The FEBS Journal

FEBS PRESS science publishing by scientists

# Crystal structure of the DNA-binding domain of the LysR-type transcriptional regulator CbnR in complex with a DNA fragment of the recognition-binding site in the promoter region

Maharani Pertiwi Koentjoro[1], Naruhiko Adachi[2,3,4], Miki Senda[2], Naoto Ogawa[1,5] and Toshiya Senda[2,3]

1 The United Graduate School of Agricultural Science, Gifu University, Japan
2 Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba, Japan
3 Department of Materials Structure Science, Accelerator Science, The Graduate University of Advanced Studies (Soken-dai), Tsukuba, Japan
4 Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, Tsukuba, Japan
5 Department of Applied Life Sciences, Faculty of Agriculture, Shizuoka University, Japan

**Keywords**
bacteria; CbnR; LysR-type transcriptional regulator; transcription; X-ray crystallography

**Correspondence**
T. Senda, Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), 1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan
Fax: +81 29 879 6179
Tel: +81 29 879 6178
E-mail: toshiya.senda@kek.jp
and
N. Ogawa, Department of Applied Life Sciences, Faculty of Agriculture, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka-Shi, Shizuoka 422-8529, Japan
Fax: +81 54 237 3028
Tel: +81 54 238 4875
E-mail: ogawa.naoto@shizuoka.ac.jp

(Received 1 September 2017, revised 17 November 2017, accepted 8 January 2018)

doi:10.1111/febs.14380

LysR-type transcriptional regulators (LTTRs) are among the most abundant transcriptional regulators in bacteria. CbnR is an LTTR derived from *Cupriavidus necator* (formerly *Alcaligenes eutrophus* or *Ralstonia eutropha*) NH9 and is involved in transcriptional activation of the *cbnABCD* genes encoding chlorocatechol degradative enzymes. CbnR interacts with a *cbnA* promoter region of approximately 60 bp in length that contains the recognition-binding site (RBS) and activation-binding site (ABS). Upon inducer binding, CbnR seems to undergo conformational changes, leading to the activation of the transcription. Since the interaction of an LTTR with RBS is considered to be the first step of the transcriptional activation, the CbnR–RBS interaction is responsible for the selectivity of the promoter to be activated. To understand the sequence selectivity of CbnR, we determined the crystal structure of the DNA-binding domain of CbnR in complex with RBS of the *cbnA* promoter at 2.55 Å resolution. The crystal structure revealed details of the interactions between the DNA-binding domain and the promoter DNA. A comparison with the previously reported crystal structure of the DNA-binding domain of BenM in complex with its cognate RBS showed several differences in the DNA interactions, despite the structural similarity between CbnR and BenM. These differences explain the observed promoter sequence selectivity between CbnR and BenM. Particularly, the difference between Thr33 in CbnR and Ser33 in BenM appears to affect the conformations of neighboring residues, leading to the selective interactions with DNA.

**Database**
Atomic coordinates and structure factors for the DNA-binding domain of *Cupriavidus necator* NH9 CbnR in complex with RBS are available in the Protein Data Bank under the accession code 5XXP.

**Abbreviations**
ABS, activation-binding site; *C. necator*, *Cupriavidus necator*; DBD, DNA-binding domain; HTH, helix–turn–helix; LTTR, LysR-type transcriptional regulator; RBS, recognition-binding site; RD, regulatory domain; wHTH, winged-helix-turn-helix.

## Introduction

The LysR-type transcriptional regulators (LTTRs) [1,2] are one of the best known families of bacterial transcriptional regulators. The members of this family regulate transcription for various genes, such as genes for degradation of aromatic compounds, amino acid biosynthesis, virulence, oxidative stress response, resistance to antibiotics, fixation of carbon dioxide, and nodulation [1–13]. LTTRs, which typically form a homo-tetramer in solution [1,3,10,14–17], recognize a promoter DNA of approximately 60 bp; the promoter DNA is composed of the recognition-binding site (RBS) and the activation-binding site (ABS). Each binding site is recognized by two subunits in the tetrameric LTTR [14,15,17]. Biochemical studies have suggested that binding of an LTTR to the promoter DNA induces a bend of the DNA. The binding, however, is not sufficient for the transcriptional activation [18–21]. LTTRs must bind an inducer molecule for the transcriptional activation. Inducer binding to the regulatory domain (RD) of an LTTR appears to evoke a quaternary structural change, resulting in a shift of the DNA-binding site, particularly in ABS, and a change of the DNA-bend angle [3,15,21–24]. These changes seem to be necessary to recruit RNA polymerase to the promoter site to initiate transcription [25,26].

To reveal the molecular mechanism of the transcriptional activation by LTTRs, the tertiary structures of LTTRs have long been studied. The first structural information on LTTRs was the crystal structure of RD of CysB [27]. RD is the binding site of the inducer molecule; upon inducer binding, RD undergoes a conformational change, the effect of which is transferred to the other portion of the LTTR, resulting in a quaternary structural change for transcriptional activation. The mechanism of the conformation change by the inducer molecule, therefore, is one of the central topics in the study of LTTRs. While CysB forms a tetramer in solution, the RD is a dimer in the crystal. The crystal structure suggests the binding site of the inducer. However, no structural information has been obtained with regard to the mechanism of the DNA recognition and associated DNA bending. After 6 years, our group determined the first crystal structure of a full-length LTTR, CbnR [14]. CbnR, which is derived from *Cupriavidus necator* (formerly *Alcaligenes eutrophus* or *Ralstonia eutropha*) NH9, is an LTTR that regulates the gene expression of the *cbnABCD* genes [23]. The crystal structure revealed that the full-length CbnR (CbnR(-full)) forms a dimer of (protomeric) dimers in the crystal. The subunit of CbnR is composed of an RD and a DNA-binding domain (DBD), and the two domains are connected by a long helix, which is the interface of two subunits in the protomeric dimer [14]. DBD has a winged-helix–turn–helix (wHTH) motif. Interestingly, two subunits in the protomeric dimer adopt different conformations, i.e., the compact and extended forms. The tetrameric CbnR is composed of the two protomeric dimers. Interestingly, four DBDs are located at the bottom of the tetrameric CbnR and arranged in a V-shape, which well explains the DNA bending and recognition mode of the RBS and ABS [14]. One protomeric dimer in tetrameric CbnR recognizes RBS and the other one recognizes ABS (Fig. 1A).

Another critical research topic in terms of the structure of LTTRs is the question of the molecular mechanisms underlying the recognition of the promoter DNA and induction of transcriptional activation via inducer binding. The conformational changes of RD upon inducer binding have been structurally analyzed with AphB [24], BenM [28], DntR [29], OxyR [10,30], PcpR [31] and TsaR [17]. While several models of transcriptional activation by LTTRs with large structural changes have been proposed [15,17,24], all structural changes are likely to arise from a local conformational change of RD upon inducer binding. However, no crystallographic information has been obtained showing the overall structural changes of LTTRs, including the promoter DNA.

Interactions between an LTTR and RBS were revealed by the crystal structure of the DBD of BenM (hereafter BenM(DBD)) in complex with its cognate RBS in the *benA* promoter (*benA*–RBS) [32]. This crystal structure demonstrated details of the interaction between the dimeric BenM(DBD) and a *benA*–RBS sequence of 25 bp. Ala28, Gln29, Pro30, Arg34, and Arg53 in the wHTH motif of BenM(DBD) specifically interact with the *benA*–RBS, TAAAA-ATACT-CC ATA-GGTAT-TTTAT (specifically recognized bases are underlined). While the crystal structure of the BenM(DBD)–RBS complex revealed details of the interaction, the mechanism responsible for the selection of a specific sequence remains elusive; RBS sequences for BenM and CbnR are similar to each other, but the specificity of the recognition of cognate RBS by each LTTR is not known.

