents the initial fluorescence, and Fi is the fluorescence at concentration i of the effector. The decrease of intrinsic fluorescence caused by titration with benzoate or muconate was plotted versus the effector concentration. Inner filter effects were corrected for each observed data point [22]. The hyperbolic curves were subsequently fitted to a ligand binding equation. These plots are representative titrations. The concentrations of BenM were 2 μM (A and C) and 0.2 μM (B).

Dissociation constant of BenM for muconate

The crystal structures of BenM EBD showed electron density consisted of only one binding site for muconate, which was called the primary binding site, located between subdomains I and II of the EBD. Based on the estimated dissociation constant of BenM for benzoate, dissociation constant of the regulator for muconate was also expected to be in the micromolar (μM) range. Therefore, the same protein concentration and similar range of ligand concentrations were used. The dissociation constant of wild-type BenM for muconate was estimated by adding muconate to achieve the final concentration of 0.001-2 mM into 2 μM purified protein solution. Quenching of intrinsic fluorescence peaks after each titration was normalized and plotted against the muconate concentrations yielding a hyperbolic curve (Figure 3.4C). Curve fitting of the plotted data into a one-site ligand binding equation yield the dissociation constant of BenM for muconate of 3.9 μM (Table 3.1). The attempt to identify a higher affinity binding site for BenM by titrating in 0.01-10 μM muconate into 0.2 μM BenM was unsuccessful due to the lack of systematic quenching of the intrinsic fluorescence of BenM (data not shown).

Additionally, it was reported from the crystal structures of BenM EBD that the regulator could bind benzoate and muconate simultaneously. Further, from *in vitro* transcription experiments, benzoate and muconate were shown to have synergistic effect on transcriptional activation of *benA* by BenM[23]. This synergistic effect could arise from an allosteric modification of BenM that will increase the affinity for muconate. To observe the effect of benzoate on the protein's affinity for muconate, fixed concentration of benzoate corresponding to the high affinity K_d of benzoate was present in a protein solution, and muconate was subsequently titrated into the protein solution. In the presence of 40 nM benzoate, the dissociation constant of BenM for muconate is approximately 3.3 μM (Table 3.1).

Dissociation constants of BenM(E226K) for benzoate and muconate

BenM(E226K) is a constitutively active BenM mutant whose activity is not altered dependent on the inducers[14]. *In vivo* experiments revealed that addition of inducers had no effects on the transcriptional activation by the mutant, but it was unclear whether the protein could accommodate benzoate or muconate in its binding sites[14]. A similar strategy was employed to determine the dissociation constants of BenM(E226K) for benzoate and muconate. Again, a solution of 2 μM BenM(E226K) was used, and benzoate or muconate was added to reach the final concentration range of 0.001-2 mM. However, the changes in fluorescence intensity were not pronounced after each addition of the ligands, and no systematic quenching of the fluorescence signals was observed, preventing a determination of the dissociation constant (data not shown).

Dissociation constants of BenM(Y293F) for benzoate and muconate

Arg160 and Tyr293 are two key residues critical for binding of benzoate to the secondary binding site as observed in the structures of wild-type BenM EBD soaked with effectors[5]. BenM(Y293F) whose Tyr293 is replaced by Phe293 represents a mutant that is likely to have disrupted secondary binding site as suggested by *in vivo* studies[14]. Titrations of benzoate and muconate into the mutant both resulted in systematic quenching of intrinsic protein fluorescence. The decreases in fluorescence intensity were plotted against ligand concentrations, and the resulting curves were also fitted into a hyperbolic binding isotherm (Figure 3.5). Dissociation constants of BenM(Y293F) estimated for benzoate and muconate were 7.2 and 5.7 μM, respectively (Table 3.1). However, when lower protein concentration (0.2 μM) and lower ligand concentrations (0.01-10 μM) were used to determine dissociation constant of the high affinity binding site, no systematic quenching was observed (data not shown).

