



Crystal structure of the effector-binding domain of *Synechococcus elongatus* CmpR in complex with ribulose 1,5-bisphosphate

Didel M. Mahounga, Hui Sun and Yong-Liang Jiang*

School of Life Sciences and Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei, Anhui 230027, People's Republic of China. *Correspondence e-mail: jyl@ustc.edu.cn

Received 4 April 2018

Accepted 16 June 2018

Edited by M. J. van Raaij, Centro Nacional de Biotecnología – CSIC, Spain

Keywords: crystal structure; LysR-type transcription factors; cyanobacteria; *Synechococcus elongatus* PCC 7942; CmpR-EBD complex with ribulose 1,5-bisphosphate; CO₂-concentrating mechanism.

PDB reference: CmpR effector-binding domain, complex with ribulose 1-5-bisphosphate, 5z49

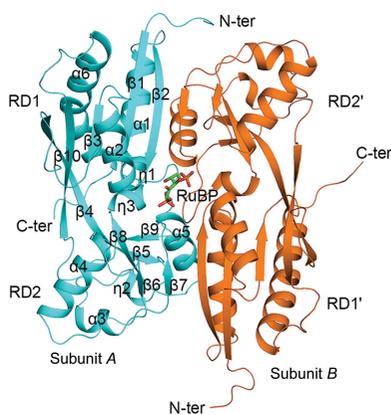
Supporting information: this article has supporting information at journals.iucr.org/f

The CO₂-concentrating mechanism (CCM) has evolved to improve the efficiency of photosynthesis in autotrophic cyanobacteria. CmpR, a LysR-type transcriptional regulator (LTTR) from *Synechococcus elongatus* PCC 7942, was found to regulate CCM-related genes under low-CO₂ conditions. Here, the dimeric structure of the effector-binding domain of CmpR (CmpR-EBD) in complex with the co-activator ribulose 1,5-bisphosphate (RuBP) is reported at 2.15 Å resolution. One RuBP molecule binds to the inter-domain cleft between the two subunits of the CmpR-EBD dimer. Structural comparison combined with sequence analyses demonstrated that CmpR-EBD has an overall structure similar to those of LTTRs of known structure, but possesses a distinctly different effector-binding pattern.

1. Introduction

Cyanobacteria are autotrophic organisms that perform photosynthesis to produce chemical energy for cell growth under a wide range of ambient CO₂ concentrations. In order to adapt to a decline in atmospheric CO₂ and an increasing O₂ level, cyanobacteria evolved a CO₂-concentrating mechanism (CCM; Kaplan & Reinhold, 1999). The CCM consists of several HCO₃⁻/CO₂-uptake systems to accumulate intracellular inorganic carbon (C_i). The absorbed C_i is then diffused into a microcompartment termed the carboxysome, which encapsulates two enzymes, carbonic anhydrase and ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), to improve the efficiency of photosynthesis (Badger & Price, 2003). Carbonic anhydrase converts the accumulated HCO₃⁻ into CO₂, which is the substrate for the carboxylation reaction of RubisCO that produces two molecules of 3-phosphoglyceric acid (3-PGA; Price *et al.*, 2008). However, RubisCO can also perform an oxygenation reaction to produce one molecule of 3-PGA and one molecule of 2-phosphoglycolate (2-PG) under low C_i conditions (Tcherkez, 2016).

The CCM needs to be tightly regulated for cyanobacteria to adapt to the fluctuating environment (Woodger *et al.*, 2007). CCM-related genes are repressed under normal conditions and are up-regulated when CO₂ becomes limited (Burnap *et al.*, 2015). In the cyanobacterium *Synechococcus elongatus* PCC 7942 the CCM is mainly regulated by the global regulator CmpR (UniProt code Q9F1R2; Takahashi *et al.*, 2004). CmpR activates the expression of the *cmp* operon encoding the HCO₃⁻ transporter under C_i-limited conditions in *S. elongatus* PCC 7942, and ribulose 1,5-bisphosphate (RuBP) may



enhance the binding of CmpR to the *cmp* operon (Nishimura *et al.*, 2008). Recently, CmpR has also been shown to act as a repressor of its own gene under high- C_i conditions (Pan *et al.*, 2016).

