Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Overexpression, purification, crystallization and preliminary X-ray diffraction analysis of the C-terminal domain of Ss-LrpB, a transcription regulator from *Sulfolobus solfataricus*

Eveline Peeters,[a] Bach Thi Mai Hoa,[a] Ingrid Zegers,[b] Daniel Charlier[a] and Dominique Maes[b]*

[a]Erfelijkheidsleer en Microbiologie, Vrije Universiteit Brussel (VUB), Pleinlaan 2, B-1050 Brussels, Belgium, and [b]Laboratorium voor Ultrastructuur, Vrije Universiteit Brussel (VUB) and Vlaams interuniversitair Instituut voor Biotechnologie (VIB), Pleinlaan 2, B-1050 Brussels, Belgium

Correspondence e-mail: dominique.maes@vub.ac.be

Received 7 September 2005
Accepted 10 October 2005
Online 20 October 2005

Ss-LrpB from *Sulfolobus solfataricus* P2 belongs to the bacterial/archaeal superfamily of Lrp-like (leucine-responsive regulatory protein-like) transcription regulators. The N-terminal DNA-binding domain contains a HTH motif and the C-terminal domain contains an $\alpha\beta$-sandwich ($\beta\alpha\beta\beta\alpha\beta$ fold). The C-terminal domain was overexpressed in *Escherichia coli*, purified and crystallized using the hanging-drop vapour-diffusion method. The crystals belong to space group $P2_12_12$, with unit-cell parameters a = 59.35, b = 74.86, c = 38.08 Å and a data set was collected to 2.0 Å resolution. Structure determination using a selenomethionine derivative is under way.

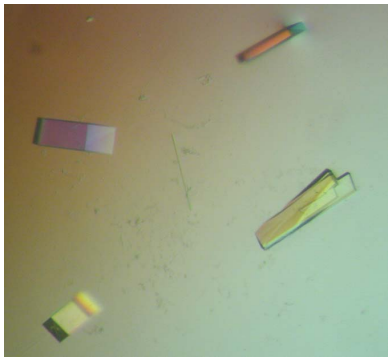
1. Introduction

Archaea are microorganisms constituting a domain of life that is distinct from the bacteria and eukarya (Woese *et al.*, 1990). Most characterized archaeal species are extremophiles, thriving at extremes of temperature, pH, salt concentration *etc.* *Sulfolobus solfataricus* P2 is a hyperthermoacidophilic archaeon, growing optimally at 353 K and pH 2–4 (Zillig *et al.*, 1980). Transcription in archaea appears to be typified by a eukaryal-like basal transcription apparatus, regulated mainly by bacteria-like helix–turn–helix (HTH) transcription regulators (Geiduschek & Ouhammouch, 2005).

Most of the characterized archaeal transcription regulators belong to the leucine-responsive regulatory protein (Lrp) family (Brinkman *et al.*, 2003). This family is composed of both global and specific regulators. To date, few archaeal regulators with known physiological function have been described (Brinkman *et al.*, 2002; Ouhammouch *et al.*, 2003). Although it has been claimed that all specific Lrp-like regulators have a role in the regulation of amino-acid metabolism (Brinkman *et al.*, 2003), this was refuted by the discovery that Ptr2 from *Methanocaldococcus jannaschii* activates a ferredoxin and a rubredoxin gene (Ouhammouch *et al.*, 2003).

All Lrp-like proteins are small basic proteins composed of two domains connected by an extended flexible hinge. Crystal structures of two Lrp-like proteins, both archaeal, have been determined: LrpA from *Pyrococcus furiosus* (Leonard *et al.*, 2001) and FL11 from *Pyrococcus* OT3 (Koike *et al.*, 2004). The N-terminal DNA-binding domain contains an HTH motif and the C-terminal domain contains an $\alpha\beta$-sandwich ($\beta\alpha\beta\beta\alpha\beta$ fold). This C-terminal domain is also called the regulation of amino-acid metabolism (RAM) domain and resembles the ACT domain, an allosteric regulatory domain of many metabolic enzymes (Ettema *et al.*, 2002). It has been shown to be responsible for oligomerization (Leonard *et al.*, 2001; Chen *et al.*, 2001) and in the case of *Escherichia coli* Lrp for ligand binding and activation of transcription (Platko & Calvo, 1993).