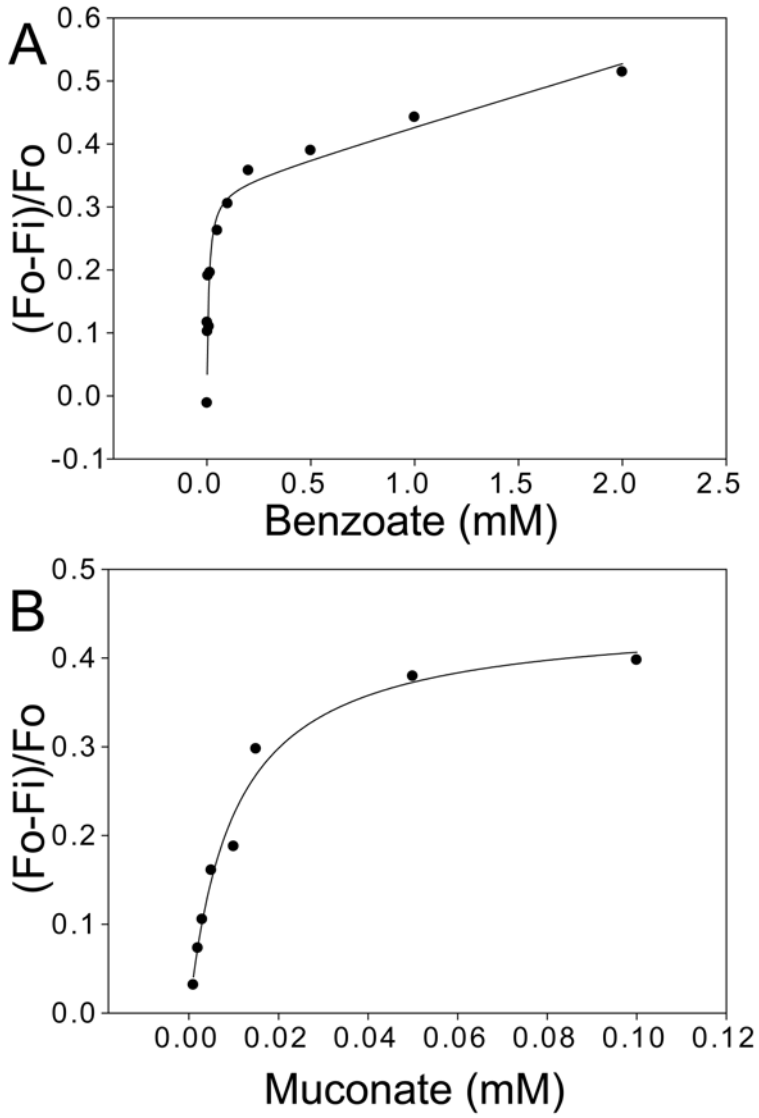

Figure 3.5. The hyperbolic plots representing changes of relative fluorescence intensity quenching, (Fo-Fi)/Fo, measured at 326 nm upon titration of BenM(Y293F) with benzoate (A) or muconate (B). The concentrations of BenM(Y293F) were 2 μM for both titrations.

Table 3.1. Dissociation constants (K_d) of BenM and BenM variants

Effector	K_d (μM)[a]		
	BenM	BenM(E226K)	BenM(Y293F)
Benzoate at low affinity site	7.5 ± 0.7	No systematic quenching	7.2 ± 1.4
Benzoate at high affinity site	0.040 ± 0.002	-	No systematic quenching
Muconate	3.9 ± 0.7	No systematic quenching	5.7 ± 0.6
Muconate + 40 nM benzoate	3.3 ± 0.6	-	-

[a]The dissociation constants were determined by non-linear regression with the one site saturation ligand binding equation, $\Delta F/Fo = (\Delta F_{max}[E]/Kd+[E]) + N_s[E]$ using SigmaPlot 11.0 to fit the data. ΔF/Fo is relative fluorescence intensity quenching. [E] is the effector concentration. N_s accounts for non-specific ligand binding.

Discussion

The single tryptophan residue located at the C-terminal end of BenM emitted a fluorescence peak at 326 nm with an excitation wavelength of 280 nm. Quenching of the tryptophan signals upon the additions of benzoate or muconate implied that binding of the natural effectors to BenM introduced alterations in the local environments of tryptophan residues and possibly the protein conformation. This allowed us to investigate the effector binding characteristics and calculate the affinity of the regulator for its effectors. The ability of different ligands to interact with BenM with dissimilar affinities at distinct binding sites observed in this

study is very intriguing. These characteristics of BenM may be important for the regulator to integrate complex signals from multiple exogenous inducers and regulate a complex metabolic pathway accordingly.

The dissociation constant of BenM for benzoate estimated from this fluorescence emission experiment is approximately 7.5 μM, which is significantly lower than the previously reported value of 1.2 mM[21] obtained by the same technique. Similar results were also observed for the dissociation constant of BenM for muconate. It is critical to note that the previously determined dissociation constants were also markedly high relative to the dissociation constants of DntR or Cbl for their effectors[9; 19; 20]. One possible explanation for this discrepancy in the K_d for BenM is the variation in the purification and preparation protocols of BenM for the experiments. Previously, after the purification step by metal chelate column, BenM peak fractions were pooled together and diluted with a new buffer to reach the final monomeric concentration of 2 μM and used directly in the assays without undergoing further dialysis[21]. However, it has been demonstrated in the crystal structures of BenM and other family members that different kinds of ions are capable of binding to the primary binding site of LTTRs[9; 11; 14]. There is a high possibility that certain ligands from early steps of overexpression or purification remain bound to the protein with adequately high affinity that a single purification step cannot remove the ions from the site. In this present study, the fluorescence spectra of purified BenM before undergoing exhaustive dialysis exhibited the two extraneous peaks at 364 and 360 nm. These findings further suggested the presence of fluorescent ligands carried over from the protein preparation steps. Therefore, in order to replace the ligands from the binding sites, exaggeratedly high concentrations of benzoate or muconate are required. In addition, the shift of fluorescence peak particularly observed in the previous experiments might result from changing

of the environment of the tryptophan residue due to displacement of the contaminate ligands from the binding site. In this present study, additional rounds of exhaustive dialysis were introduced after BenM was purified by affinity column. It is likely that the more exhaustive dialysis was a highly crucial step that helped eliminate any ions that potentially stay bound in the binding sites of BenM and, consequently, allowed better estimation of the dissociation constants.

The crystal structures of BenM EBD soaked with benzoate and with muconate revealed specific interactions between residues of the regulator and its effectors[5]. In the primary binding site, a number of interactions were observed in the muconate-bound BenM structure, while binding of benzoate to the same site involved fewer interactions[21]. These findings are consistent with the estimated dissociation constants of wild-type BenM for benzoate and muconate from our study. Higher dissociation constant for benzoate (7.5 μM) than that for muconate (3.9 μM) implies that the primary site of BenM has a higher affinity for muconate than for benzoate, which consistent with the difference in the number of specific interaction utilized in muconate binding than in benzoate binding.