CmpR belongs to the family of LysR-type transcriptional regulators (LTTRs), which contain the most ubiquitous prokaryotic transcriptional factors involved in regulating various biological processes (Henikoff *et al.*, 1988). The members of this family feature an N-terminal winged helix–turn–helix (HTH) DNA-binding domain (DBD) and a C-terminal effector-binding domain (EBD). To elucidate the structure and the regulatory mechanism of CmpR, we purified and crystallized the EBD domain of CmpR (CmpR-EBD). We solved the crystal structure of CmpR-EBD in complex with the co-activator RuBP at 2.15 Å resolution. Structural analyses of CmpR-EBD reveal that RuBP binds to the dimeric interface of CmpR-EBD with a distinctly different binding pattern. These findings provide structural insights into the transcriptional regulation of CmpR in the CCM.

2. Materials and methods

2.1. Cloning and expression

The coding region for CmpR-EBD (residues 94–323) was amplified by PCR from the genomic DNA of *S. elongatus* PCC 7942. NdeI and XhoI restriction sites were incorporated into the sequences of the forward primer 5'-CGCCATATGGG-TCAACTGCGCCTAGCAGTGATCACC-3' and the reverse primer 5'-CGGCTCGAGTTAAACCCCGACGACCGGCT-GAGGGGC-3', respectively (Table 1). The PCR product was cloned into a pET-28a-derived vector with an N-terminal 6×His tag. The recombinant vector was transformed into *Escherichia coli* strain BL21 (DE3) cells by heat shock, and the cells were cultured overnight at 37°C on a Luria–Bertani (LB) agar plate (10 g tryptone, 5 g yeast extract, 10 g NaCl, pH 7) containing 30 µg ml⁻¹ kanamycin. The recombinant colonies were validated by colony PCR and the construct was further confirmed by sequencing. The recombinant bacteria were grown at 37°C in 1 l LB medium with 30 µg ml⁻¹

