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

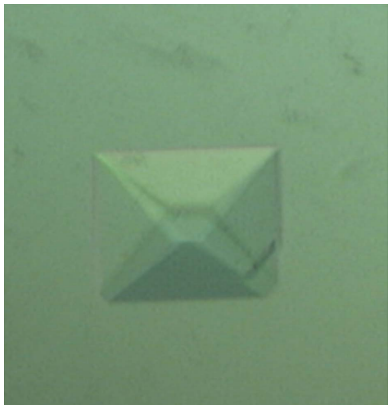
Purification, crystallization and preliminary X-ray crystallographic analysis of ST1022, a putative member of the Lrp/AsnC family of transcriptional regulators isolated from *Sulfolobus tokodaii* strain 7

Noboru Nakano,[a] Thirumananseri Kumarevel,[a]* Emiko Matsunaga,[a] Akeo Shinkai,[a] Seiki Kuramitsu[a,b] and Shigeyuki Yokoyama[a,c,d]*

[a]RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan, [b]Department of Biological Sciences, Graduate School of Science, Osaka University, Tayonaka, Osaka 560-0043, Japan, [c]Genomic Sciences Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan, and [d]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Correspondence e-mail: tskvel@spring8.or.jp, yokoyama@biochem.s.u-tokyo.ac.jp

Received 5 July 2007
Accepted 11 October 2007



The Lrp/AsnC family of transcriptional regulators, also known as feast/famine transcriptional regulators, are widely distributed among bacteria and archaea. This family of proteins are likely to be involved in cellular metabolism, with exogenous amino acids functioning as effectors. Here, the crystallization and preliminary X-ray diffraction analysis of ST1022, a member of the Lrp/AsnC family of proteins, is reported with and without exogenous glutamine as the effector molecule. The crystals of native ST1022 and of the putative complex belong to the tetragonal space group $I422$, with unit-cell parameters $a = b = 103.771$, $c = 73.297$ Å and $a = b = 103.846$, $c = 73.992$ Å, respectively. Preliminary X-ray diffraction data analysis and molecular-replacement solution revealed the presence of one monomer per asymmetric unit.

1. Introduction

The Lrp/AsnC or feast/famine family of transcriptional regulators are widely distributed in bacteria and archaea as an important regulatory system of amino-acid metabolism and related processes (Newman & Lin, 1995; Calvo & Matthews, 1994; Brinkman *et al.*, 2003; Suzuki, 2003) and this family of proteins is most probably restricted to only prokaryotes (Brinkman *et al.*, 2003). Of the Lrp/AsnC family members, the *Escherichia coli* leucine-responsive regulatory protein (Lrp) has been extensively studied (Newman & Lin, 1995; Calvo & Matthews, 1994). *E. coli* Lrp is an abundant protein that functions as a global regulator of amino-acid biosynthesis, transport, protein degradation and intermediary metabolism and responds to leucine (Chen & Calvo, 2002; Brinkman *et al.*, 2003). Recently, proteins similar to Lrp/AsnC family members have also been characterized from archaea (Bell & Jackson, 2001; Ouhammouch, 2004; Geiduschek & Ouhammouch, 2005).

The molecular weights of the Lrp/AsnC family members range between 15 and 17 kDa. In organisms such as *E. coli*, *Agrobacterium tumefaciens*, *Pseudomonas aeruginosa* and *Pyrococcus furiosus*, different multimeric forms of the family members have been reported (Willins *et al.*, 1991; Madhusudhan *et al.*, 1995; Jafri *et al.*, 1999; Brinkman *et al.*, 2000; Chen *et al.*, 2001; Koike *et al.*, 2004). To date, only a few crystal structures have been reported for members of the Lrp/AsnC family: those of proteins from *Pyrococcus* sp. OT3, *E. coli* and *Bacillus subtilis* (Leonard *et al.*, 2001; Koike *et al.*, 2004; Thaw *et al.*, 2006). All of the protein structures share a similar topology containing an N-terminal helix–turn–helix DNA-binding domain and a C-terminal effector-binding domain. A sequence-comparison analysis of ST1022 with other proteins using the *MULTALIN* program (Corpet, 1998) showed it to have 40% identity to the *Pyrococcus* sp. OT3 FL11 protein. Gel-filtration experiments revealed that FL11 forms a higher order assembly in the presence of L-glutamine (Gln; Koike *et al.*, 2004). Similar proteins modulated by binding to the cognate ligands have been described, including Lrp by L-leucine, AsnC by L-asparagine, LysM by L-lysine, the glutamate-uptake regulatory protein by L-glutamate, PutR by L-proline, BkdR by L-valine and MdeR by methionine (Brinkman *et al.*, 2003).

Table 1
Data-collection and refinement statistics of ST1022 protein in the absence (native) and presence (putative complex) of Gln.