Structural studies of BenM-EBD in complex with its natural effectors revealed unexpected secondary binding pocket located near the primary binding site which had never been identified in any other LTTRs. This secondary binding site was considered to have relatively high affinity for benzoate considering both the presence of the effector at this site even when no benzoate were purposely added to the crystallization conditions and the significant hydrophobic and ionic interactions at the site[5]. The fluorescence spectroscopy study confirmed the presence of a second site for inducer binding in BenM. Our results correspondingly revealed the dissociation constant of the high affinity binding site or, in other words, the secondary

binding site for benzoate of 40 nM, which is significantly lower than the binding constant of the lower affinity site, which is likely the primary site for either benzoate or muconate.

Simultaneous binding of muconate at the primary binding site and benzoate in the secondary binding site was demonstrated in the crystal structure of BenM EBD soaked with muconate[5]. This structure was proposed to represent how the two effectors simultaneously act on BenM to create synergistic effect on the transcriptional activation. Synergism potentially resulted from the increased affinity of the protein for muconate at the primary binding site or enhanced conformational changes after binding of benzoate to the adjacent secondary binding pocket. In this study, the presence of small amount of benzoate (40 nM) yielded a lower dissociation constant of BenM for muconate of 3.3 μM when using 2 μM BenM. However, this change in the affinity for muconate was not pronounced potentially due to too low concentrations of benzoate were used relative to the protein concentrations. On the contrary, in the previous studies, the presence of 1 mM benzoate was shown to reduce the affinity of the protein to muconate[21]. This possibly because, at 1 mM which exceeds the dissociation constant of the primary site of BenM for benzoate as identified in the present study, benzoate molecules are able to bind to both the secondary and primary binding site resulting in competition between benzoate and muconate for the primary binding site. The presence of high amount of benzoate is thus able to lower the apparent affinity of the regulator for muconate at the primary binding site.

The dissociation constants of two BenM mutants, BenM(E226K) and BenM(Y293F), for benzoate and muconate were also determined by fluorescence spectroscopy. As reported previously, BenM(E226K) is a constitutively active mutant that activates high-level *benA* expression without requirements for exogenous effectors[14]. Additionally, the presence of benzoate or muconate has no effects on transcriptional activation by the protein[14]. It has been

suggested that BenM(E226K) already assumes an activated conformation, and binding of muconate or benzoate is dispensable for the activity of this mutant. Addition of benzoate or muconate did not appear to alter the conformation of the protein or the tryptophan environments as no systematic alterations in fluorescence signal quenching were observed from the experiments. Consequently, we were unable to obtain dissociation constants for BenM(E226K). These findings implied that the mutation in BenM(E226K) locked the protein in the conformation that resembled liganded BenM, and thus, addition of effector could not introduced additional changes in tryptophan conformation. However, this does not mean that benzoate binding is impossible. Instead, there is no conformational change that could be detect in this study to reveal binding. It is still unclear how the mutation outside of the binding sites affects the ability of the protein to recognize its cognate effectors and the ability to constitutively activate the transcription of the target genes.

BenM(Y293F), whose secondary site is missing the hydroxyl group from tyrosine residue crucial for interacting with benzoate, displays the dissociation constant for benzoate in the low affinity site of 7.2 μM, which is comparable to the dissociation constant of wild-type BenM (7.5 μM). The 5.7 μM dissociation constant of this mutant for muconate is only a slightly higher than the dissociation constant of wild-type BenM for muconate (3.9 μM). When low protein and low benzoate concentrations were used in the experiments to determine binding constants of the high affinity binding site of the protein, no systematic quenching was detected resulting in an inability to derive binding constants for benzoate at the secondary pocket. These findings imply that the mutation of Tyr293 to Phe293 potentially generated a nonfunctional high affinity binding site and annulled the ability of the regulator to bind benzoate at this site. However, it is also possible that the protein may still bind benzoate, but the binding event does not induce a change in the

protein conformation or the tryptophan environments. Binding of benzoate in the secondary site has been proposed to be absolutely essential for response of BenM to benzoate as well as to impact the synergistic effect of the regulator based on previous *in vivo* experiments. The results from this study further support this hypothesis[14].

Acknowledgements

We gratefully acknowledge the staff at Olis Inc. (Stoyan Milev) for suggestions and assistance with experimental design and data analysis.

Materials and Methods

Full-length BenM with a C-terminal histidine tag was overexpressed and purified using metal chelating column as described in Chapter 2. The histidine tagged protein was previously shown to function normally in gene transcription regulation [4; 21]. After the metal chelating column purification step, the fractions containing BenM were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis, pooled together and dialyzed exhaustively to a buffer containing 20 mM Tris, 10% glycerol, and 100 mM NaCl adjusted to pH 7.9 with concentrated HCl. Protein concentrations were determined by Bradford assay using bovine serum albumin (BSA) as a standard. For initial dissociation constant estimations, BenM was diluted to 2 μM (monomer) with freshly prepared dialysis buffer to reach a final volume of 2 ml, and stock solutions of benzoate or muconate (1-2 μl) were titrated to reach final effector concentrations of 0.001, 0.002, 0.003, 0.005, 0.01, 0.015, 0.05, 0.1, 0.2, 0.5, 1 and 2 mM. After each titration, samples were incubated at 25 °C for 10 min before being measured for fluorescence emission spectra. Fluorescence spectra were collected using Olis DM 45 Spectrofluorimeter. Samples

were excited with 280 nm wavelength, and emission spectra from 500-300 nm were collected using a 200 increment and 0.5 second integration time. Data reduction mode was set as photon counting. Slit widths before the excitation monochromator, after the excitation monochromator, and before the detector were all set as 5 nm. Data were plotted in SigmaPlot software version 11.0 (SPSS Inc.), and dissociation constants were determined using a hyperbolic regression function. Averaged dissociation constants were estimated and standard deviations were calculated from at least three separate experiments. To determine the dissociation constant of the high affinity site of BenM, the experiments were performed using 0.2 μM BenM, and benzoate was added to reach final concentration of 0.01, 0.02, 0.03, 0.05, 0.1, 0.15, 0.5, 1, 5 and 10 μM. Similar instrument settings and strategy were employed.