Ss-LrpB from *S. solfataricus* binds its own control region (Peeters *et al.*, 2004), as do most Lrp-like proteins. This binding occurs cooperatively at three regularly spaced 15 bp binding sites with as consensus the palindromic sequence 5′-TTGYAWWWWTRCAA-3′ (Y = pyrimidine, R = purine, W = weak bp). This is in contrast to most other Lrp-like proteins, which recognize clusters of strongly degenerate binding sites (Wang & Calvo, 1993; Ouhammouch & Geiduschek, 2001). It has also been suggested that a secondary DNA structure might be recognized rather than a consensus sequence (for



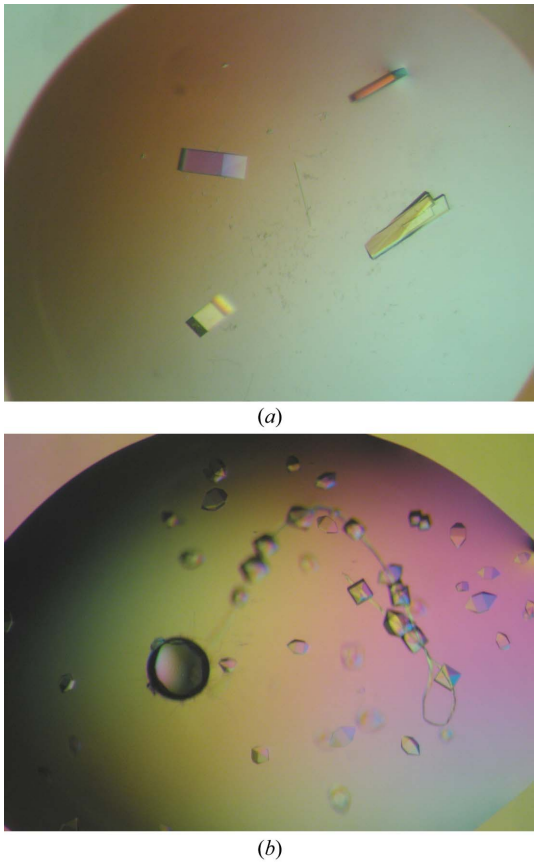
example, a pattern of repeating weak base pairs; Calvo & Matthews, 1994; Beloin *et al.*, 2000). Ss-LrpB is therefore an interesting candidate for function prediction: the consensus binding sequence is being used to scan all promoter/operator regions in the *S. solfataricus* P2 genome. This might also provide useful information regarding a potential ligand.

Here, we report the cloning, overexpression, purification and crystallization of the C-terminal domain of Ss-LrpB comprising residues 69–147 of the full-length protein, omitting the last seven residues. This approach was chosen because full-length Ss-LrpB has a strong tendency to aggregate in solution. We aim towards a better understanding of the molecular regulation mechanism, protein oligomerization, potential cofactor binding and thermostability of this transcription regulator.