In this study, we focused on the mechanism of the DNA binding of two LTTRs, CbnR and BenM. While these two proteins have a similar primary structure, sharing 48% amino acid identity in their DBDs, the CbnR(DBD) only interacts with its cognate promoter DNA. To understand the detailed mechanism of the specific interaction between CbnR(DBD) and its promoter DNA, we determined the crystal structure of the CbnR(DBD)–RBS complex.

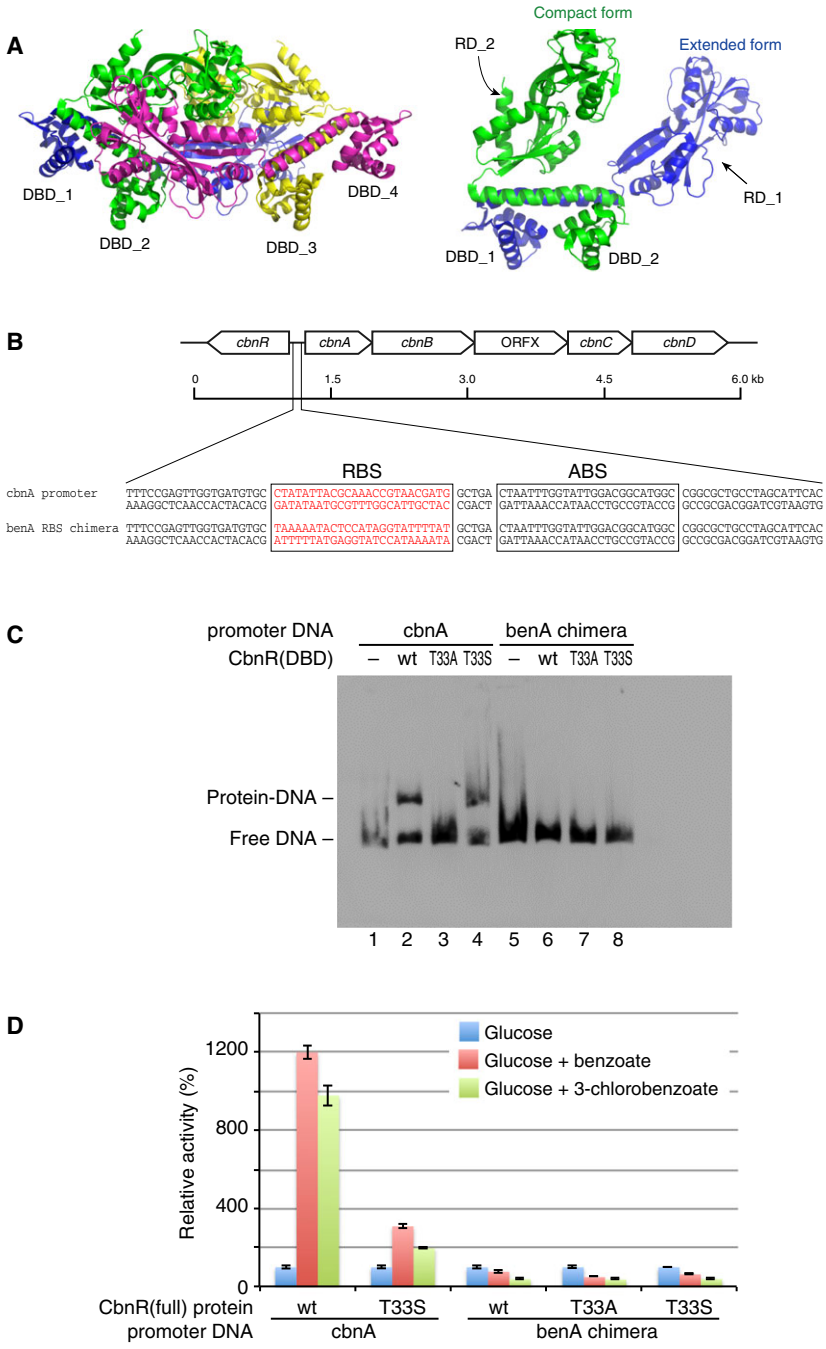


**Fig. 1.** DNA-binding activity of CbnR. (A) Overall structure of tetrameric CbnR (left panel) and a dimer of CbnR (right panel) containing compact (green) and extended (blue) forms. (B) Schematic drawing of the gene-coding regions of *cbnRAB*, *ORFX*, and *cbnCD*. The DNA sequence of the *cbnA* promoter region is shown. RBS (red) and ABS (black) are boxed. The DNA sequence of the *benA* promoter region is also shown. (C) The EMSA assay with CbnR(DBD) and its mutant proteins. The assays with the *cbnA* promoter (lanes 1-4) and the *benA* chimera promoter (lanes 5–8) are shown. (D) The LacZ assay with CbnR(full) and its mutant proteins. Activities were measured as relative activity, where that of wild-type CbnR(full) without inducer was calculated as 100%. Benzoic acid and 3-chlorobenzoic acid are transformed into *cis*, *cis*-muconate and 2-chloro-*cis*, *cis*-muconate, respectively, in the medium. Error bars represent means ± SD ($n$ = 9).

## Results

### *In vitro* sequence specificity assay and transcriptional activity assay

The sequence specificity of CbnR was examined using an electrophoretic mobility shift assay (EMSA) with *cbnA* and chimeric *benA* promoters. The intergenic region of *cbnR* and *cbnA* genes containing the *cbnA* promoter spans 149 bp and includes the *cbnA*–RBS and ABS sequences. The chimeric *benA* promoter was prepared by replacing the *cbnA*–RBS sequence of the *cbnA* promoter with the *benA*–RBS sequence (Fig. 1B). The EMSA using CbnR(DBD) revealed a specific interaction between CbnR(DBD) and *cbnA* promoter DNA (lane 2 in Fig. 1C); CbnR(DBD) showed no interaction with the chimeric *benA*

promoter (lane 6 in Fig. 1C). While the ABS sequence of the chimeric *benA* promoter is derived from that of the *cbnA* promoter, CbnR(DBD) showed no significant binding to the chimeric *benA* promoter (lane 6 in Fig. 1C). This fact suggests that the affinity between CbnR(DBD) and *benA*–RBS was weak under the present conditions.

These results are consistent with those of the LacZ assay of CbnR (Fig. 1D). The wild-type CbnR showed 12- and 10-fold transcription activation of the *cbnA* promoter in the presence of benzoate and 3-chlorobenzoate, respectively. *cis*, *cis*-Muconate and 2-chloro-*cis*, *cis*-muconate converted from benzoate and 3-chlorobenzoate in the bacterial cells, respectively, served as the actual inducers for CbnR [23]. However, no transcription activation of the *benA* chimera promoter was observed with these two inducer molecules.

## Crystal structure of the CbnR(DBD)–RBS complex

To understand the molecular mechanism of the sequence specificity of CbnR, the crystal structure of the CbnR(DBD)–RBS complex was determined at 2.55 Å resolution by the molecular replacement method (Table 1). The amino acid sequence of CbnR (DBD) and DNA sequence of *cbnA*–RBS are shown in Fig. 2A,B, respectively. The crystal structure shows that the CbnR(DBD) dimer binds the *cbnA*–RBS (Fig. 2C). The structure of the dimeric CbnR(DBD) is essentially the same as that in the crystal structure of the full-length CbnR (PDB ID: 1IXC), with root mean square deviation (rms deviation) of 1.06 Å (165 Cα atoms) (Fig. 2D). While the tertiary structure of CbnR (DBD) is well superposed on that of the DBD portion of the DNA-free form, the winged region of the wHTH motif in CbnR(DBD) showed a large shift, approximately 4 Å, to interact with a minor groove of the *cbnA*–RBS (red arrows in Fig. 2D).