BenM(E226K) and BenM(Y293F) were purified according to the same protocol. Titrations of ligands to determine dissociation constants of both binding sites were performed following the same method used for wild-type BenM.

References

1. Collier, L. S., Gaines Iii, G. L. & Neidle, E. L. (1998). Regulation of benzoate degradation in Acinetobacter sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *Journal of bacteriology* **180**, 2493.
2. Craven, S. H., Ezezika, O. C., Momany, C. & Neidle, E. L. (2008). LysR homologs in Acinetobacter: insights into a diverse and prevalent family of transcriptional regulators. *Acinetobacter Molecular Biology*, 163–202.
3. Harwood, C. S. & Parales, R. E. (1996). The -ketoadipate pathway and the biology of self-identity. *Annual Reviews in Microbiology* **50**, 553-590.
4. Bundy, B. M., Collier, L. S., Hoover, T. R. & Neidle, E. L. (2002). Synergistic transcriptional activation by one regulatory protein in response to two metabolites. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 7693.
5. Ezezika, O. C., Haddad, S., Clark, T. J., Neidle, E. L. & Momany, C. (2007). Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. *Journal of molecular biology* **367**, 616-629.
6. Lochowska, A., Iwanicka-Nowicka, R., Plochocka, D. & Hryniewicz, M. M. (2001). Functional dissection of the LysR-type CysB transcriptional regulator. *Journal of Biological Chemistry* **276**, 2098.
7. Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K. & Senda, T. (2003). Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. *Journal of molecular biology* **328**, 555-566.

8. Schell, M. A., Brown, P. H. & Raju, S. (1990). Use of saturation mutagenesis to localize probable functional domains in the NahR protein, a LysR-type transcription activator. *Journal of Biological Chemistry* **265**, 3844.

9. Smirnova, I. A., Dian, C., Leonard, G. A., McSweeney, S., Birse, D. & Brzezinski, P. (2004). Development of a bacterial biosensor for nitrotoluenes: the crystal structure of the transcriptional regulator DntR. *Journal of molecular biology* **340**, 405-418.

10. Monferrer, D., Tralau, T., Kertesz, M. A., Dix, I., Solι, M. & Usσn, I. Structural studies on the full-length LysR-type regulator TsaR from Comamonas testosteroni T-2 reveal a novel open conformation of the tetrameric LTTR fold. *Molecular Microbiology* **75**, 1199-1214.

11. Tyrrell, R., Verschueren, K. H. G., Dodson, E. J., Murshudov, G. N., Addy, C. & Wilkinson, A. J. (1997). The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. *Structure* **5**, 1017-1032.

12. Lang, G. & Ogawa, N. (2009). Mutational analysis of the inducer recognition sites of the LysR-type transcriptional regulator TfdT of Burkholderia sp. NK8. *Applied microbiology and biotechnology* **83**, 1085-1094.

13. Cebolla, A., Sousa, C. & de Lorenzo, V. (1997). Effector specificity mutants of the transcriptional activator NahR of naphthalene degrading Pseudomonas define protein sites involved in binding of aromatic inducers. *Journal of Biological Chemistry* **272**, 3986.

14. Craven, S. H., Ezezika, O. C., Haddad, S., Hall, R. A., Momany, C. & Neidle, E. L. (2009). Inducer responses of BenM, a LysR-type transcriptional regulator from Acinetobacter baylyi ADP1. *Molecular microbiology* **72**, 881-894.

15. Maddocks, S. E. & Oyston, P. C. F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology-Sgm* **154**, 3609-3623.

16. Lochowska, A., Iwanicka-Nowicka, R., Plochocka, D. & Hryniewicz, M. M. (2001). Functional Dissection of the LysR-type CysB Transcriptional Regulator. *Journal of Biological Chemistry* **276**, 2098-2107.

17. Zhou, X., Lou, Z., Fu, S., Yang, A., Shen, H., Li, Z., Feng, Y., Bartlam, M., Wang, H. & Rao, Z. (2009). Crystal Structure of ArgP from Mycobacterium tuberculosis Confirms Two Distinct Conformations of Full-length LysR Transcriptional Regulators and Reveals Its Function in DNA Binding and Transcriptional Regulation. *Journal of Molecular Biology*.

18. Lakowicz, J. R. & Masters, B. R. (2008). Principles of fluorescence spectroscopy. *Journal of Biomedical Optics* **13**, 029901.

19. Lönneborg, R., Smirnova, I., Dian, C., Leonard, G. A. & Brzezinski, P. (2007). In vivo and in vitro investigation of transcriptional regulation by DntR. *Journal of molecular biology* **372**, 571-582.

20. Stec, E., Witkowska-Zimny, M., Hryniewicz, M. M., Neumann, P., Wilkinson, A. J., Brzozowski, A. M., Verma, C. S., Zaim, J., Wysocki, S. & D Bujacz, G. (2006). Structural basis of the sulphate starvation response in E. coli: crystal structure and mutational analysis of the cofactor-binding domain of the Cbl transcriptional regulator. *Journal of molecular biology* **364**, 309-322.

21. Clark, T. J., Phillips, R. S., Bundy, B. M., Momany, C. & Neidle, E. L. (2004). Benzoate decreases the binding of cis, cis-muconate to the BenM regulator despite the synergistic effect of both compounds on transcriptional activation. *Journal of bacteriology* **186**, 1200.

