Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic studies of the cofactor-binding domain of the LysR-type transcriptional regulator Cbl from *Escherichia coli*

Emilia Stec,[a] Malgorzata Witkowska,[b] Monika M. Hryniewicz,[b] Andrzej M. Brzozowski,[c] Anthony J. Wilkinson[c] and Grzegorz D. Bujacz[a]*

[a]Faculty of Biotechnology and Food Sciences, Technical University of Lodz, Stefanowskiego 4/10, 90-924 Lodz, Poland, [b]Department of Microbial Biochemistry, Polish Academy of Sciences, Pawinskiego 5A, 02-106 Warsaw, Poland, and [c]Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York Y01 5DD, England

Correspondence e-mail: gdbujacz@mail.p.lodz.pl

Received 25 May 2004
Accepted 9 July 2004

Cbl (CysB-like protein) is a member of the family of LysR-type transcriptional regulators (LTTRs) and controls genes engaged in sulfur assimilation in *Escherichia coli*. It has been postulated that adenosine 5-phosphosulfate (APS) is responsible for abolishing Cbl-activated transcription from the *ssu* promoter (Bykowski *et al.*, 2002). To elucidate the structural basis of Cbl function and to confirm the role of APS as an anti-inducer, the cofactor-binding domain of Cbl (c-Cbl, MW = 26 kDa) was cloned, purified and crystallized in the presence of APS. The crystals belong to space group $C222_1$, but show substantial variation of the unit-cell parameters and diffraction anisotropy. Despite this, X-ray data extending to 3.0 Å resolution have been collected and solution of the structure by molecular replacement is in progress.

1. Introduction

LysR-type transcriptional regulators (LTTRs) comprise the largest family of prokaryotic DNA-binding transcription factors and are broadly represented in Bacteria and also found in Archaea and chloroplasts (Henikoff *et al.*, 1988; Schell, 1993; Perez-Rueda & Collado-Vides, 2000, 2001). LTTRs control expression of diverse genes involved in a wide range of cellular processes such as amino-acid biosynthesis, CO_2 fixation, ion transport, antibiotic resistance, initiation of nodulation, chromosomal replication and synthesis of virulence factors. The LTTR family groups similarly sized proteins (270–330 amino acids) sharing amino-acid sequence similarities over approximately 280 residues, with the highest conservation in the N-terminal 65 residues, which contains a helix–turn–helix motif. The functional significance of the N-terminal domain for DNA binding has been confirmed by mutational studies performed on several LTTRs. Most LTTRs bind DNA as dimers or tetramers and engage rather long sequences (40–60 bp) in these interactions; many regulators occupy multiple binding sites within the regulatory regions of their target promoters. A requirement for a small molecule to act as a coinducer in the transcriptional activation of responsive genes is a common feature of most of the family members. The cofactor-binding/response functions have been ascribed by mutational analyses to the C-terminal domain, in which subregions critical for this function have been identified (Schell, 1993). Several studies suggest that interaction with cognate cofactors promotes a conformational change in the LysR-type proteins (Akakura & Winans, 2002; Tyrrell *et al.*, 1997). 'Conformational switching' may involve alterations of DNA-binding topology or/and stoichiometry at responsive promoters to ensure appropriate contacts with RNA polymerase. Structural studies on LTTRs have been hampered by the generally poor solubility of these proteins. To date, the only crystal structures that have been determined are of the C-terminal domains of *Klebsiella aerogenes* CysB (Tyrrell *et al.*, 1997) and *Escherichia coli* OxyR (Choi *et al.*, 2001) (dimeric forms) and *Ralstonia eutropha* full-length CbnR (tetrameric form; Muraoka *et al.*, 2003).

In this paper, we describe the crystallization of the C-terminal fragment of the transcription factor Cbl from *E. coli* (MW = 26 kDa, 229 amino acids). Cbl is the closest relative of CysB among the LysR-family members. These proteins share 41% amino-acid identity and 60% similarity (Iwanicka-Nowicka & Hryniewicz, 1995). Both CysB and Cbl are involved in the control of sulfur assimilation in *E. coli*: CysB is an activator of the genes involved in the assimilatory sulfate-reduction pathway (Kredich, 1996), while Cbl is required for the expression of the *tau* and *ssu* operons, representing sulfate-starvation-induced (*ssi*) genes (van der Ploeg *et al.*, 1997, 1999). The basic functional difference between Cbl and most LTTRs (including CysB) is the lack of a requirement for a coinducer in transcriptional activation. In contrast, the activity of Cbl has been shown to be negatively regulated by adenosine 5′-phosphosulfate (Bykowski *et al.*, 2002).