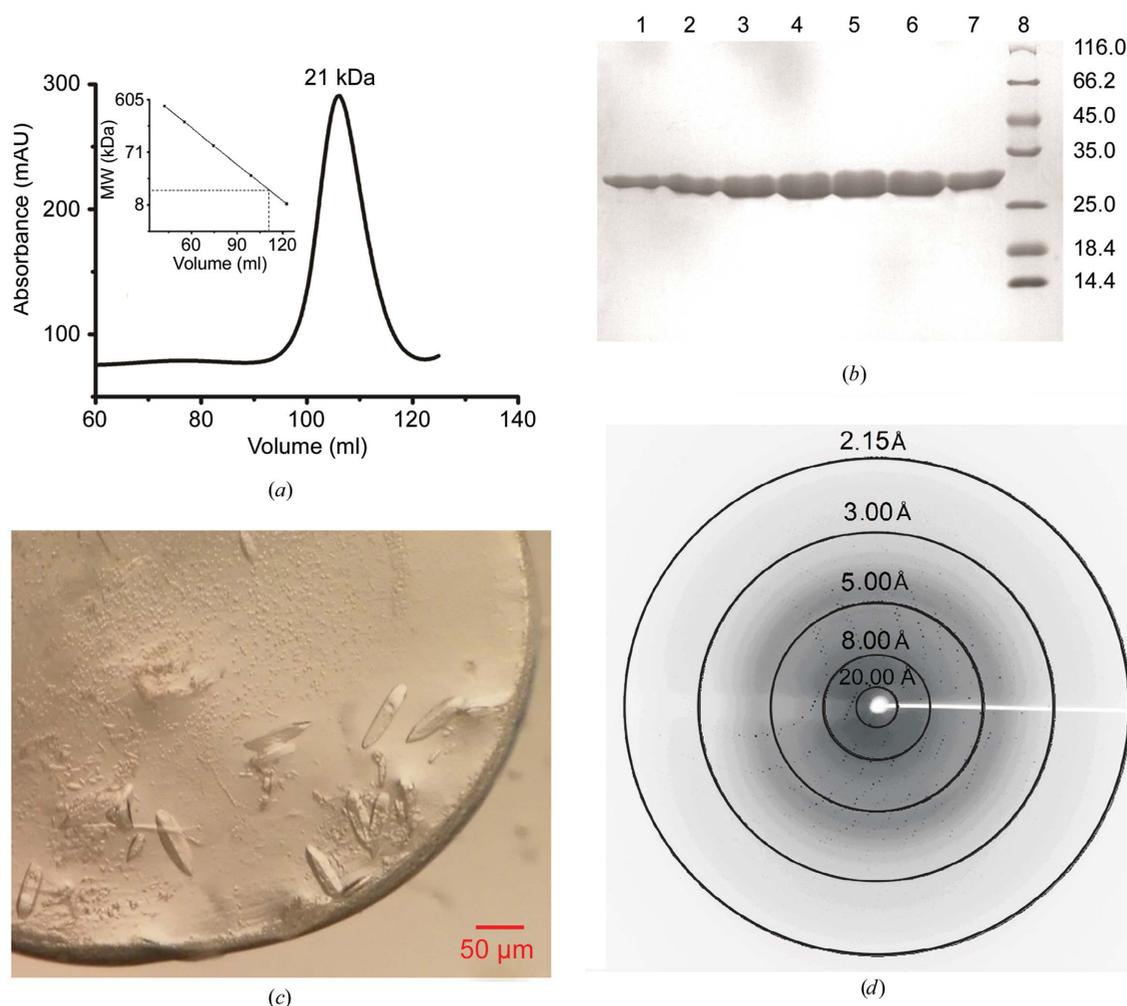


Figure 1
 (a) Size-exclusion chromatography profile of the CmpR-EBD protein obtained using a HiLoad 16/60 Superdex 200 column, showing a single peak for the target protein. The flow rate was 1 ml min⁻¹. (b) Gel electrophoresis profile of protein fractions from size-exclusion chromatography. Lanes 1–7, fractions collected from the peak; lane 8, low-molecular-mass markers (labelled in kDa). (c) Crystals of CmpR-EBD. (d) X-ray diffraction pattern from a single crystal of CmpR-EBD.

Table 1
Macromolecule-production information.

Source organism	<i>S. elongatus</i> strain PCC 7942
DNA source	Genomic DNA from PCC 7942
Forward primer†	5'-CGCCATATGGGTCAACTGCCCTAGCA GTGATCACC-3'
Reverse primer†	5'-CGGCTCGAGTTAAACCCCGACGACCG CTGAGGGGC-3'
Cloning vector	pET-28a-derived
Expression vector	pET-28a-derived
Expression host	<i>E. coli</i> BL21 (DE3)
Complete amino-acid sequence of the construct produced	MGHHHHHMQQLRLAVITTAKYFIPRLIGP FCQRYPGINVSCLKVTNHEGLINRINDNL DDLVLVSRPPSGFDITVQFFLDNPLVVV GPASHPLANQRGISLERLAQEPFILRER GSGTREATQLFAAHNLNLNVKLDLGSN EAIKQAILGGLGLAVLSYHTLTSAGATP ELKMFVEVEGFPPIHRQWHAVYPAGQLST VAATFLDYLLTESQRIAAADIQIPESSTT DPELDAPQPVVGV

† Restriction sites are underlined.

Table 2
Crystallization of CmpR-EBD.

Method	Hanging-drop vapour diffusion
Plate type	96-well plate
Temperature (K)	289
Protein concentration (mg ml ⁻¹)	11
Buffer composition of protein solution	20 mM Tris-HCl pH 8.0, 200 mM NaCl
Composition of reservoir solution	0.1 M Tris pH 8.5, 25% <i>t</i> -butanol
Volume and ratio of drop	1 µl:1 µl
Volume of reservoir (µl)	500

kanamycin. When the OD_{600 nm} reached 0.8, protein expression was induced by adding 0.02 mM isopropyl β-D-1-thiogalactopyranoside and the culture was grown at 16°C for a further 20 h.

2.2. Extraction and purification

To extract the protein, the recombinant *E. coli* cells were collected by centrifugation at 6000g for 10 min at 4°C and the pellets were resuspended in 30 ml loading buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl). After sonication for 15 min on ice, the cell lysate was centrifuged at 12 000g for 25 min at 4°C. The supernatant containing the soluble protein was loaded onto an Ni-NTA column (GE Healthcare) pre-chelated with 0.1 M Ni²⁺ and then equilibrated with loading buffer. The target protein was eluted with loading buffer containing 500 mM imidazole. The eluted protein was further purified by size-exclusion chromatography on an ÄKTAprime plus system (GE Healthcare) using a HiLoad 16/60 Superdex 200 column pre-equilibrated with loading buffer (Fig. 1a). The purity and the homogeneity of the protein were further confirmed by gel electrophoresis (Fig. 1b). The fractions containing CmpR-EBD were pooled and concentrated using a concentrator with a cutoff of 10 kDa (Millipore Amicon) to a final concentration of 11 mg ml⁻¹ in a volume of 2 ml. To obtain the complex of CmpR-EBD with RuBP, the CmpR-EBD protein was soaked with 10 mM RuBP and incubated on ice for 1 h.