22. Palmier, M. O. & Van Doren, S. R. (2007). Rapid determination of enzyme kinetics from fluorescence: Overcoming the inner filter effect. *Analytical biochemistry* **371**, 43-51.

23. Bundy, B. M., Collier, L. S., Hoover, T. R. & Neidle, E. L. (2002). Synergistic transcriptional activation by one regulatory protein in response to two metabolites. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 7693-7698.

CHAPTER 4

CONCLUSIONS

The structural characterization of BenM presented in Chapter 2 of this dissertation provides the first look at the structure of full-length BenM. Several interesting observations were derived from the detailed investigation of the crystal structure. The protein fold and domain organization of BenM substantially resembled that of CbnR and all other full-length LTTR structures, yet each regulator exhibited unique DBD orientations. Flexibility at the hinge region connecting the LH to the EBD is proposed as an origin of these different DBD placements. Nevertheless, the fundamental of this flexibility remains uncharacterized. It is possible that certain residues or specific sequence motifs in this region play an exclusive role in directing the LH into a definite orientation. Comprehensive analysis of the residues comprising the hinge may ultimately reveal how sequence variations affect the orientations of DBD and may assist in predicting the trend of DBD orientation from a given LTTR amino acid sequence.

Gel filtration studies of BenM implied that the protein formed tetramers in solution. However, in the present study, full-length BenM was not crystallized as a tetrameric unit as expected from the chromatography results. Instead, a network of interlocking BenM molecules was found to constitute the entire crystal. These findings clearly suggested that the protein was not crystallized in its biologically functional form and, therefore, the possibility to draw any solid conclusion on how the protein actually functions *in vivo* was limited. Although the crystal structure of BenM did not give an accurate portrayal of the biologically active BenM, the

comprehensive investigation of the structure allowed us to identified interfaces involved in BenM oligomerization. The discovery of these interfaces encouraged us to pursue a more general assessment of the interfaces utilized in oligomerization of other full-length LTTRs. The categorization of LTTRs based on their schemes of oligomerization was introduced. Among all the full-length structures available, three major schemes of LTTR oligomerization were proposed. These findings are highly intriguing considering all LTTRs bear remarkably similar folds and domain organization. Again, the flexibility of LTTR structures may be an important factor in allowing the regulators to adopt multiple modes of oligomerization. Nonetheless, it is still a mystery how a protein defines which scheme to follow. The rationale behind these differences and the importance of various modes of oligomerization in transcriptional regulation are still areas of research that are yet to be explored.

Structural investigation of two BenM mutants, BenM(E226K) and BenM(R156), was expected to shed light on how these mutations affect the overall structures of the protein and how the conformations of BenM mutants differ from that of unliganded wild-type regulator. Although these two mutants are both able to actively turn on *benA* gene transcription without the presence of benzoate or muconate, their responses to effectors are dissimilar. Initially, the two BenM variants were suggested to assume distinct conformations that allowed the proteins to turn on gene transcription without effectors. Unexpectedly, the structures of both variants turned out to be almost identical to each other and to the structure of wild-type BenM. These surprising findings bring about novel interesting hypotheses on how the mutants possibly obtain their constitutive activities. Instead of being locked in an active conformation, the BenM variants may be able to recognize different effectors, may oligomerize differently and/or may exhibit better interactions with RNA polymerase. However, because both BenM(E226K) and BenM(R156H)

were not crystallized in their functional tetrameric units, the observed conformations of these proteins might merely be an artifact of the crystallization conditions. There is a possibility that the crystallization conditions used in the studies constrained both mutants in a definite conformation that did not represent the *in vivo* state of the proteins. In this case, it is inconclusive that BenM(E226K) and BenM(R156) actually adopt the identical conformation to that of wild-type BenM *in vivo*.

Another important aspect derived from the work presented in this dissertation regarding crystallization is that structures solved from X-ray crystallographic technique do not always bear biologically relevant conformations. This implication is based on the discovery of the infinite array of interlocking molecules instead of tetramers in the crystals of BenM, BenM(E226K) and BenM(R156H). As a result, interpretation of a crystal structure of an LTTR, especially that crystallized in an unusual oligomerization mode, should be conducted with extreme care.

Although the attempt to structurally characterize BenM only allowed visualization of the protein in a non-biological oligomeric form, biophysical studies of BenM indicated that the regulator could still bind to its cognate effectors and most likely retained its activity in solution. As reported in Chapter 3, the values of the dissociation constants of wild-type BenM for benzoate and muconate conform to the nature of interactions observed in the binding site of the BenM-EBD crystal structures. Muconate, as expected, binds to only one site of BenM with the K_d of 3.9 μM while benzoate binds to two sites with dissimilar K_ds of 7.5 and 0.04 μM. The studies also confirmed the necessity of Tyr293 for the binding of benzoate at the secondary site. The effects on inducer binding of mutations at other residues of BenM still need to be evaluated. Another interesting aspect that could be clarified by tryptophan fluorescence studies of BenM is the effects of the DNA on the affinity of BenM for its effectors.