Crystallization trials with full-length Cbl have been thwarted by the low solubility of the protein. Cbl, like the other LTTRs that have

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

been crystallized, has very low solubility in the absence of glycerol. Most LTTRs need to be stored either in glycerol or in high NaCl and imidazole concentrations to improve their solubility. For example, CbnR, the only known full-length LysR-type protein structure, was concentrated under highly unusual conditions: 1.0 *M* imidazole, 0.5 *M* NaCl and 20 m*M* Tris–HCl pH 7.9 (Muraoka *et al.*, 2003). Therefore, we focused here on a truncated form (residues 88–316) representing the C-terminal cofactor-binding domain of Cbl (c-Cbl).

2. Methods

2.1. Cloning

N-terminal sequencing (by the Edman degradation method at The University of Sheffield, England) of an ~26 kDa fragment of Cbl (residues 1–316) generated by limited proteolysis with chymotrypsin revealed the presence of a fragment beginning at residue 88.

MALDI–TOF spectrometry showed the molecular weight of this fragment to be 25 599.45 Da and establishes it as a C-terminal fragment encompassing residues 88–318 (c-Cbl). Crystals of c-Cbl have been obtained from 0.05 *M* $(NH_4)_2SO_4$ and 0.1 *M* MES pH 6.5 (Fig. 1*a*). However, these dissolved during crystal manipulation. Since subsequent attempts to reproduce the proteolytic cleavage of Cbl were unsuccessful, the production of recombinant c-Cbl was undertaken.

Using pMH243 plasmid DNA (harbouring *cbl*) as a template, the sequence encoding c-Cbl was amplified in a polymerase chain reaction (PCR). The PCR product was cloned into the expression vector pET28 (Novagen). The resulting recombinant plasmid (designated pES1) expresses c-Cbl tagged with a hexahistidine peptide at the N-terminus. The correctness of the construct was confirmed by sequencing. Plasmids encoding His_6-c-Cbl were transformed into competent *E. coli* BL21-CodonPlus cells (Stratagene) for protein expression.

2.2. Expression and purification

E. coli BL21-CodonPlus cells harbouring pES1 were grown in LB medium containing 35 μg ml^{-1} kanamycin. Cells were grown at 310 K to an A_{600} of 0.7 before induction with 1 m*M* IPTG and further growth for 4 h. Harvested cells were washed with 50 m*M* Tris–HCl buffer pH 8.0 containing a protease-inhibitor cocktail (EDTA-free) and repelleted. The cells were incubated with 1 mg ml^{-1} hen egg-white lysozyme at 277 K for 30 min and sonicated. Cell debris was pelleted by centrifugation at 15 000*g* for 20 min and the supernatant was dialysed against lysis buffer (50 m*M* Tris–HCl buffer, 0.3 *M* NaCl, 10 m*M* imidazole pH 8.0) and applied to a column with His-bind resin. The protein was eluted with 50 m*M* Tris–HCl buffer, 0.3 *M* NaCl and 20 m*M* imidazole pH 8.0. Conditions for the preparation and operation of this column were as described in the QIAexpressionist Manual. The purified protein was dialysed against 50 m*M* Tris–HCl and 5% glycerol pH 8.0 and concentrated (using a 10 kDa Microsep concentrator, PALL Life Sciences). The typical yield of soluble His_6-c-Cbl protein was 10–15 mg per litre of culture. The apparent molecular weight of the protein was estimated to be ~56 kDa based on its elution volume upon Superdex HR 200 (Pharmacia) gel-filtration column chromatography. This indicates that the His_6-c-Cbl protein is a dimer. Non-denaturing (native) polyacrylamide gel electrophoresis (PAGE) of both the His_6-c-Cbl protein and His_6-c-Cbl + APS was also carried out. The presence of APS has an effect on the migration in native gel: a second abundant band appears (data not shown).