Table 3
Crystal parameters and data-collection and refinement statistics.

Values in parentheses are for the highest resolution bin.	
Data collection	
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
<i>a</i> , <i>b</i> , <i>c</i> (Å)	62.55, 86.79, 87.18
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	43.35–2.15 (2.28–2.15)
No. of unique reflections	26440 (4178)
<i>R</i> _{merge}	0.10 (0.47)
CC _{1/2}	99.7 (95.7)
<i>I</i> (σ(<i>I</i>))	15.6 (6.1)
Completeness (%)	99.6 (99.3)
Multiplicity	6.9 (7.2)
Refinement statistics	
Resolution (Å)	43.35–2.15
<i>R</i> _{work} / <i>R</i> _{free}	0.19/0.24 (0.22/0.26)
No. of atoms	
Total	3644
Protein	3436
Ligand	18
Water	190
<i>B</i> factors (Å ²)	
Protein	37.6
Ligand	37.4
Water	39.9
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.148
Ramachandran statistics (%)	
Most favoured regions	98.62
Additional allowed regions	1.38
Disallowed regions	0
PDB code	5z49

2.3. Crystallization, data collection and processing

Crystallization of CmpR-EBD was performed by the hanging-drop vapour-diffusion method in 96-well plates at 16°C. Crystals were grown using 1 µl protein solution (11 mg ml⁻¹ protein in loading buffer) mixed with an equal volume of reservoir solution (0.1 M Tris-HCl pH 8.5, 25% *t*-butanol; Table 2). The crystals appeared in 2 d and reached maximum size in one week (Fig. 1c). A single crystal was transferred into the cryoprotectant (reservoir solution supplemented with 30% glycerol) and flash-cooled with liquid nitrogen. X-ray diffraction data were collected at 100 K in a liquid-nitrogen stream using an ADSC Q315r CCD on beamline BL17U1 at the Shanghai Synchrotron Radiation Facility (SSRF). All diffraction data were integrated and scaled with *HKL-2000* (Otwinowski & Minor, 1997).

2.4. Structure determination and refinement

The crystal structure of CmpR-EBD was determined by molecular replacement with *MOLREP* (Vagin & Teplyakov, 2010) using the coordinates of NdhR from *Synechocystis* sp. strain PCC 6803 (PDB entry 5y2w; Jiang *et al.*, 2018), which has 57% sequence identity, as the search model. The initial model was further refined using the maximum-likelihood method as implemented in *REFMAC5* (Murshudov *et al.*, 2011) as part of the *CCP4i* program suite, and the model was rebuilt interactively using *Coot* (Emsley *et al.*, 2010) until the free *R* factor converged (Table 3). The final model was

evaluated with *MolProbity* (Chen *et al.*, 2010) and *PROCHECK* (Laskowski *et al.*, 1993).

3. Results and discussion

3.1. Overall structure of CmpR-EBD

X-ray diffraction data were collected from a single crystal of CmpR-EBD at SSRF (Fig. 1*d*). The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 62.55$, $b = 86.79$, $c = 87.18$ Å, $\alpha = \beta = \gamma = 90.00^\circ$ (Table 3). The Matthews coefficient suggests that there are two molecules in the asymmetric unit, with a V_M value of 2.23 Å³ Da⁻¹ and a solvent content of 44.84%. The crystal structure of CmpR-EBD was solved by molecular replacement and refined to 2.15 Å resolution, with R and R_{free} values of 0.19 and 0.24, respectively. Data-collection and refinement statistics are listed in Table 3. Each asymmetric unit contains two subunits, which form a stable dimer with a buried interface area of about 1600 Å². The two subunits are quite similar to each other, with a root-mean-square deviation (r.m.s.d.) of 1.66 Å over 200 C α atoms. Each subunit of CmpR-EBD adopts an α/β structure with two Rossmann-like subdomains (termed RD1 and RD2). RD1 includes the N-terminal residues Gly94–Asp165 and the C-terminal residues Gln270–Ser323, whereas RD2 consists of residues Asn166–Arg269. In more detail, the RD1 subdomain is composed of five β -strands ($\beta 1$ – $\beta 3$, the N-terminal moiety of $\beta 4$ and the C-terminal moiety of $\beta 10$), which are surrounded by three α -helices and one 3_{10} -helix (Fig. 2*a*). The RD2 subdomain is similar to RD1 and contains five β -strands ($\beta 5$ – $\beta 9$), three α -helices ($\alpha 3$ – $\alpha 5$) and two 3_{10} -helices ($\eta 1$ and $\eta 3$) (Fig. 2*a*). The RD1 and RD2 subdomains are connected by a crossover region of two anti-parallel β -sheets $\beta 4$ and $\beta 10$ (Fig. 2*a*).