Binding properties of BenM to the *ben* promoter region investigated by electrophoretic mobility shift assay (EMSA) (Appendix A) suggested that the regulator maintained its function at low protein concentrations in solution. Wild-type BenM was found to shift a band of DNA encompassing *benMA* intergenic region with an apparent K_d between 20-35 nM. Smeary bands of the complex might indicate that the protein, although bound to DNA, still displayed some flexibility and was not fixed in one single conformation in solution. Addition of muconate into binding reactions did not appear to change the affinity of the regulator for DNA, but yielded sharper protein-DNA complex bands potentially because BenM was locked in one uniform active conformation. Therefore, introduction of muconate into the crystallization conditions may reduce the flexibility of BenM and assist in obtaining uniform, biologically active crystal structures. Moreover, it is important to note that, at micromolar BenM concentration, EMSA displayed BenM-DNA complex bands that were shifted to markedly higher positions implying that at these concentrations, the protein tended to assemble into large oligomers or even precipitate but was still able to bind to DNA. As the concentrations of BenM used in crystallization studies were in the high micromolar range, it is very likely that the unnaturally large oligomers are present in the solution.

Although the full-length structures of BenM did not truly elucidate the mechanism of transcriptional regulation by LTTRs, structural and biophysical studies of BenM and its variants revealed several interesting characteristics of the proteins and introduced novel idea regarding how transcriptional regulators in the LysR family function. To achieve an ultimate answer of how BenM functions in transcriptional regulation, the structure of the full-length protein in complex with DNA and effectors and additional analyses including structural characterization of the complex of BenM, DNA, and effectors with RNA polymerase are unquestionably required.

APPENDIX A

ELECTROPHORETIC MOBILITY SHIFT ASSAYS OF BENM AND BENM VARIANTS

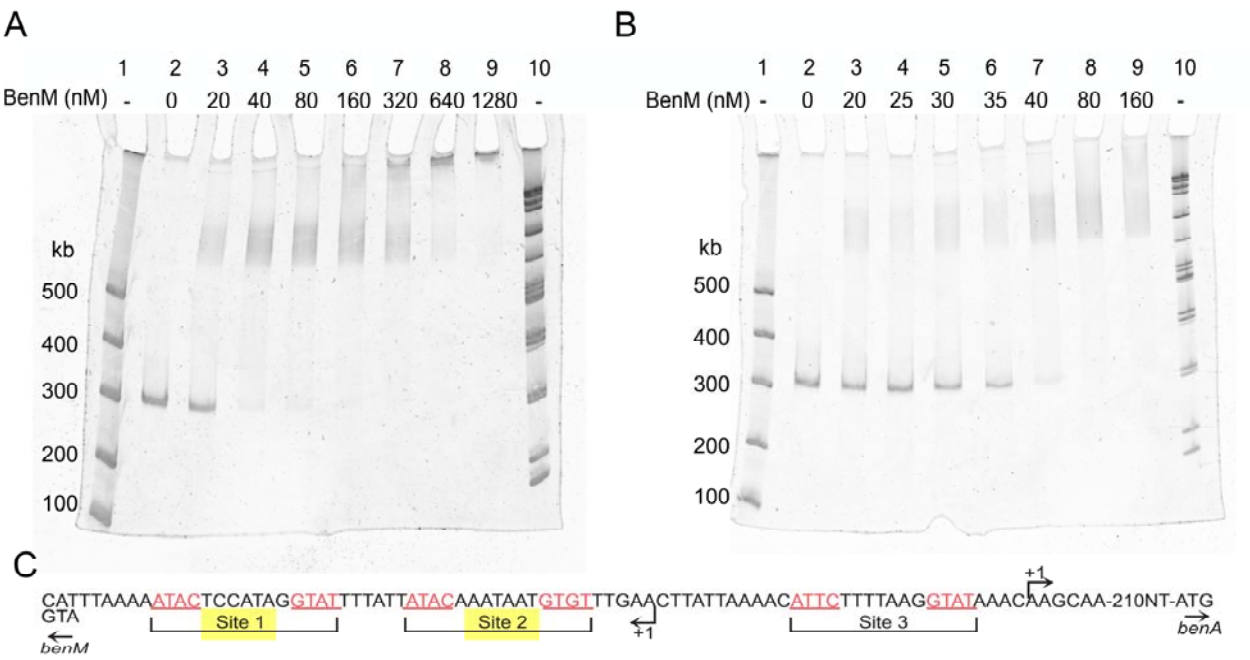

Figure A.1. Binding of wild-type BenM to the *ben* promoter region as measured by electrophoretic mobility shift assay (EMSA). The 440-bp DNA fragment harboring Site 1, Site 2, and Site 3 for BenM binding was incubated with varying concentrations of full-length BenM-His, and the binding reactions were resolved on 6% polyacrylamide gels (A and B). Lanes 1 and 10 are the PCR markers (New England Biolabs) and the 1-kb DNA ladder (Promega), respectively. DNA bands were stained with SYBR Green and visualized by a Storm 865 fluorescence imaging system (GE Healthcare). Under these conditions, BenM is proposed to interact with Site 1 and Site 3 of the *benM-benA* intergenic region (C).