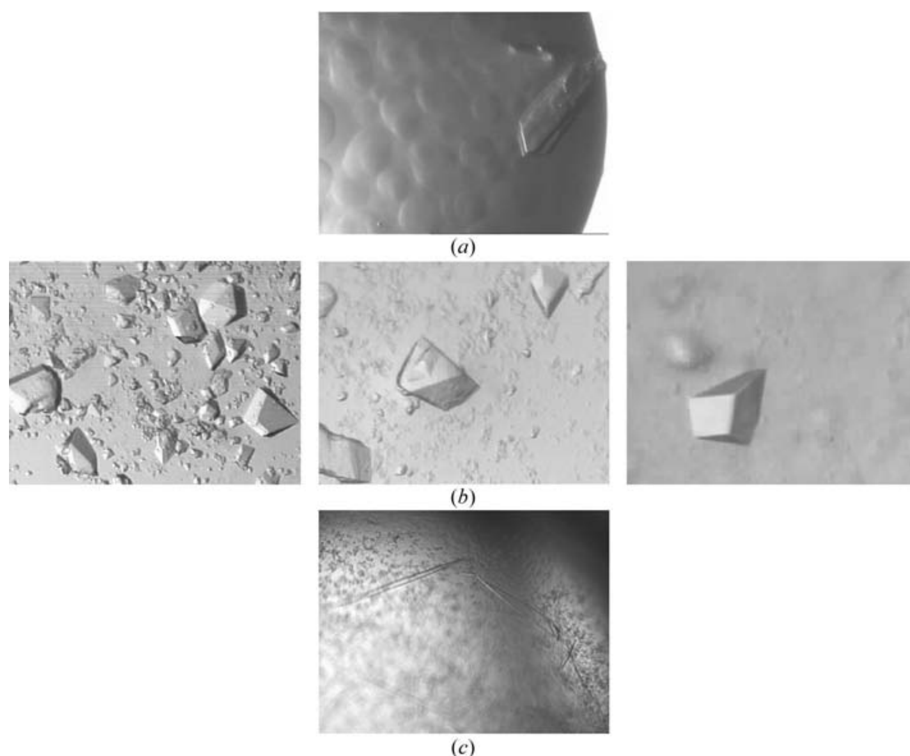

Figure 1
Crystals of the cofactor-binding domain of Cbl: (*a*) proteolytic digestion fragment of Cbl (c-Cbl), (*b*) cofactor-binding domain of c-Cbl (His_6-c-Cbl) grown in the presence of APS (progress in the optimization of the crystallization conditions) and (*c*) cofactor-binding domain of c-Cbl with the His-tag removed (without APS).

2.3. Crystallization and data collection

The purified protein was used for crystallization trials in the presence and absence of a fivefold molar excess of adenosine 5-phosphosulfate (APS). His_6-c-Cbl protein in the presence and absence of APS was concentrated to approximately 60 mg ml^{-1} using a 10 kDa Microsep concentrator (PALL, Life Sciences). The purity of the resulting sample was assessed in Coomassie Brilliant Blue stained SDS polyacrylamide gels.

The His_6-c-Cbl protein and His_6-c-Cbl + APS were also subjected to electrophoresis under non-denaturing conditions (12% polyacrylamide gel). In of the presence of APS, an additional band is seen.

Hampton Research Crystal Screens (Jancarik & Kim, 1991) and Clear Strategy Screens I and II (Brzozowski & Walton, 2001) were used to determine initial crystallization conditions for His_6-c-Cbl + APS. Crystallization was carried out by the sitting-drop vapour-diffusion technique at 312 K using a Mosquito nanolitre pipetting station (LabTech). 0.15 μl aliquots of precipitant solutions were added to 0.15 μl drops containing His_6-c-Cbl and His_6-c-Cbl + APS solutions. After 1 d, initial crystals of His_6-c-Cbl + APS grew in 4.3 *M* NaCl, 0.1 *M* HEPES pH 7.5. In order to optimize the

quality of the crystals, finer intervals of pH and precipitant concentration were employed.