Values in parentheses are for the highest resolution shell.

	Native	Putative complex
Space group	*I*422	*I*422
Unit-cell parameters (Å)	$a = b = 103.771$, $c = 73.297$	$a = b = 103.846$, $c = 73.992$
No. of molecules per ASU	1	1
Solvent content (%)	56.3	56.8
Wavelength	1.0000	1.0000
Resolution range (Å)	50.0–1.82 (1.89–1.82)	50.0–1.80 (1.86–1.80)
Unique reflections	18202	19055
Redundancy	13.9 (13.6)	13.6 (11.9)
Completeness (%)	99.9 (99.9)	99.9 (100)
$I/\sigma(I)$	40.2 (6.1)	41.5 (7.7)
$R_{\rm merge}$† (%)	0.065 (0.484)	0.063 (0.343)

† $R_{\rm merge} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where $I(h,i)$ is the intensity value of the ith measurement of h and $\langle I(h)\rangle$ is the corresponding mean value of $I(h)$ for all i measurements.

To date, the only structure of a ligand-bound complex to be reported is that of *E. coli* AsnC with its effector L-asparagine (Thaw *et al.*, 2006). However, the structure in the absence of the ligand (L-asparagine) is not available in order to reveal insights into any conformational change that might occur upon binding to the cognate ligand. In order to clarify the effector-mediated conformational changes, we have crystallized the ST1022 protein from one of our structural genomics targets, the thermoacidophilic crenarchaeon *Sulfolobus tokodaii* strain 7, in the presence and absence of its effector. We now report the preliminary X-ray diffraction data analyses of the ST1022 protein in order to identify the conformational changes that accompany ligand binding.

2. Materials and methods

2.1. Expression and purification of ST1022

The gene encoding the hypothetical regulator protein *ST1022* (SwissProt ID Q972W6) from *S. tokodaii* strain 7 was amplified from the genomic DNA by PCR using the primers 5′-ggaattCATATGG-ATGAAATAGATTTAAGAATTTT-3′ and 5′-ggaattGGATCCTT-ATTAAAAGATAACTATATTTGGTGATTC-3′. The PCR fragment was digested with *Nde*I and *Bam*HI and cloned into the pET-21a(+) expression vector (Novagen). The resultant plasmid was transformed into *E. coli* BL21-CodonPlus (DE3)-RIL-X (Stratagene) strain and the ST1022 protein was overexpressed at mid-log phase by the addition of IPTG to a final concentration of 1 m*M*. Harvested cells (35.5 g) were suspended in lysis buffer (20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl, 5 m*M* β-mercaptoethanol) and sonicated for 15 min. The sonicated cell lysate was incubated with DNase I (3.55 units) and RNase A (100 µg ml^{-1}) with 5 m*M* $CaCl_2$ and 25 m*M* $MgCl_2$ at 310 K for 30 min. An equal volume of preheated buffer (20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl) at 343 K was added to the lysate and this solution was heat-treated at 343 K for 10 min in order to denature most of the nonthermophile contaminant proteins. After centrifugation at 40 000 rev min^{-1} for 60 min at 277 K, the supernatant was subjected to 80% ammonium sulfate precipitation and the pellet was dialyzed against buffer (20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl). The resulting solution was then subjected to a high-throughput purification protocol, the first step of which was a hydrophobic column (Resource ISO, GE Healthcare Biosciences) which was pre-equilibrated with 50 m*M* Na_2HPO_4 buffer pH 7.0. The bound protein was eluted with an ammonium sulfate gradient (0–1.2 *M*) and most of the ST1022 protein eluted at 0.08 *M*. The protein was loaded onto another hydrophobic column (Resource PHE1, GE Healthcare Biosciences) to remove the additional contaminants. The bound sample was washed and eluted using a linear gradient of ammonium sulfate. The ST1022 protein-containing fractions were pooled and concentrated and the salt was removed using a HiPrep Desalting column (GE Healthcare Biosciences). The sample was applied onto a hydroxyapatite column (CHT10; GE Healthcare Biosciences) and the bound sample was eluted with a sodium phosphate gradient (0.01–0.5 *M*). The eluted sample was concentrated and chromatographed on a gel-filtration column (Superdex 75; GE Healthcare Biosciences) which was pre-equilibrated with 20 m*M* Tris–HCl pH 8.0 and 150 m*M* NaCl. The homogeneity of the final purified protein was over 99%, as determined by SDS–PAGE. A total of 1.2 mg protein was purified from 35.5 g bacteria, concentrated to 5.4 mg ml^{-1} in buffer (20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 1 m*M* DTT) and used for crystallization studies.