2. Material and methods

2.1. Cloning and overexpression

A DNA fragment containing residues 205–441 of the *Ss-lrpB* gene was amplified by PCR with pET24aSs-lrpB as template (Peeters *et al.*, 2004). The forward primer 5′-GGAATTC*CATATG*AATGTCTC-ATTAGTAATGG-3′ introduced an *Nde*I restriction site (indicated in italics) and placed an ATG start codon in front of the first codon that translates into the C-terminal domain of Ss-LrpB. The reverse primer (5′-CCG*CTCGAG*CTTATATCTCTCAATTATCCTTGAAATC-3′) introduced an *Xho*I restriction site (indicated in italics). The restricted PCR fragment was ligated into the expression vector pET24a, previously digested with *Nde*I and *Xho*I. Using this cloning site, the protein was expressed including an additional C-terminal His-tag sequence, LEHHHHHH, allowing Ni^{2+} ion affinity purification (see below). This construct was designated pET24aSs-lrpBC-term. The DNA sequence of the insert was confirmed by sequence analysis with an ABI PRISM 3100 Genetic Analyzer (Applied Biosciences). The recombinant plasmid was transformed into *E. coli* BL21(DE3) in order to overexpress the protein, named Ss-LrpBC-term.



(*a*)

(*b*)

Figure 1
(*a*) Crystals obtained from condition I. (*b*) Crystals obtained from condition II.

Table 1
Crystallization conditions.

Condition	I	II
Salt	100 m*M* NaH_2PO_4, 100 m*M* KH_2PO_4	0.5 *M* $(NH_4)_2SO_4$
Buffer	100 m*M* MES pH 6.5	100 m*M* sodium citrate pH 5.6
Precipitant	2.0 *M* NaCl	1 *M* Li_2SO_4
Crystal dimensions (µm)	400 × 50 × 50	150 × 150 × 250

A 2.5 l culture was grown at 303 K in rich medium complemented with 30 µg ml^{-1} kanamycin. Overexpression of Ss-LrpBC-term was induced at a cell density of 6×10^8 cells ml^{-1} by adding 1.0 m*M* IPTG followed by overnight growth. Cells were collected by centrifugation, resuspended in 35 ml extraction buffer (20 m*M* sodium phosphate buffer, 0.2 *M* NaCl, 40 m*M* imidazole pH 7.4) and disrupted by sonication for 1 h with a VibraCell sonicator (Bioblock Scientific) at 20% of the maximal amplitude in a cell cooled at 277 K. Cell debris was removed by ultracentrifugation at 23 000 rev min^{-1} for 1 h (Beckman L8-70 ultracentrifuge, Ti-60 rotor).

2.2. Protein purification

The crude extract was incubated at 328 K for 5 min followed by centrifugation of the denatured *E. coli* proteins at 10 000 rev min^{-1} for 10 min (microcentrifuge). In contrast to full-length Ss-LrpB, Ss-LrpBC-term proved to be sensitive to temperatures exceeding 333 K. The supernatant was loaded onto a HisTrap HP column (1 ml, Amersham Biosciences) allowing Ni^{2+} ion affinity chromatography purification of the His-tagged protein. Equilibration of the column was performed with 20 m*M* sodium phosphate buffer, 0.2 *M* NaCl, 20 m*M* imidazole pH 7.4. Elution of Ss-LrpBC-term was obtained by applying a linear gradient from 20 to 500 m*M* imidazole. The purest fractions were pooled.

2.3. Crystallization

Ss-LrpBC-term was dialyzed against 20 m*M* sodium phospate buffer pH 7.4 and concentrated to a final concentration of 4 mg ml^{-1} as determined by a MicroBCA Protein assay (Pierce). Higher concentrations resulted in protein precipitation. The initial crystallization conditions were screened using the hanging-drop vapour-diffusion method with sparse-matrix crystallizations kits (Crystal Screen II, Hampton Research, Riverside, CA, USA). Immediately prior to setting up crystallization, the protein was centrifuged at 13 000 rev min^{-1} for 10 min to remove precipitates (micro-centrifuge). 2 µl of protein (4 mg ml^{-1}) was mixed with an equal amount of reservoir solution and equilibrated against 500 µl reservoir solution. The crystallization trays were kept at 293 K. Crystals appeared after 2 d in two different conditions (Table 1; Figs. 1*a* and 1*b*). Crystallization under these conditions was repeated and optimized. In many drops too much nucleation occurred, resulting in a large number of small crystals. In order to increase the crystal size these drops were hung above a reservoir of water for 2 d. After dissolution of the crystals the drops were equilibrated again against the same reservoir solution as before to which 50 µl water was added.