3.2. The binding site for RuBP

As suggested in a previous report (Dangel & Tabita, 2015), CmpR and other regulators of the Calvin–Bassham–Benson pathway (CbbRs) play important roles in regulating CCM. The binding of metabolic effectors to CbbRs may trigger conformational changes and thus coordinate the transcriptional activity (Schell, 1993). The *S. elongatus* PCC 7942 CmpR has been reported to regulate the expression of the *cmp* operon encoding the HCO₃⁻ transporter that is required for CCM. The metabolic effector RuBP could enhance the binding of the *cmp* operon by CmpR (Nishimura *et al.*, 2008), indicating direct binding of RuBP to CmpR. Indeed, in the structure of CmpR-EBD we found that one molecule of the co-activator RuBP binds to the dimeric interface between the subunits of CmpR-EBD, with the two phosphate groups of RuBP inserting into the two inter-subdomain clefts (Fig. 2*a*). The inter-subdomain cleft is mainly formed by the helices $\eta 1$ from RD1 and $\alpha 5$ from RD2 of each subunit. The cleft is positively charged, which favours the binding of the acidic moieties of RuBP. In more detail, one phosphate group of the RuBP molecule is stabilized by Lys106, Ser226 and Asn227, in addition to Lys106' from subunit *B* (Fig. 2*b*), whereas another

phosphate group is fixed by the counterparts from subunits *A* and *B* (Fig. 2*b*). Previous structures of LTTR in complex with an effector revealed two different binding sites, termed site 1 and site 2, respectively (Craven *et al.*, 2009). The primary binding site (site 1) is located in the cleft between RD1 and RD2 of one subunit (Maddocks & Oyston, 2008), whereas the secondary binding site (site 2) is adjacent to the C-terminal region of RD1. The binding of two different effectors to site 1 and site 2 of *Acinetobacter baylyi* BenM had a synergistic