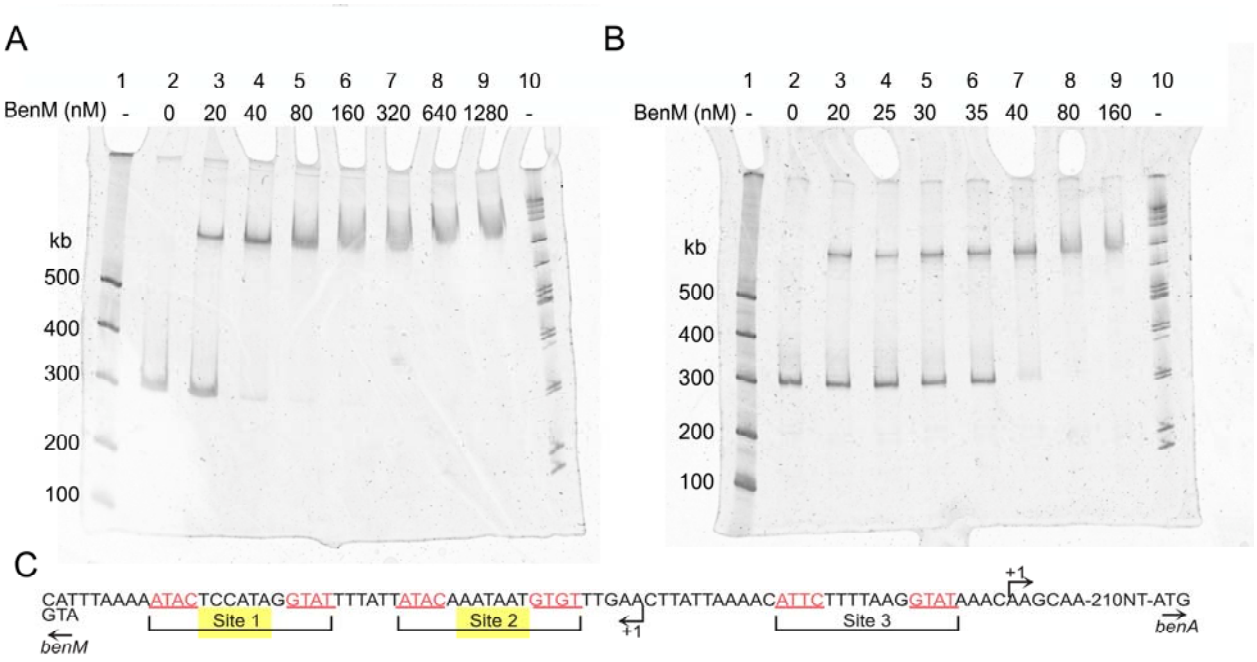

Figure A.2. EMSA of wild-type BenM with the *ben* promoter region in the presence of 2 mM muconate. The protein samples, DNA fragment, and percentage of the gels used are described in the caption of Figure A.1. The gels were pre-run in running buffer containing 2 mM muconate before loading the samples. Muconate was also added to the binding reactions and the running buffer to achieve the final concentration of 2 mM. Lanes 1 and are the PCR markers (New England Biolabs) and the 1-kb DNA ladder (Promega), respectively. The presence of muconate is proposed to induce conformational changes of BenM that allow the protein to interact with Site 1 and Site 2 of the *ben* gene operator region (C).

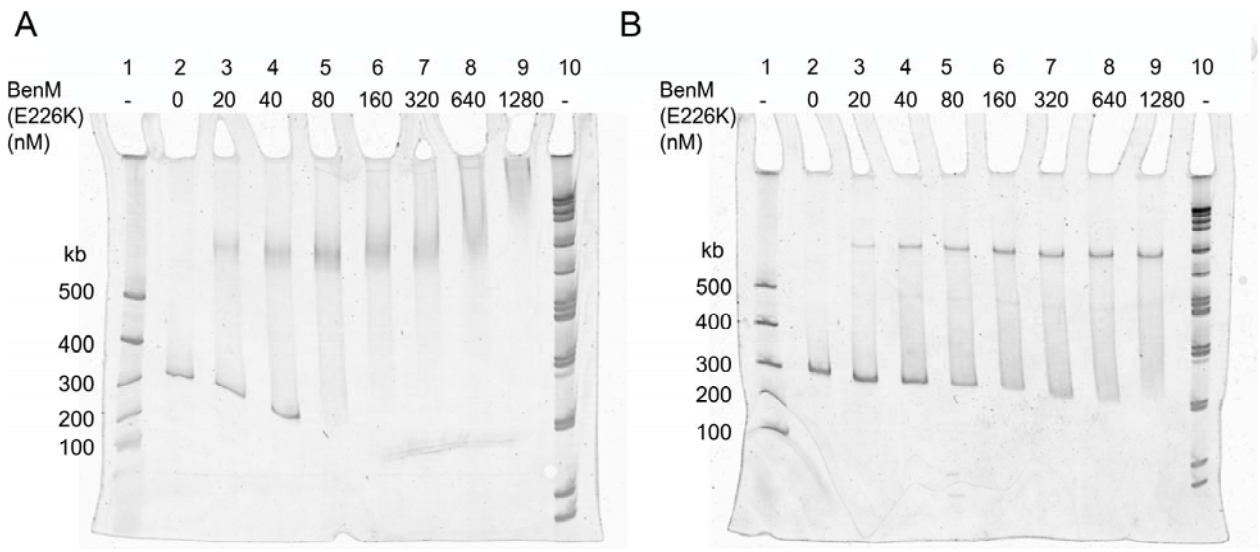

Figure A.3. Binding of BenM(E226K) to the 440-bp *ben* promoter region (A) in the absence and (B) in the presence of 2 mM muconate as measured by EMSA. The protein samples, DNA fragment, and percentage of the gels used are described in the captions of Figure A.1 and Figure A.2. Lanes 1 and 10 are the PCR markers (New England Biolabs) and the 1-kb DNA ladder (Promega), respectively.