The structure of the CbnR(DBD)–RBS complex was similar to that of the BenM(DBD)–RBS complex, with an rms deviation of 1.2 Å (172 Cα atoms) (Fig. 2E). As observed for the BenM(DBD)–RBS complex, the bound DNA in the CbnR(DBD)–RBS complex is a B-form and curved by approximately 30 degrees around the central portion of the bound DNA. The helix α3 in the wHTH motif is embedded in a major groove of the DNA and helix α2 also interacts with the same major group. The winged motif, showing a significant shift in the DNA complex, interacts with a minor groove. These interactions generate an interacting surface area of 780 $Å^2$ between one DBD and DNA, and thus the dimeric CbnR(DBD) has an interacting surface area of approximately 1560 $Å^2$ with the *cbnA*–RBS sequence. No significant interactions were found between the two wHTH motifs except a π–π interaction of Phe3; a corresponding π–π interaction was also observed in the DNA-free form of CbnR (full-length).

**Table 1.** Crystallographic data summary of the CbnR(DBD)–RBS complex.

| | |
|---|---|
| Data collection[a] | |
| Space group | *C*2 |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 186.5, 28.7, 73.2 |
| α, β, γ (°) | 90, 111.7, 90 |
| Resolution (Å) | 67.99–2.55 (2.69–2.55)[b] |
| $R_{merge}$ | 0.059 (0.439) |
| $\langle I/\sigma(I)\rangle$ | 13.5 (3.3) |
| Completeness (%) | 97.1 (95.5) |
| Redundancy | 3.3 (3.3) |
| Crystallographic refinement | |
| Resolution (Å) | 67.99–2.55 |
| No. reflections | 11 461 |
| $R_{work}/R_{free}$ | 0.219/0.263 |
| No. atoms | |
| Protein | 1316 |
| DNA | 1019 |
| Water | 2 |
| B-factor ($Å^2$) | |
| Protein | 48.5 |
| DNA | 69.2 |
| Water | 44.5 |
| RMS deviations | |
| Bond lengths (Å) | 0.008 |
| Bond angles (°) | 1.366 |
| PDB code | 5XXP |

[a] Data collected from one crystal. [b] Values in parentheses are for the highest resolution shell.

The structure of the CbnR(DBD)–RBS complex has a pseudo–twofold axis. Since the DNA sequence of *cbnA*–RBS is a palindromic one, the two DBD domains in the complex can be well superposed with an rms deviation of 0.20 Å (85 Cα atoms). Not only the main chain atoms but also the side chain atoms are well superposed between the two subunits of CbnR (DBD) (rms deviation = 0.82 Å for all atoms in 85 residues), showing that the palindromic sequences in *cbnA*–RBS are recognized in nearly the same manner by the two DBDs.

In the present crystal structure, we found only two water molecules in the electron density map under our criterion for water picking (see the Materials and methods section). Despite this relatively low number of water molecules at the interface of CbnR(DBD) and RBS, we cannot exclude the possibility that more water molecules would be found in a crystal structure of the CbnR(DBD)–RBS complex with higher resolution.

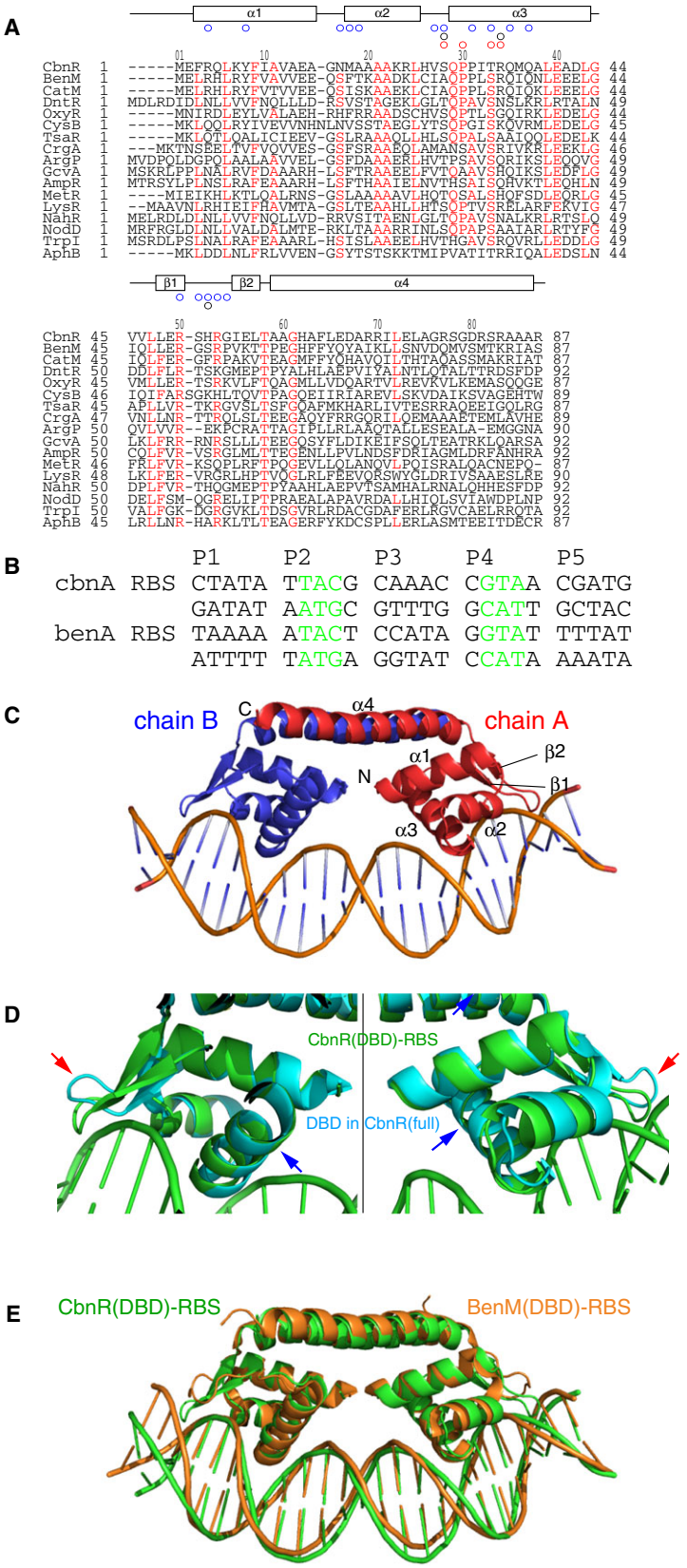


**Fig. 2.** Crystal structure of the CbnR(DBD)-RBS complex. (A) Amino acid sequence alignment of the DNA-binding domain of LTTR family proteins. The secondary structures of CbnR are shown above. In CbnR(DBD), the residues involved in sugar-phosphate backbone recognition are shown in blue circles. The residues involved in nonselective or selective interaction are shown in black and red circles, respectively. (B) DNA sequences of *cbnA* RBS and *benA* RBS. The bases conserved between *cbnA* RBS and *benA* RBS are shown in green. The sequences can be divided into five parts, P1 to P5. (C) The crystal structure of the CbnR(DBD)-RBS complex. CbnR(DBD) binds to DNA as a dimer. The subunits A and B in the dimer are colored in red and blue, respectively. The N- and C-terminal ends and the secondary structures are labeled. (D) Structural comparison of the DNA-binding domains of CbnR(DBD)-RBS (green) and CbnR(full) (cyan) (PDB code: 1IZ1). Significant structural differences are observed in the winged region (red arrows) and α3 (blue arrows). (E) Structural comparison of the CbnR(DBD)-RBS (green) and BenM(DBD)-RBS (orange) (PDB code: 4IHT) complexes.