2.4. X-ray diffraction collection and processing

X-ray diffraction data were collected from a single crystal at 100 K using a MAR CCD detector (MAR Research, Norderstadt, Germany) at beamline BW7A at the DORIS storage ring, DESY, Hamburg. 206 images were collected at a wavelength of 0.9116 Å and with a rotation angle of 1°. The images were processed using the *HKL* programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The molecular-replacement trials were performed using *PHASER* (from *CCP*4; Read, 2001).

3. Results and discussion

Ss-LrpBC-term was overexpressed and purified from pET24aSs-lrpBC-term transformants of strain BL21 (DE3). A combination of heat treatment and affinity chromatography resulted in electrophoretically pure protein (3.2 mg per litre of culture) which has a theoretical molecular weight of 10.22 kDa and migrated accordingly on SDS–PAGE (Fig. 2). Crystallization of the protein fragment proved to be easy. Crystals were found in two conditions of the screening set (conditions I and II; Table 1). This probably indicates that the C-terminal part has a stable tightly packed structure, as for the LrpA from *P. furiosus* (Leonard *et al.*, 2001). Indeed, the same secondary structure (RAM domain: $\beta\alpha\beta\beta\alpha\beta$ fold) was predicted (3*DPSSM*; Kelley *et al.*, 2000). The exact formulation of conditions I and II and the estimated dimensions of the crystals are given in Table 1. An attempt was made to optimize crystal growth under these conditions, but changing the protein, salt or precipitant concentration frequently resulted in too much nucleation.

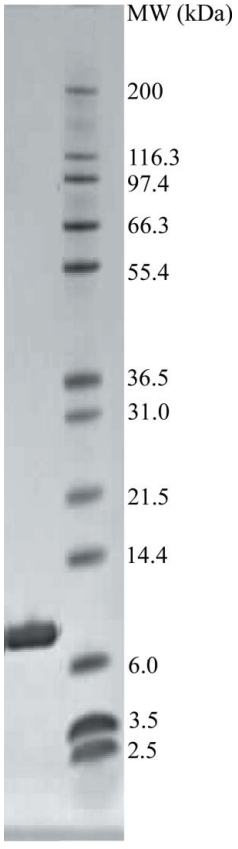


Figure 2
SDS–PAGE of a sample of purified Ss-LrpBC-term after heat treatment and Ni^{2+} ion affinity purification.

Table 2
Data-collection statistics.

Values in parentheses correspond to the highest resolution shell.

Unit-cell parameters (Å, °)	a = 59.35, b = 74.86, c = 38.08, $\alpha = \beta = \gamma$ = 90
Space group	$P2_12_12$
Wavelength (Å)	0.9116
Resolution (Å)	30–2.0 (2.07–2.00)
Total No. of reflections	96635
No. of unique reflections	11970
Completeness (%)	98.6 (98.7)
Redundancy	8.1 (8.1)
Percentage of reflections with $I/\sigma(I) < 3$ (%)	12.4 (36.6)
R_{sym} (%)	4.3 (23.5)
Mosaicity (°)	0.65

A crystal grown from condition I (Fig. 1*a*) diffracted to 1.8 Å. PEG 400 was used as a cryoprotectant at a concentration of 30%(*v*/*v*). A data set to a resolution of 2 Å was collected from a single crystal. The crystals belong to the orthorhombic space group $P2_12_12$, with unit-cell parameters a = 59.35, b = 74.86, c = 38.08 Å. Data-collection information and statistics are given in Table 2. The values of the Matthews coefficient V_M (Matthews, 1968), determined assuming the presence of a dimer in the asymmetric unit, is 2.1 $Å^3$ Da^{-1}. The corresponding value for the solvent content is 41%. Crystals grown from condition II (Fig. 1*b*) were not stable when transferred into cryoprotectant solution, either PEG 400 or glycerol, at concentrations ranging from 10–40%(*v*/*v*). Other cryoprotectants were not tested owing to the limited availability of crystals.