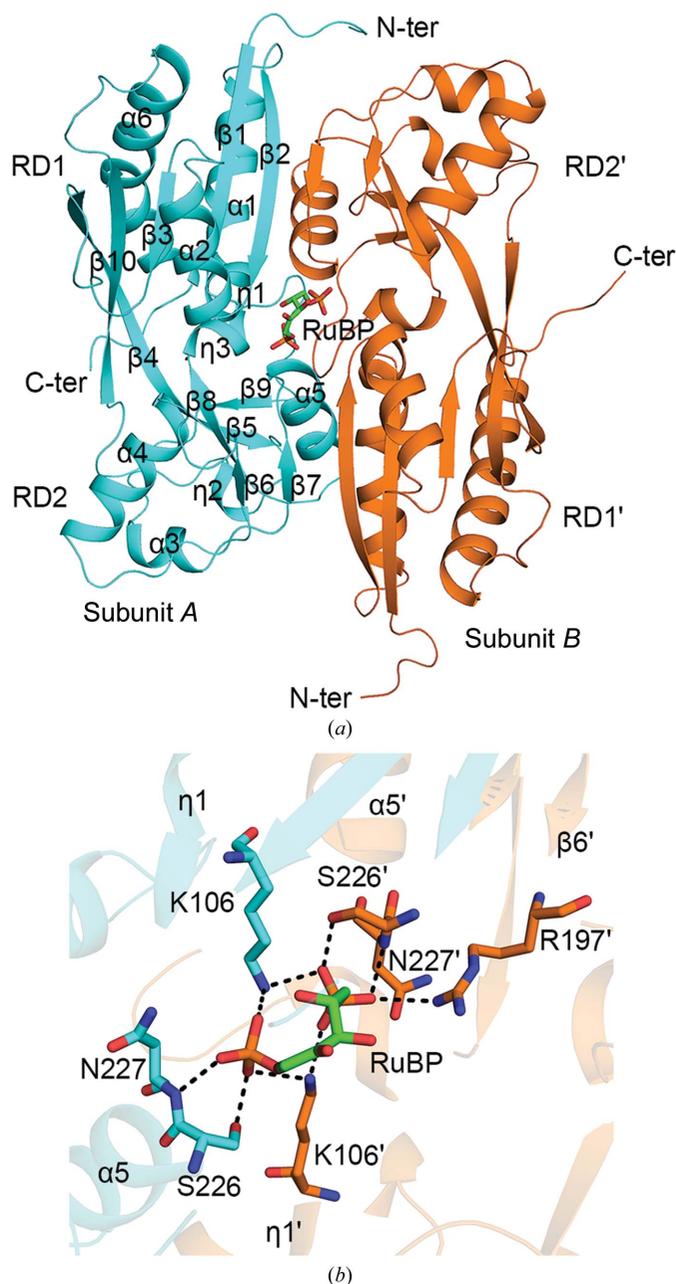


Figure 2
Crystal structure of CmpR-EBD in complex with the co-activator RuBP. (a) The dimeric structure of CmpR-EBD. The two subunits are coloured cyan and orange, respectively. The secondary-structural elements are labelled sequentially. A single RuBP molecule binds to the interface of the CmpR-EBD dimer. The RuBP molecule is shown in stick representation. (b) The RuBP-binding site. The RuBP-binding residues of different subunits are shown as cyan and orange sticks, respectively.

effect on transcriptional regulation (Ezezika *et al.*, 2007). During structural refinement, no obvious electron density was found at the putative site 1 or site 2 of CmpR-EBD. Instead, a single molecule of RuBP binds to the openings of the inter-subdomain clefts between two subunits. It represents a unique ligand-binding pattern that differs from those in previously reported LTTRs of known structure. In addition, the CmpR-EBD structure provides solid evidence for the direct binding of RuBP to CmpR, which is in agreement with previous

studies (Daley *et al.*, 2012; Nishimura *et al.*, 2008). Moreover, the structure of CmpR-EBD further proved the previous findings that most LTTRs require a metabolic effector to control transcriptional activity (Schell, 1993; Maddocks & Oyston, 2008). The binding of effectors to LTTRs may induce conformational changes that are required to regulate the transcriptional activity of the related operon. In our case, the binding of RuBP should trigger conformational changes of CmpR and eventually enhance the expression of the *cmp*