### Sugar–phosphate backbone recognition of CbnR

To understand the RBS-recognition mechanism of CbnR(DBD), a list of amino acid residues interacting with RBS was prepared. Since only a limited number of water molecules were identified in the present crystal structure, it was rather difficult to prepare a complete list of the protein–DNA interactions including water molecules. All residues in a range of 4 Å distance from DNA atoms were therefore listed (Table S1). In addition, a similar list was prepared for the BenM(DBD)–RBS complex (PDB ID: 4IHT) (Table S2). Then, based on these lists, the interactions between CbnR(DBD) and RBS were analyzed.

Initially, interactions between CbnR(DBD) and the sugar–phosphate backbone of DNA were analyzed. Since the chemical structure of the sugar–phosphate backbone is shared in all DNAs, these interactions mainly contribute to the binding strength between protein and DNA rather than the specificity for the promoter sequences. In the crystal, the CbnR(DBD) dimer interacts with sugar–phosphate moieties of 18 bp out of the 25 bp RBS sequence (blue in Fig. 3A). The interactions are nearly symmetrical as described above. This binding mode is similar to that of the BenM(DBD)–RBS interactions (blue in Fig. 3B); 18-bp sugar–phosphate moieties from the 25 bp RBS sequence were recognized by the BenM (DBD) dimer in a nearly symmetric manner.

### Base–moiety recognition of CbnR

Next, interactions with base moieties were analyzed. Sequence-specific interactions can occur through a combination of direct-readout (e.g., base-moiety recognition) and indirect-readout (e.g., phosphate-backbone recognition) mechanisms [33,34]. In this study, we primarily focused on direct-readout mechanisms. Interactions with base moieties were categorized into two types – i.e., nonselective and selective interactions – based on earlier studies [35] (Fig. 3C). The protein–base interactions could be further categorized by the location of the interaction, major and minor grooves. In the crystal structures of both the CbnR(DBD)–RBS and BenM(DBD)–RBS complexes, the 25-bp RBS sequences can be divided into five parts, P1 to P5, each of which has 5 bp (Fig. 3A). The base pairs of P1 and P5, which are located at both ends of the RBS, interact with the winged motif of CbnR(DBD) using the minor groove side. The central 5 bp, P3, have fewer interactions with the proteins. Base pairs in the remaining two parts, P2 and P4, interact intensively with the helix–turn–helix motif of the DBDs using major grooves. Of the five parts, the DNA sequences in P2 and P4 are best conserved between the *cbnA* and *benA* promoters; the central 3 bp are identical and the two flanking bases show differences (green in Fig. 3A, B). It is intriguing that the best-conserved portions have several selective interactions, which seem to be responsible for the sequence specificity.

In the CbnR(DBD)–RBS complex, nonselective interactions were found in 4 bp in the major grooves (P2 and P4) and 2 bp in the minor groove (P5) (black in Fig. 3A). On the other hand, in the BenM(DBD)–RBS complex, nonselective interactions were found in 8 bp in the major groove and 4 bp in the minor groove (black in Fig. 3B). In P2, P4, and P5, the positions of base moieties with nonselective interactions are nearly the same in CbnR(DBD) and BenM(DBD). In the P1 site, however, BenM(DBD) has more nonselective interactions with DNA.

We next analyzed selective interactions between DBD and DNA (red in Fig. 3A,B). In the CbnR (DBD)–RBS complex, helix α3 takes part in selective interactions with 10 bp (TTACG-CAAAC-CGTAA) in the major grooves. Residues Ser28, Pro30, Thr33, and Arg34 selectively interact with base moieties. In the BenM(DBD)–RBS complex, helix α3 contributes to the selective interactions with 10 bp in the major grooves (ATACT-CCATA-GGTAT-T); Ala28, Gln29, Pro30, Pro31, and Arg34 are all involved in the selective interactions (the protein–DNA interactions are summarized in Fig. 3D). While no selective interactions were found for the base moieties of the minor grooves in CbnR(DBD), Arg53 of BenM(DBD) makes a selective interaction with a base pair in a minor groove. Selective interactions in the CbnR(DBD)–RBS complex have a twofold symmetry probably due to a palindromic structure of the *cbnA* promoter. On the other hand, selective interactions in BenM(DBD) were nonsymmetric, probably reflecting the nonsymmetric structure of the *benA* promoter sequence.

### Mechanism of the sequence selectivity between the cbnA and benA promoters

Next, the sequence specificity of CbnR(DBD) was analyzed. The comparison of DNA sequences between the *cbnA* and *benA* promoters showed that TAC and GTA in P2 and P4, respectively, are conserved between the two promoters (green in Fig. 3A, B). Therefore, the interactions with these base pairs cannot contribute to the specificities. To understand the sequence specificity of CbnR(DBD) and BenM (DBD), it is necessary to focus on the DBD residues interacting with nonconserved base pairs of the two