Ss-LrpBC-term shares 27.8% amino-acid identity (55.7% similarity) with the corresponding C-terminal part of LrpA and only 16.5% amino-acid identity (46.8% similarity) with the corresponding C-terminal part of FL11. Preliminary attempts at molecular replacement using the C-terminal part of the LrpA and FL11 as a model, using the dimer as well as the monomer, have not been successful. No clear solution was found. In order to solve the structure a SeMet derivative is being prepared.

This research was financed by a PhD grant from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen) to EP, grants from the Fund for Joint Basic Research-Flanders (FWO-Vlaanderen) and the Research Council of the Vrije Universiteit Brussel, from the Flanders inter-university Institute for Biotechnology (VIB) and the Vlaamse Gemeenschapscommissie. We thank the ESA for finance in the context of Prodex project C90035. The authors gratefully acknowledge the use of synchrotron beamtime at the EMBL beamline BW7A at the DORIS storage ring, DESY, Hamburg.

References

Beloin, C., Exmey, R., Mahé, A.-L., Zouine, M., Cubashi, S. & Le Hégarat, F. (2000). *J. Bacteriol.* **182**, 4414–4424.
Brinkman, A. B., Bell, S. D., Lebbink, R. J., de Vos, W. M. & van der Oost, J. (2002). *J. Biol. Chem.* **277**, 29537–29549.
Brinkman, A. B., Ettema, T. J. G., de Vos, W. M. & van der Oost, J. (2003). *Mol. Microbiol.* **48**, 287–294.
Calvo, J. M. & Matthews, R. G. (1994). *Microbiol. Rev.* **58**, 466–490.
Chen, S., Rosner, M. H. & Calvo, J. M. (2001). *J. Mol. Biol.* **312**, 625–635.
Ettema, T. J. G., Brinkman, A. B., Tani, T. H., Rafferty, J. B. & van der Oost, J. (2002). *J. Biol. Chem.* **277**, 37464–37468.
Geiduschek, E. P. & Ouhammouch, M. (2005). *Mol. Microbiol.* **56**, 1397–1407.
Kelley, L. A., MacCallum, R. M. & Sternberg, M. J. (2000). *J. Mol. Biol.* **299**, 499–520.
Koike, H., Ishijima, S. A., Clowney, L. & Suzuki, M. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 2840–2845.

Leonard, P. M., Smits, S. H. J., Sedelnikova, S. E., Brinkman, A. B., de Vos, W. M., van der Oost, J., Rice, D. W. & Rafferty, J. B. (2001). *EMBO J.* **20**, 990–997.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Ouhammouch, M., Dewhurst, R. E., Hausner, W., Thomm, M. & Geiduschek, E. P. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 5097–5102.
Ouhammouch, M. & Geiduschek, E. P. (2001). *EMBO J.* **20**, 146–156.
Platko, J. & Calvo, J. (1993). *J. Bacteriol.* **177**, 1110–1117.
Peeters, E., Thia-Toong, T. L., Gigot, D., Maes, D. & Charlier, D. (2004). *Mol. Microbiol.* **54**, 321–336.
Read, R. J. (2001). *Acta Cryst.* D**57**, 1373–1382.
Wang, Q. & Calvo, J. M. (1993). *J. Mol. Biol.* **229**, 306–318.
Woese, C., Kandler, O. & Wheels, M. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 4576–4579.
Zillig, W., Stetter, K. O., Wunderl, S., Schulz, W., Priess, H. & Scholz, I. (1980). *Arch. Microbiol.* **125**, 259–269.