The crystals of His_6-c-Cbl + APS, despite their relatively small dimensions (maximum 0.1 mm in length), were of diffraction quality. As they grew in a high sodium chloride concentration (3.8–4.5 *M* NaCl), no additional cryoprotection was required. Two comparable X-ray diffraction data sets have been collected using two synchrotron sources (ESRF, Grenoble and EMBL, Hamburg). Crystals for data collection in Grenoble were flash-cooled in liquid nitrogen and crystals for data collection in Hamburg were mounted in a capillary for transport and flash-cooled in a nitrogen stream (100 K) before data collection. Hamburg synchrotron data from the His_6-c-Cbl + APS crystal were collected on a MAR CCD detector (165 mm) using a wavelength of 0.8042 Å. The crystal-to-detector distance was 270 mm and each image was collected using a 0.8° oscillation. The data were processed with *DENZO* and scaled using *SCALEPACK* (Otwinowski & Minor, 1997). The diffraction limit for these data was ~2.8 Å and $I/\sigma(I)$ at this resolution was >2, but owing to a high R_{merge} value (~70%) the 2.8–3.0 Å resolution shell was discarded.

Grenoble synchrotron data for the His_6-c-Cbl + APS crystal were collected with an ADSC Q4 CCD detector using a radiation wavelength of 0.934 Å. The crystal-to-detector distance was 162 mm and each image was collected using a 0.5° oscillation. The data were processed and scaled in the same way as for the Hamburg data.

The data statistics of the His_6-c-Cbl + APS crystals are summarized in Table 1.

Table 1
Essential crystallographic data.

Values in parentheses refer to the highest resolution shell.

Data collection	EMBL Hamburg	ESRF Grenoble
Wavelength (Å)	0.8042	0.934
Resolution (Å)	30–3.0 (3.11–3.0)	30–3.0 (3.11–3.0)
Space group	$C222_1$	$C222_1$
Unit-cell parameters (Å)		
a	169.7	167.9
b	242.4	245.1
c	101.6	101.8
Total No. reflections	176398	165373
Unique reflections	42288	41571
Completeness (%)	100 (100)	98.9 (99.2)
R_{sym} (%)	11.5 (50.2)	11.7 (49.2)
$\langle I/\sigma(I) \rangle$	12.7 (2.8)	9.6 (3.1)
Redundancy	4.17 (4.22)	3.98 (3.95)
Unit-cell volume (Å^3)	4179603.5	4187817.0
Molecules per AU	8	8
V_M (Å^3 Da^{-1})	2.5	2.5
Solvent content (%)	50.9	50.9

3. Discussion

As attempts to crystallize full-length Cbl were unsuccessful, crystallization trials of the cofactor-binding C-terminal domain of Cbl (residues 88–316) were carried out.

As subsequent attempts to reproduce the proteolytic digestion of Cbl failed, the coding sequence for the c-Cbl was cloned into the pET28a expression vector. Expression and purification experiments (purification by affinity chromatography using an N-terminal hexahistidine tag) enabled us to prepare His_6-c-Cbl protein suitable for crystallization trials. The His_6-c-Cbl protein was crystallized in the presence and absence of APS. Crystals of His_6-c-Cbl were obtained in the presence but not in the absence of APS (His_6-c-Cbl + APS) (Fig. 1*b*). These crystals appear after 1 d in Crystal Screen II solution No 36: 4.3 *M* NaCl, 0.1 *M* HEPES pH 7.5. Since crystals do not appear without APS, we anticipate that the crystals consist of the His_6-c-Cbl–APS complex, but this is not yet established and we refer to these crystals as His_6-c-Cbl + APS. The best-shaped His_6-c-Cbl + APS crystals grew from 4.2 *M* NaCl and 0.1 *M* Tris–HCl pH 7.5 (Fig. 1*b*). If it is assumed that they contain between four and eight molecules (two to four dimers) of protein per asymmetric unit, the Matthews coefficient (V_M) is in the range 5.1–2.5 Å^3 Da^{-1}, corresponding to a solvent content of 75.5–50.9% (Matthews, 1968). However, the self-rotation function strongly indicates that four dimers are present in the asymmetric unit. Calculations are under way to solve the structure by molecular replacement (MR) using the coordinates of the *K. aerogenes* CysB dimer, which shares significant (41%) sequence identity with Cbl (Tyrrell *et al.*, 1997), as a search model. If these are not successful, we will explore alternative approaches to structure determination involving heavy-atom substitution, either by exposing the crystals to solutions of heavy-metal salts or through selenomethionine incorporation (there are two Met residues in the c-Cbl amino-acid sequence).