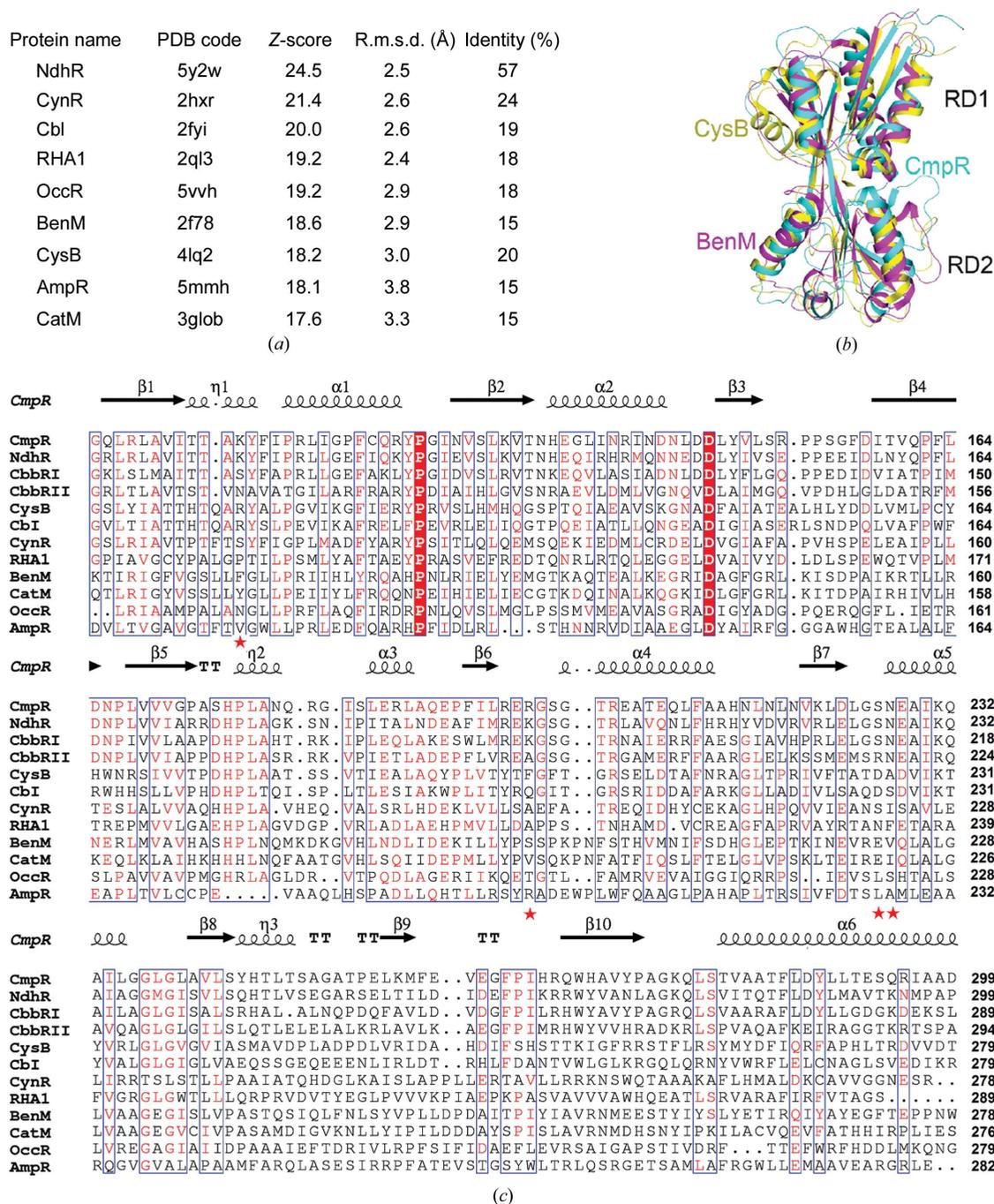


Figure 3 Comparison of the CmpR-EBD structure with those of other LTTRs. (a) A list of CmpR-EBD homologue structures obtained from *DALI*. (b) Superposition of CmpR-EBD (cyan) with BenM (magenta) and CysB (yellow). (c) Multiple-sequence alignment of the EBDs of LTTR proteins. The alignment was performed using *ESPrpt* (<http://esprpt.ibcp.fr/ESPrpt/cgi-bin/ESPrpt.cgi>). The RuBP-binding residues are indicated by red stars.

operon. We propose that the accumulation of intracellular RuBP might be an indicator of C_i deficiency, and the up-regulation of the *cmp* operon via CmpR will in turn lead to a higher intracellular level of the HCO_3^- transporter and a higher efficiency of C_i uptake.

3.3. Structural comparisons with other LTTRs

To obtain structural insights, we compared the CmpR-EBD structure with other LTTR structures using the DALI server. The results gave dozens of hits, all of which belonged to the LTTR family. Following the first hit, the homologue NdhR with 57% sequence identity (PDB entry 5y2w, Z-score 24.5, r.m.s.d. of 2.5 Å for 202 C^α atoms; Jiang *et al.*, 2018), the top homologues include *E. coli* CynR (PDB entry 2hxr, Z-score 21.4, r.m.s.d. of 2.6 Å for 203 C^α atoms; Midwest Center for Structural Genomics, unpublished work), *E. coli* CbI (PDB entry 2fyi, Z-score 20, r.m.s.d. of 2.6 Å for 202 C^α atoms; Stec *et al.*, 2006) and *Rhodococcus jostii* RHA1 (PDB entry 2ql3, Z-score 19.2, r.m.s.d. of 2.4 Å for 195 C^α atoms; Midwest Center for Structural Genomics, unpublished work) (Fig. 3a). Structural comparison of CmpR-EBD with *Salmonella typhimurium* CysB (PDB entry 4lq2; M. Mittal, A. K. Singh & S. Kumaran, unpublished work) and *A. baylyi* BenM (PDB entry 2f78; Ezezika *et al.*, 2007) revealed that these EBDs share a similar overall structure with two Rossmann-fold subdomains (Fig. 3b). The multiple sequence alignment showed that the RuBP-binding residues Ser226 and Asn227 are conserved in CbbRs such as *S. elongatus* CmpR, *Synechocystis* NdhR and *Thiobacillus denitrificans* CbbR_I and CbbR_{II}, whereas significant sequence variations of the effector-binding cleft could be found in other LTTRs, including *E. coli* CynR and CbI, *A. baylyi* BenM and CatM, *R. jostii* RHA1, *S. typhimurium* CysB, *Pseudomonas aeruginosa* AmpR and *Agrobacterium tumefaciens* OccR (Fig. 3c). This suggests that the RuBP-binding pattern of CmpR-EBD is common among CbbRs, which adopt a transcriptional mechanism that differs from other LTTRs of known structure.