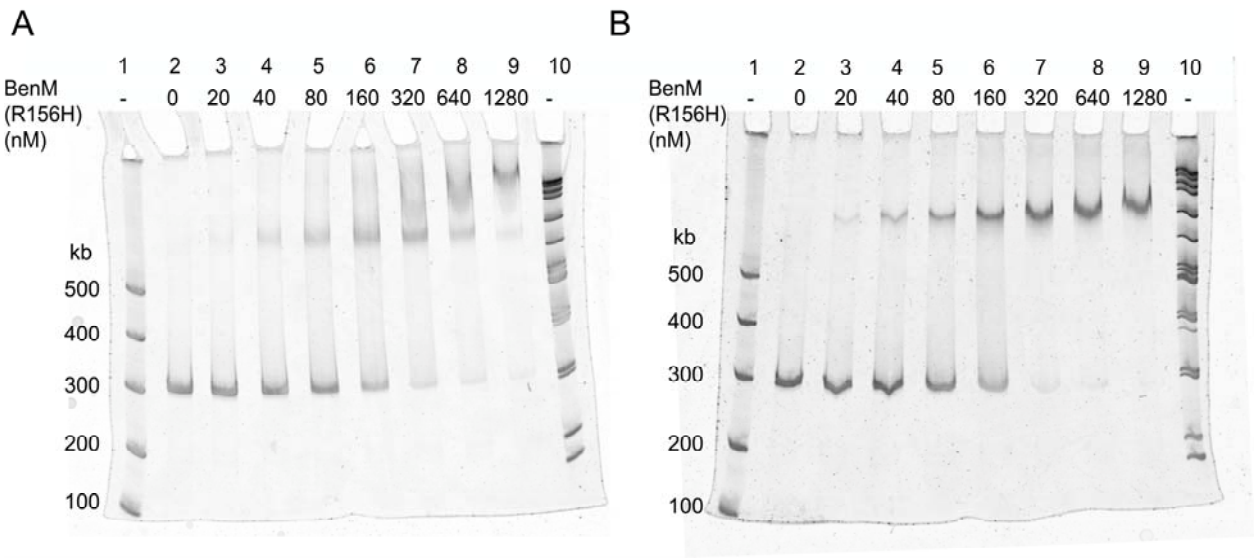

Figure A.4. Binding of BenM(R156H) to the 440-bp *ben* promoter region (A) in the absence and (B) in the presence of 2 mM muconate as measured by EMSA. The protein samples, DNA fragment, and percentage of the gels used are described in the captions of Figure A.1 and Figure A.2. Lanes 1 and 10 are the PCR markers (New England Biolabs) and the 1-kb DNA ladder (Promega), respectively.

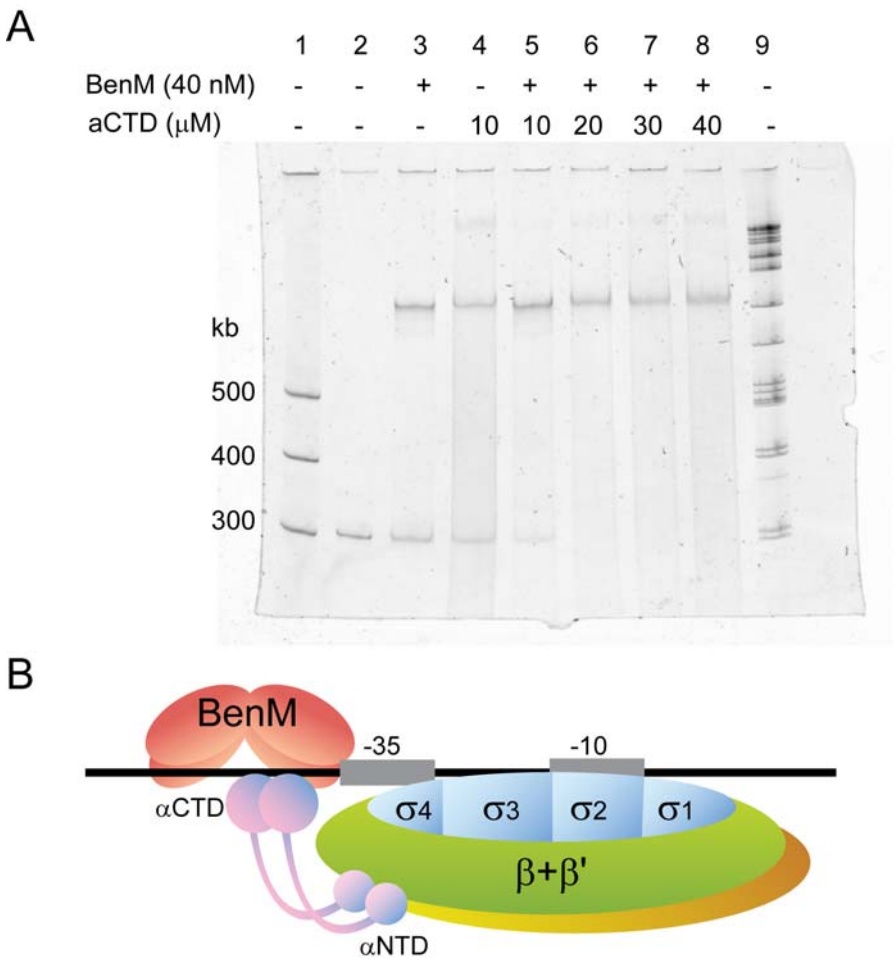

Figure A.5. (A) Binding of wild-type BenM and αCTD of *Acinetobacter* ADP1 to the *ben* promoter region in the presence of 2 mM muconate as investigated by EMSA. Binding reactions contained the 440-bp DNA fragment of the *benM-benA* intergenic region, 40 nM BenM and/or indicated amounts of αCTD. Lane 1 and lane 10 are the PCR marker (NEB) and the 1-kb DNA ladder (Promega), respectively. (B) Predicted model of BenM binding to the *ben* promoters in the presence of muconate. Upon muconate binding, the regulator is suggested to assume activated conformation that is able to bind to Site 2 within the *benM-benA* intergenic region which is located adjacent to the -35 element allowing the protein to interact with αCTD of the RNA polymerase and turn on the transcription from the *benA* promoter.