In addition to these studies, we are currently seeking to (i) improve the His_6-c-Cbl + APS crystals and (ii) establish crystallization conditions for c-Cbl in the absence of APS. Preliminary crystals of c-Cbl (without APS) have already been obtained following removal of the His_6 tag proteolytically (chymotrypsin digestion). These crystals grew from 35% *t*-butanol, 0.1 *M* sodium citrate pH 5.6 (Fig. 1*c*).

Determination of the structure of His_6-c-Cbl in the presence of APS will allow the identification of the cofactor-binding site and the nature of the cofactor–protein interactions. Additionally, solving the structures of both His_6-c-Cbl and His_6-c-Cbl–APS would allow comparison of these structures and characterization of the response of Cbl to the cofactor (APS) at a molecular level.

It is noteworthy that CysB, in contrast to Cbl, interacts with two types of small ligands: *N*-acetylserine and thiosulfate, whose effects were explained by subtle conformational changes within the tetrameric protein (Kredich, 1992). Moreover, the solved structure of the CysB dimer (Tyrrell *et al.*, 1997) confirmed the presence of a cavity in each CysB monomer that would accommodate each of these small ligands. We expect that the native cofactor-binding domain of Cbl forms a similar fold to that of CysB and it would be of great interest to observe its conformational transition upon interaction with bulky ligand such as APS. Despite these differences, Cbl and CysB represent a unique pair of closely related LTTRs, since they function together in the regulation of some promoters (*e.g. tau*, *ssu*). Structural studies of Cbl in the presence and absence of APS will advance our understanding of the mechanisms by which quite different signalling molecules determine the specificity of DNA binding and the cooperation of CysB and Cbl in transcriptional control.

We are grateful to the Royal Society East European Consortium and the BBSRC, UK for support. We thank Dr Szymon Krzywda for his assistance and helpful advice.

References

Akakura, R. & Winans, S. C. (2002). *J. Biol. Chem.* **277**, 15773–15780.

Brzozowski, A. M. & Walton, J. (2001). *J. Appl. Cryst.* **34**, 97–101.

Bykowski, T., van der Ploeg, J. R., Iwanicka-Nowicka, R. & Hryniewicz, M. M. (2002). *Mol. Microbiol.* **43**, 1347–1358.

Choi, H., Kim, S., Mughopadhyay, P., Cho, S., Woo, J., Storz, G. & Ryu, S. (2001). *Cell*, **105**, 103–113.

Henikoff, S., Haughn, G. W., Calvo, J. M. & Wallace, J. C. (1998). *Proc. Natl Acad. Sci. USA*, **85**, 6602–6606.

Iwanicka-Nowicka, R. & Hryniewicz, M. M. (1995). *Gene*, **166**, 11–17.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Kredich, N. M. (1992). *Mol. Microbiol.* **6**, 2747–2753.

Kredich, N. M. (1996). *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd ed., edited by F. C. Neidhardt, pp. 514–527. Washington DC: ASM Press.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K. & Senda, T. (2003). *J. Mol. Biol.* **328**, 555–566.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Perez-Rueda, E. & Collado-Vides, J. (2000). *Nucleic Acids Res.* **28**, 1838–1847.

Perez-Rueda, E. & Collado-Vides, J. (2001). *J. Mol. Evol.* **53**, 172–179.

Ploeg, J. R. van der, Iwanicka-Nowicka, R., Bykowski, T., Hryniewicz, M. M. & Leisinger, T. (1999). *J. Biol. Chem.* **274**, 29358–29365.

Ploeg, J. R. van der, Iwanicka-Nowicka, R. & Hryniewicz, M. M. (1997). *J. Bacteriol.* **179**, 7671–7678.

Schell, M. A. (1993). *Annu. Rev. Microbiol.* **47**, 597–626.

Tyrrell, R., Verschueren, K. G., Dodson, E. J., Murshudov, G. N., Addy, C. & Wilkinson, A. J. (1997). *Structure*, **5**, 1017–1032.