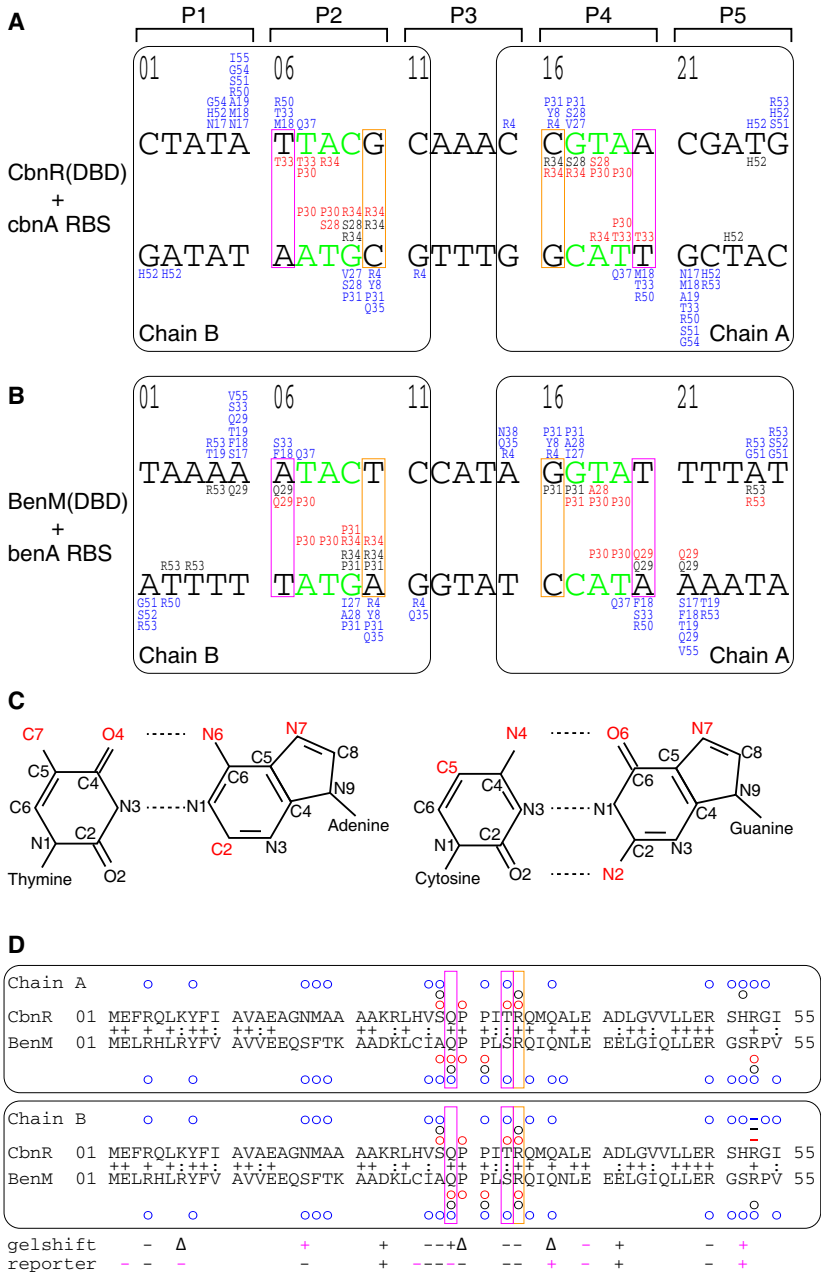


**Fig. 3.** Protein-nucleotide interactions of the CbnR(DBD)-RBS and BenM(DBD)-RBS complexes. (A, B) Diagrams showing the interactions between CbnR(DBD) and RBS (A) and BenM(DBD) and RBS (B). The residues involved in sugar-phosphate backbone recognition are shown in blue. The residues involved in nonselective or selective interaction with bases are shown in black and red, respectively. (C) Schematic picture of the base pairs of DNA. The atoms involved in nonselective and selective interaction are shown in black and red, respectively. (D) Sequence alignment of CbnR(DBD) and BenM(DBD). The residues involved in sugar-phosphate backbone recognition are shown in blue circles. The residues involved in the nonselective or selective interaction are shown in black and red circles, respectively. Magenta and orange boxes indicate the residues that recognize nonconserved base pairs of the two promoters. The results of EMSA and LacZ assays from a previous report [36] are summarized at the bottom of the sequence alignment; +, Δ, and – represent binding/activation, weak-binding/weak-activation, and no-binding/no-activation, respectively. The results that cannot be explained by the crystal structure alone are colored in magenta.

promoters, T6/T6′, G10/G10′, C16/C16′, and A20/A20′ in *cbnA* RBS, and A6/A6′, T10/C10′, G16/A16′, and T20/T20′ in *benA* RBS. In CbnR(DBD), Thr33 and Arg34 recognize T6/T6′ and C16′/C16, respectively (Fig. 4A and upper panels in Fig. 4B). These interactions are observed in both subunits of CbnR (DBD), because conformations of the sidechains of the two subunits are nearly the same in CbnR(DBD). In BenM(DBD), Gln29, Arg34, and Arg53 undergo specific interactions with A6/A6′/A5′, A16′, and A24, respectively (Fig. 4A and lower panels in Fig. 4B). These interactions, however, are slightly different between the two subunits of BenM(DBD); Arg34 in chain A and Arg53 in chain B were not involved in these interactions. Of the four residues that interact with nonconserved base pairs (Gln29, Thr33/Ser33, Arg34, Arg53), Gln29, Arg34, and Arg53 are conserved both in CbnR and BenM (Fig. 2A), implying that these residues cannot play a primary role in the promoter sequence selectivity. Therefore, Thr33 in CbnR and Ser33 in BenM could contribute to the selective interactions of CbnR and BenM.

In the crystal structure of the BenM(DBD)–RBS complex, Ser33 interacts with Gln29, which in turn forms hydrogen bonds with A6/A6′/A5′ (lower panels in Fig. 4B). The interaction with Ser33 makes the side chain of Gln29 move toward A6/A6′/A5′ (Fig. 4C). On the other hand, Gln29 in the CbnR(DBD)–RBS complex adopts a conformation different from that of BenM(DBD)–RBS due to the lack of a hydrogen bond with Thr33 (Fig. 4D, Fig. S1). This result is consistent with our earlier mutational analysis, in which Gln29 in CbnR was shown to have no impact on the DNA-binding activity; Gln29Ala retained the *cbnA* promoter-binding activity [36].

The side chain methyl group of Thr33 in CbnR (DBD) makes a hydrophobic interaction with the C7 atom of T6/T6′ in a selective manner. To validate the functional role of Thr33, two mutants of CbnR(DBD), Thr33Ala and Thr33Ser, were prepared and examined for their DNA-binding activity. The EMSA showed that Thr33Ala lacked the DNA-binding activity (lane 3 in Fig. 1C). Thr33Ser significantly reduced the DNA-binding activity (lane 4 in Fig. 1C). These results suggested that the hydrophobic interaction between Thr33 and T6/T6′ is of critical importance to the *cbnA*-promoter binding. While the Thr33Ser mutant had been expected to show *benA*-promoter binding activity, no such activity was observed (lane 8 in Fig. 1C). To make a *benA*-binding surface on CbnR(DBD), the effects of other neighboring residues should be required. Moreover, other effects such as the indirect-readout mechanism [33,34] and interactions between the winged motif and nucleotides in the minor groove should be considered for a precise discussion of the RBS recognition.

These results of EMSA are consistent with the LacZ assay using the full-length CbnR. The wild-type full-length CbnR showed 12-fold transcription activation by the inducer, *cis*, *cis*-muconate, which is derived from benzoate. The Thr33Ser mutant with a reduced affinity for the promoter DNA showed only threefold activation by the inducer (Fig. 1D). Thr33Ala, which lacks the promoter-binding activity, could not activate the transcription of the *cbnA* promoter [36]. Similarly, wild-type CbnR, Thr33Ala, and Thr33Ser all failed to bind to the *benA* chimera promoter and could not activate its transcription (Fig. 1D).

## Discussion

In this study, we determined the crystal structure of the CbnR(DBD)–RBS complex to analyze the specificity difference between CbnR and BenM. Our structural analysis and comparison with the crystal structure with that of the BenM(DBD)–RBS complex suggested some critical interactions for the specific interaction between CbnR(DBD) and RBS of the *cbnA* promoter. Moreover, the crystal structure of the CbnR (DBD)–RBS could reasonably explain the results of our earlier EMSA with Gln29 mutant proteins, which could not be reasonably explained based on the crystal structure of BenM(DBD)–RBS.

We would like to further analyze the interactions between CbnR(DBD) and RBS based on our earlier mutational experiments on CbnR [36]. In our earlier experiments, a series of single amino acid substitutions were introduced into the full length of CbnR (Fig. 3D). The crystal structure of the CbnR(DBD)–RBS complex showed that our prediction of amino acids interacting with DNA was essentially correct. Next, we analyzed the results of our EMSA using the present crystal structure. Of the 15 mutants, the EMSA results for 12 mutants could be reasonably explained by a simple rule based on the distance from the bound DNA (black in Fig. 3D). Alanine substitutions for residues within 4 Å distance from the bound DNA reduced the DNA-binding activity (Arg4Ala, Val27Ala, Ser28Ala, Pro30Ala, Thr33Ala, Arg34Ala, Gln37Ala, and Arg50Ala), and those located more than 4 Å from the bound DNA had smaller effects on the DNA-binding activity (Lys7Ala, Lys23Ala, Gln29Ala, and Asp42Ala). While the results of EMSA for the remaining three mutants (Asn17Ala, Glu40Ala, and His52Ala) did not follow the simple rule described above, they could also be explained with the crystal structure (magenta in Fig. 3D). Asn17 and His52 are located within 4 Å from the DNA, but alanine substitutions for these two residues did not affect the DNA-binding activity. Since Asn17 interacts selectively with DNA via the Cβ atom, an alanine substitution that retains the Cβ atom would not affect the DNA-binding activity significantly. Alanine substitution of His52 results in a loss of the interactions with DNA. However, Arg53 could take part in an interaction with DNA, as observed in the BenM(DBD)–DNA complex. Finally, Glu40 appears to have an indirect effect on DNA binding. Glu40 interacts with Arg50 and stabilizes the Arg50 conformation to interact with DNA. The substitution of Glu40 with alanine, therefore, destabilizes the conformation of Arg50, resulting in a reduction of the affinity for DNA.