Acknowledgements

We thank Professor Cong-Zhao Zhou of the University of Science and Technology of China for his supervision of this project. We appreciate the assistance of the staff at the Shanghai Synchrotron Radiation Facility (SSRF) and the Core Facility Center for Life Sciences at the University of Science and Technology of China.

Funding information

This work was supported by the National Natural Science Foundation of China (<http://www.nsf.gov.cn>; grant Nos. 31630001 and 31500598) and the CAS–TWAS President's PhD Fellowship.

References

- Badger, M. R. & Price, G. D. (2003). *J. Exp. Bot.* **54**, 609–622.
- Burnap, R. L., Hagemann, M. & Kaplan, A. (2015). *Life*, **5**, 348–371.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst. D* **66**, 12–21.
- Craven, S. H., Ezezika, O. C., Haddad, S., Hall, R. A., Momany, C. & Neidle, E. L. (2009). *Mol. Microbiol.* **72**, 881–894.
- Daley, S. M. E., Kappell, A. D., Carrick, M. J. & Burnap, R. L. (2012). *PLoS One*, **7**, e41286.
- Dangel, A. W. & Tabita, F. R. (2015). *J. Bacteriol.* **197**, 3488–3498.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.
- Ezezika, O. C., Haddad, S., Clark, T. J., Neidle, E. L. & Momany, C. (2007). *J. Mol. Biol.* **367**, 616–629.
- Henikoff, S., Haughn, G. W., Calvo, J. M. & Wallace, J. C. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 6602–6606.
- Jiang, Y.-L., Wang, X.-P., Sun, H., Han, S.-J., Li, W.-H., Cui, N., Li, G.-M., Zhang, J.-Y., Cheng, W., Cao, D.-D., Zhang, Z.-Y., Zhang, C.-C., Chen, Y. & Zhou, C.-Z. (2018). *Proc. Natl Acad. Sci. USA*, **2**, 403–408.
- Kaplan, A. & Reinhold, L. (1999). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 539–570.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Maddocks, S. E. & Oyston, P. C. F. (2008). *Microbiology*, **154**, 3609–3623.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Nishimura, T., Takahashi, Y., Yamaguchi, O., Suzuki, H., Maeda, S. & Omata, T. (2008). *Mol. Microbiol.* **68**, 98–109.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pan, L.-L., Onai, K., Uesaka, K., Ihara, K., Natsume, T., Takatani, N., Ishiura, M. & Omata, T. (2016). *Biomed. Genet. Genomics*, **1**, <https://doi.org/10.15761/BGG.1000123>.
- Price, G. D., Badger, M. R., Woodger, F. J. & Long, B. M. (2008). *J. Exp. Bot.* **59**, 1441–1461.
- Schell, M. A. (1993). *Annu. Rev. Microbiol.* **47**, 597–626.
- Stec, E., Witkowska-Zimny, M., Hryniewicz, M. M., Neumann, P., Wilkinson, A. J., Brzozowski, A. M., Verma, C. S., Zaim, J., Wysocki, S. & Bujacz, G. D. (2006). *J. Mol. Biol.* **364**, 309–322.
- Takahashi, Y., Yamaguchi, O. & Omata, T. (2004). *Mol. Microbiol.* **52**, 837–845.
- Tcherkez, G. (2016). *Plant Cell Environ.* **39**, 983–997.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D* **66**, 22–25.
- Woodger, F. J., Bryant, D. A. & Price, G. D. (2007). *J. Bacteriol.* **189**, 3335–3347.