Interestingly, as described in our earlier report, the affinity for RBS does not directly relate to the transcriptional activation activity of CbnR (or its mutant) [36], suggesting a multistep mechanism of the transcription activation by LTTRs [15,17,30]. In particular, the results of mutational analysis on Gln29 are enigmatic. As shown in Fig. 4B, this residue is not directly

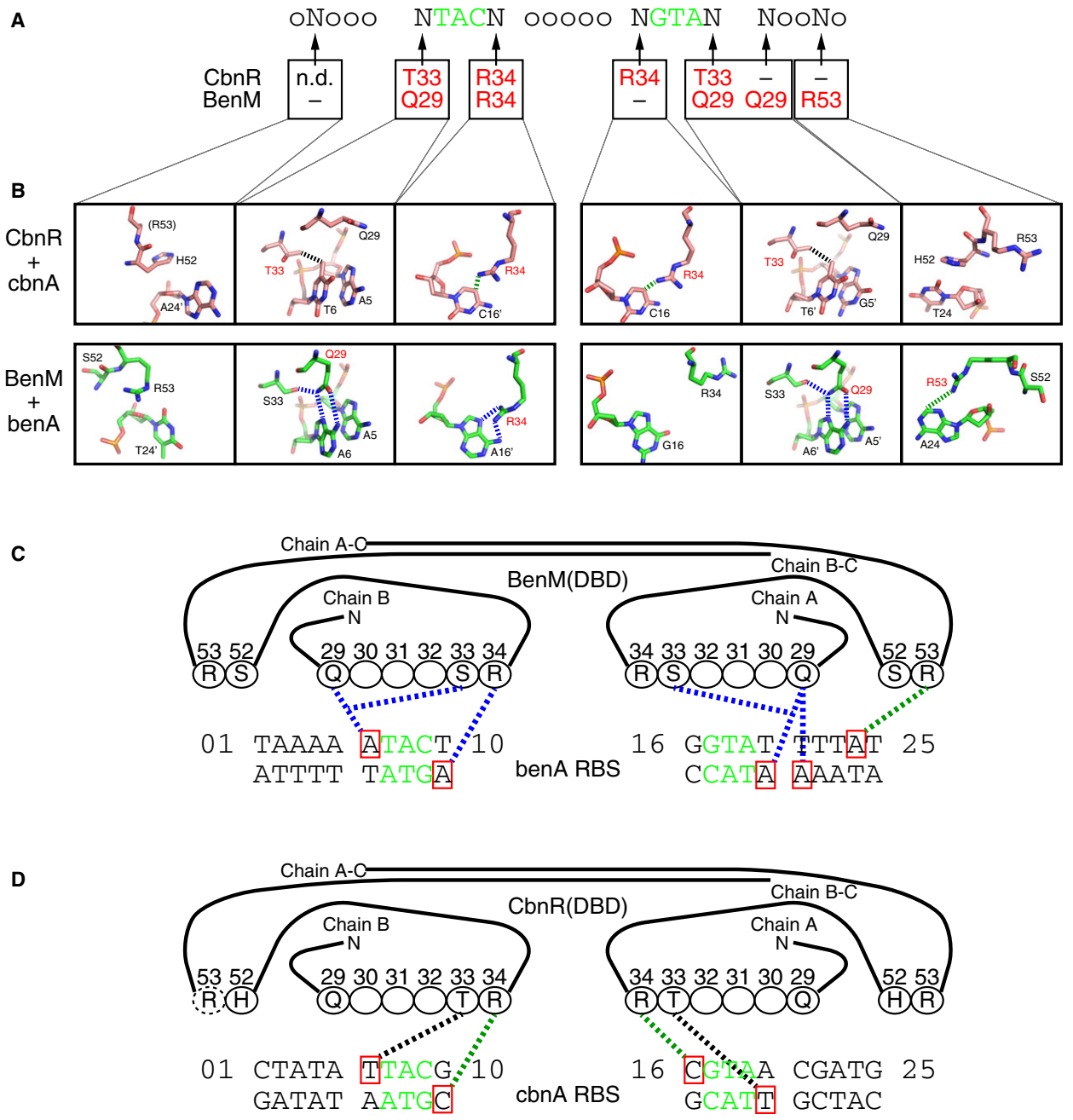


**Fig. 4.** Detail of the interactions of CbnR(DBD)-RBS and BenM(DBD)-RBS. (A) Summary of the interactions between the residues in CbnR (DBD)/BenM(DBD) and DNA. Bases conserved between the *cbnA* and *benA* promoters are shown in green. (B) Close-up view of the residues in CbnR(DBD)/BenM(DBD) that are involved in the protein–DNA interaction. The residues involved in selective interaction are labeled in red. Hydrogen bonds and Van der Waals interactions are shown as blue and black dotted lines, respectively. Protein and DNA atoms within a distance of 4 Å are connected by green dotted lines. (C) Schematic drawing of the residues important for the selective interaction between BenM(DBD) and RBS in the *benA* promoter. Hydrogen bonds and Van der Waals interactions are shown as blue and black dotted lines, respectively. Protein and DNA atoms within a distance of 4 Å are connected by green dotted lines. (D) Schematic drawing of the residues important for selective interaction between CbnR(DBD) and RBS in the *cbnA* promoter. Color usage is the same as in panel (C).

involved in the interaction with DNA, and its mutant proteins exhibit only limited differences by EMSA [36]. However, the inducer was unable to activate transcription of most of the mutant proteins of Gln29. It is, thus, reasonable to predict that Gln29 in CbnR plays a critical role other than in the interaction with RBS; Gln29, which is one of the well-conserved amino acids in DBD of various LTTRs, may contribute to the interactions with ABS or other factors including RNA polymerase. To address this possibility, it will be necessary to analyze not only the interaction between CbnR(DBD) and ABS but also the crystal structure of the CbnR(full) in complex with the *cbnA* promoter with and without the inducer molecule. These studies would provide deep insight into the mechanism of the transcriptional activation by CbnR.

## Materials and methods

### Vector cloning and subcloning

Synthetic DNA encoding the CbnR(DBD) gene from the *C. necator* NH9 [23], which encodes an 87 amino acid protein, was purchased from Eurofins (Luxembourg). The synthesized CbnR(DBD) gene, whose codons were optimized for expression in *Escherichia coli*, was subcloned into a pET47b vector (Novagen, Darmstadt, Germany) by using *Nde*I and *Not*I restriction sites. pET47b–DBD–CbnRHis was designed to express the CbnR(DBD) protein with a TEV protease recognition site (ENLYFQG) and a His6–tag sequence (AHHHHHH).

### Overexpression and purification of CbnR(DBD)

The pET47b–DBD–CbnRHis was introduced into *E. coli* BL21(DE3) and grown at 37 °C in 200 mL Luria Bertani medium supplemented with 2% (w/v) glucose and 50 $\mu g \cdot mL^{-1}$ kanamycin (Kan). The culture medium was inoculated into 2 L of fresh medium. The culture was grown at 27 °C with constant shaking at 110 rpm until the optical density ($OD_{600}$) reached 0.6–0.8 and then cooled at 16 °C. The CbnR(DBD) expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and culturing overnight at 16 °C. The bacterial cells were harvested by centrifugation (3993 ***g*** for 10 min at 4 °C) and then the cell pellet was resuspended in 20 mL of buffer A (20 mM Tris–HCl; 500 mM NaCl; and 5 mM imidazole, pH 7.9). The cells were lysed by sonication using 1 min pulses and cooling on ice for 3 min (1 min on/3 min off) for three cycles. In addition, the cell debris was removed by centrifugation at 20 400 ***g*** for 30 min.

CbnR(DBD) proteins in the soluble fraction were purified by standard affinity chromatography using a HisTrap HP column (GE Healthcare, Chicago, IL, USA). The resin was washed three times in series with buffer A, B (20 mM Tris–HCl, 500 mM NaCl, and 1.0 M Imidazole, pH 7.9), and C (20 mM Tris–HCl, 500 mM NaCl, and 50 mM Imidazole, pH 7.9). The CbnR(DBD) was eluted with a linear gradient of imidazole (50 mM to 1 M). The peak fractions containing the CbnR(DBD) were collected and the concentration of the protein was determined using a NanoDrop ND-1000 (Thermo Scientific, Waltham, MA, USA) at 280 nm with an absorption coefficient of 2980 $M^{-1}$ $cm^{-1}$. The purity of the protein was analyzed using SDS/PAGE (NuPAGE 4–12% Bis–Tris) with Coomassie Brilliant Blue G–250 (Wako, Osaka, Japan) staining.

### Preparation and crystallization of the CbnR (DBD)–RBS complex

Oligonucleotides with the RBS sequences were synthesized for the co-crystallization experiments (Eurofins Genomic, Ebersberg, Germany). DNA duplexes were prepared by annealing complementary oligonucleotides (5′-CTATA TTACG CAAAC CGTAA CGATG-3′ and 5′-CATCG TTACG GTTTG CGTAA TATAG-3′) (0.1 mM each), by first heating the solution to 95 °C for 10 min and then cooling slowly to room temperature in a heat block. For preparation of the CbnR(DBD)–RBS complex, the double-stranded DNA and the CbnR(DBD) protein were mixed and incubated in a molar ratio of 1 : 2, respectively, for 30 min on ice. The CbnR(DBD)–RBS complex was then concentrated using Amicon Ultracel–10K (Merck Millipore, Burlington, MA, USA) in the range of 7–10 $mg \cdot mL^{-1}$ before the crystallization-screening set-up.

Crystallization conditions for the CbnR(DBD)–RBS complex were screened with Crystal Screen 1 and 2 (Hampton Research, Aliso Viejo, CA, USA); Wizard I and II (Molecular Dimensions, Suffolk, UK); Cryo I and II (Hampton Research); Footprint/MembFac (Hampton Research); PEG-sII (Qiagen, Hilden, Germany); Index (Hampton Research); PEGIon/PEGIon2 (Hampton Research); and a Protein complex suite (Qiagen) using a Protein Crystallization System (PXS) at the Structural Biology Research Center, High Energy Accelerator Research Organization (KEK) (Tsukuba, Japan) [37]. The sitting-drop vapor-diffusion method was used with crystallization drops consisting of 0.5 μL protein solutions (7 $mg \cdot mL^{-1}$ in DNA-binding buffer) and 0.5 μL precipitants at 20 °C. Crystals were apparent after more than 1 week. Various crystals obtained in the initial crystallization screening were selected based on snapshots of the X-ray diffraction images. Crystals that diffracted to the highest resolution (3.1 Å) were further optimized by changing the cryoprotectant conditions. The final conditions of the crystallization solution were 32.5% (w/v) PEG 4000, 0.1 M Tris–HCl pH 8.5, 0.2 M sodium acetate. Crystals of the CbnR(DBD)–RBS complex were obtained at 20 °C by mixing 1 μL of the CbnR(DBD)–RBS complex with 1 μL of crystallization solution. The maximum crystal growth was

achieved after 13 days of incubation. Before diffraction data collection, crystals were cryoprotected using 30% (v/v) glycerol. The crystals were flash frozen in a liquid nitrogen stream and stored in a UniPuck cassette for X-ray diffraction experiments.

### X-ray data collection, processing, and structure determination

-ray diffraction data were collected at 95 K using an ADSC Quantum 270 CCD detector on beamline BL-17A of the Photon Factory, High Energy Accelerator Research Organization (KEK) (Tsukuba, Japan). Diffraction data were processed and scaled by XDS and XSCALE, respectively [38]. The phases were determined by the molecular replacement method using the program BALBES [39]. The amino acid sequence of the CbnR(DBD) and the diffraction data of the Cbn(DBD)-RBS complex were provided for BALBES. BALBES automatically selected the coordinates of full-length BenM (the R156H mutant, PDB entry 3K1M) [16] and utilized DBD (residues 1–87) of the BenM mutant as an initial model. After obtaining the initial coordinates of the CbnR(DBD) dimer, the model was refined at 2.55 Å resolution with PHENIX.refine [40] and modified with Coot [41]. The model of RBS was constructed into a 2*m*Fo-*D*Fc map using Coot [41]. After several crystallographic refinement cycles, the crytallographic refinement at 2.55 Å resolution was completed with REFMAC5 [42] in the CCP4 suite [43]. Water molecules in the crystal structure were picked using a *m*Fo-*D*Fc map. When a difference density indicated a spherical shape at the 3 σ contour level, the density was considered to identify a water molecule. The crystal structure was refined to $R_{free}/R_{work}$ values of 0.263 and 0.219, respectively.

### EMSA

EMSA was conducted as described previously [36]. To prepare the chimeric *benA* promoter, the 25 bp RBS sequence of the *cbnA* promoter in the plasmid pBLcbnR-AB′ [36] was replaced with a corresponding 25-bp RBS sequence of BenM. The resultant plasmid was used as a template to generate a DNA fragment for the EMSA assay by polymerase chain reaction amplification and was also used to further produce plasmid constructs for reporter analysis (below). The conditions of electrophoresis to separate free DNA and the complex of CbnR-DBD and the promoter DNA were modified to 240 min and 80 V in a cold room.

### *In vivo* determination of the ability to activate transcription (LacZ analysis)

Plasmid construction and LacZ assays were carried out as described previously [36]. To prepare a pQF50-based plasmid construct containing either the mutant *cbnR* gene or the *benA* chimeric promoter, the appropriate fragment of pBLcbnR-AB′ or pBLcbnRmutant-AB′ was replaced with the fragment of interest and further cloned into pQF50 as an *Xba*I-*Hin*dII fragment. The assays were carried out essentially as described by Miller [44].

## Acknowledgements

We thank Simon Miller for provided TEV protease, Ryuichi Kato for assistance with crystallization screening using a crystallization robot, and Yusuke Yamada for assistance with the data collection. This work was partially supported by PDIS and Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS) from AMED, Japan.

## Conflict of interest

The authors declare that they have no conflict of interest.

## Author contributions

MPK and NA performed the protein purification. MPK and MS performed the crystallographic study. MPK performed the biochemical assay. NO and TS conceived the study. NA, NO, and TS wrote the manuscript.

## References

1 Schell MA (1993) Molecular biology of the LysR family of transcriptional regulators. *Annu Rev Microbiol* **47**, 597–626.
2 Maddocks SE & Oyston PCF (2008) Structure and function of the LysR–type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**, 3609–3623.
3 Tropel D & van der Meer JR (2004) Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol Mol Biol Rev* **68**, 474–500.
4 Parsek MR, McFall SM, Shinabarger DL & Chakrabarty AM (1994) Interaction of two LysR–type regulatory proteins CatR and ClcR with heterologous promoters: functional and evolutionary implications. *Proc Natl Acad Sci U S A* **91**, 12393–12397.
5 Lochowska A, Iwanicka-Nowicka R, Plochocka D & Hryniewicz MM (2001) Functional dissection of the LysR–type CysB transcriptional regulator. *J Biol Chem* **276**, 2098–2107.
6 Sperandio B, Gautier C, Pons N, Ehrlich DS, Renault P & Guédon E (2010) Three paralogous LysR–type transcriptional regulators control sulfur amino acid supply in *Streptococcus mutans*. *J Bacteriol* **192**, 3464–3473.

7 Deghmane AE, Petit S, Topilko A, Pereira Y, Giorgini D, Larribe M & Taha MK (2000) Intimate adhesion of *Neisseria meningitidis* to human epithelial cells is under the control of the *crgA* gene, a novel LysR–type transcriptional regulator. *EMBO J* **19**, 1068–1078.
8 Heroven AK & Dersch P (2006) RovM, a novel LysR–type regulator of the virulence activator gene *rovA*, controls cell invasion, virulence and motility of *Yersinia pseudotuberculosis*. *Mol Microbiol* **62**, 1469–1483.
9 Mandal RS, Ta A, Sinha R, Theeya N, Ghosh A, Tasneem M, Bhunia A, Koley H & Das S (2016) Ribavirin suppresses bacterial virulence by targeting LysR–type transcriptional regulators. *Sci Rep* **6**, 39454.
10 Choi HJ, Kim SJ, Mukhopadhyay P, Sayeon C, Woo JR, Storz G & Ryu SE (2001) Structural basis of the redox switch in the OxyR transcription factor. *Cell* **105**, 103–113.
11 Vadlamani G, Thomas MD, Patel TR, Donald LJ, Reeve TM, Stetefeld J, Standing KG, Vocadlo DJ & Mark BL (2015) The β-lactamase gene regulator AmpR is a tetramer that recognizes and binds the D-Ala-D-Ala motif of its repressor UDP-*N*-acetylmuramic acid (MurNAc)-pentapeptide. *J Biol Chem* **290**, 2630–2643.
12 Dangel AW & Tabita FR (2015) CbbR, the master regulator for microbial carbon dioxide fixation. *J Bacteriol* **197**, 3488–3498.
13 Peck MC, Fisher RF, Bliss R & Long SR (2013) Isolation and characterization of mutant *Sinorhizobium meliloti* NodD1 proteins with altered responses to luteolin. *J Bacteriol* **195**, 3714–3723.
14 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.
15 Lerche M, Dian C, Round A, Lönneborg R, Brzezinski P & Leonard GA (2016) The solution configurations of inactive and activated DntR have implications for the sliding dimer mechanism of LysR transcription factors. *Sci Rep* **6**, 19988.
16 Ruangprasert A, Craven S, Neidle E & Momany C (2010) Full–length structures of BenM and two variants reveal different oligomerization schemes for LysR–type transcriptional regulators. *J Mol Biol* **404**, 568–586.
17 Monferrer D, Tralau T, Kertesz MA, 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. *Mol Microbiol* **75**, 1199–1214.
18 Huang JZ & Schell MA (1991) In vivo interactions of the NahR transcriptional activator with its target sequences. *J Biol Chem* **266**, 10830–10838.
19 Platero AI, García-Jaramillo M, Santero E & Govantes F (2012) Transcriptional organization and regulatory elements of a *Pseudomonas* sp. strain ADP operon encoding a LysR–type regulator and a putative solute transport system. *J Bacteriol* **194**, 6561–6573.
20 McFall S, Chugani S & Chakrabarty A (1998) Transcriptional activation of the catechol and chlorocatechol operons: variations on a theme. *Gene* **223**, 257–267.
21 Lönneborg R & Brzezinski P (2011) Factors that influence the response of the LysR type transcriptional regulators to aromatic compounds. *BMC Biochem* **12**, 49.
22 Porrúa O, López-Sánchez A, Platero AI, Santero E, Shingler V & Govantes F (2013) An A–tract at the AtzR binding site assists DNA binding, inducer–dependent repositioning and transcriptional activation of the PatzDEF promoter. *Mol Microbiol* **90**, 72–87.
23 Ogawa N, McFall SM, Klem TJ, Miyashita K & Chakrabarty A (1999) Transcriptional activation of the chlorocatechol degradative genes of *Ralstonia eutropha* NH9. *J Bacteriol* **181**, 6697–6705.
24 Taylor JL, De Silva RS, Kovacikova G, Lin W, Taylor RK, Skorupski K & Kull FJ (2012) The crystal structure of AphB, a virulence gene activator from *Vibrio cholerae*, reveals residues that influence its response to oxygen and pH. *Mol Microbiol* **83**, 457–470.
25 Fritsch P, Urbanowski M & Stauffer G (2000) Role of the RNA polymerase α subunits in MetR–Dependent activation of metE and metH: important residues in the C-terminal domain and orientation requirements within RNA polymerase. *J Bacteriol* **182**, 5539–5550.
26 Lochowska A, Iwanicka–Nowicka R, Zaim J, Witowska–Zimny M, Bolewska K & Hryniewicz 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. *Mol Microbiol* **53**, 1199–1214.
27 Tyrrell R, Verschueren KH, Dodson EJ, Murshudov GN, Addy C & Wilkinson AJ (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.
28 Ezezika OC, Haddad S, Clark TJ, Neidle EL & Momany C (2007) Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. *J Mol Biol* **367**, 616–629.
29 Devesse L, Smirnova I, Lönneborg R, Kapp U, Brzezinski P, Leonard GA & Dian C (2011) Crystal structures of DntR inducer binding domains in complex with salicylate offer insights into the activation of LysR–type transcriptional regulator. *Mol Microbiol* **8**, 354–367.
30 Jo I, Chung IY, Bae HW, Kim JS, Song S, Cho YH & Ha NC (2015) Structural details of the OxyR peroxide–

sensing mechanism. *Proc Natl Acad Sci U S A* **112**, 6443–6448.
31 Hayes RP, Moural TW, Lewis KM, Onofrei D, Xun L & Kang C (2014) Structures of the inducer–binding domain of pentachlorophenol–degrading gene regulator PcpR from *Sphingobium chlorophenolicum*. *Int J Mol Sci* **15**, 20736–20752.
32 Alanazi AM, Neidle EL & Momany C (2013) The DNA–binding domain of BenM reveals the structural basis for the recognition of a T-N11–A sequence motif by LysR–type transcriptional regulators. *Acta Crystallogr D* **69**, 1995–2007.
33 Brennan RG & Matthews BW (1989) Structural basis of DNA-protein recognition. *Trends Biochem Sci* **14**, 286–290.
34 Napoli AA, Lawson CL, Ebright RH & Berman HM (2006) Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: recognition of pyrimidine-purine and purine-purine steps. *J Mol Biol* **357**, 173–183.
35 Branden C & Tooze J (1999) Introduction to Protein Structure. Garland Science, New York, NY.
36 Moriuchi R, Takada K, Takabayashi M, Yamamoto Y, Shimodaira J, Kuroda N, Akiyama E, Udagawa M, Minai R, Fukuda M *et al.* (2017) Amino acid residues critical for DNA binding and inducer recognition in CbnR, a LysR-type transcriptional regulator from *Cupriavidus necator* NH9. *Biosci Biotechnol Biochem* **81**, 2119–2129.
37 Hiraki M, Kato R, Nagai M, Satoh T, Hirano S, Ihara K, Kudo N, Nagae M, Kobayashi M, Inoue M *et al.* (2006) Development of an automated large–scale protein–crystallization and monitoring system for high–throughput protein–structure analyses. *Acta Crystallogr D* **62**, 1058–1065.
38 Kabsch E (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Cryst* **26**, 795–800.
39 Keegan RM, Long F, Fazio VJ, Winn MD, Murshudov GN & Vaginb AA (2011) Evaluating the solution from MrBUMP and BALBES. *Acta Crystallogr D* **67**, 313–323.
40 Afonine P, Grosse-Kunstleve RW, Echols N, Headd J, Moriarty N, Mustyakimov M, Terwilliger T, Urzhumtsev A, Zwart P & Adams P (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D* **68**, 352–367.
41 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132.
42 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53**, 240–255.
43 Collaborative computational project Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* **50**, 760–763.
44 Miller JH (1972) Experiments in Molecular Genetics. CSH Laboratory Press, Cold Spring Harbor, NY.

## Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

**Fig. S1.** SA-omit maps for Q29 and T33 of chain A and B of CbnR(DBD).

**Table S1.** Base and sugar-phosphate atoms within 4 Å from the protein in the crystal structure of CbnR (DBD)-RBS (PDB code: 5XXP).

**Table S2.** Base and sugar-phosphate atoms within 4 Å from the protein in the crystal structure of BenM (DBD)-RBS (PDB code: 4IHT).