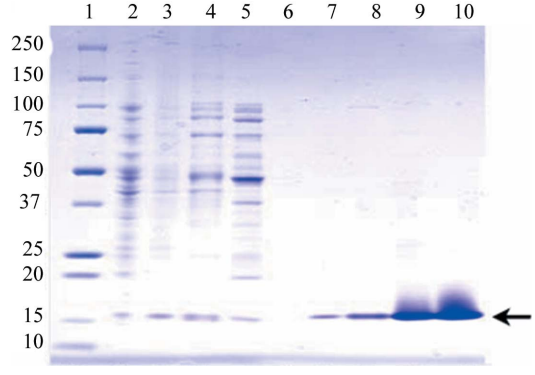
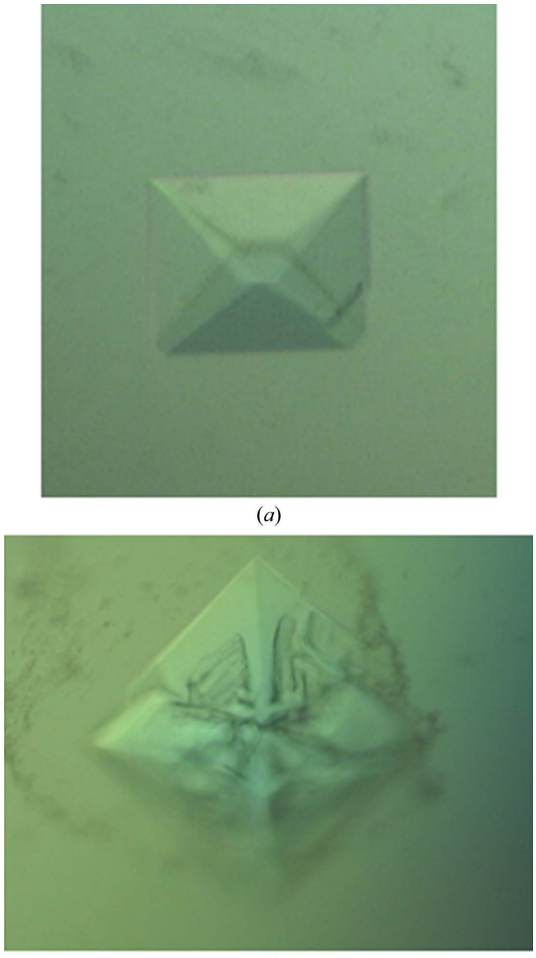


Figure 1
Purification of the ST1022 protein. SDS–PAGE of the purified ST1022 protein. Lane 1, molecular-weight markers (values are in kDa); lane 2, total lysate after sonication; lane 3, precipitate after sonication; lane 4, precipitate from ammonium sulfate precipitation; lane 5, supernatant of ammonium sulfate precipitation; lane 6, fractions of Resource ISO column washing buffer after loading; lane 7, fractions from Resource ISO column; lane 8, fractions after Resource Phe1; lane 9, fractions from hydroxyapatite column; lane 10, fractions from gel-filtration column.

2.2. Crystallization and data collection

Preliminary screening of crystallization conditions was carried out using Hampton Crystal Screen kits. We used Cryschem sitting-drop plates (Hampton Research) for crystallization and aliquoted 500 µl of the Crystal Screen solution into the wells. Initial crystals of native ST1022 were grown at 293 K by the sitting-drop vapour-diffusion method (McPherson, 1990) by adding 0.5 µl protein solution to 0.5 µl well solution consisting of 30% 2-propanol, 0.2 *M* sodium citrate and 0.1 *M* sodium cacodylate pH 6.5 (condition No. 8). Native crystals of ST1022 grew within a week. To make the effector complex, the protein was crystallized in the presence of 33 m*M* Gln. We tried several cryoprotectants, including PEG 400, MPD, glycerol and ethylene glycol, and found that trehalose worked well for this protein. The native and putative complex crystals were soaked into precipitant buffer plus 30% trehalose for cryoprotection and data sets were obtained at 100 K using a Jupiter210 CCD detector (Rigaku) on the RIKEN Structural Genomics Beamline II (BL26B2) at SPring-8, Hyogo, Japan. All crystal data were processed with the *HKL*-2000 program suite (Otwinowski & Minor, 1997). The crystals belong to the tetragonal space group *I*422. Crystal data statistics are provided in Table 1.



(*a*)

(*b*)

Figure 2
Representatives of ST1022 crystals grown using the sitting-drop method (*a*) in the absence and (*b*) in the presence of Gln.

3. Results and discussion

The Lrp/AsnC family of transcriptional regulatory proteins are involved in cellular metabolism, with exogenous amino acids functioning as the effectors. To determine the crystal structure of ST1022 in the presence and absence of its effector, we expressed it in *E. coli* using the vector pET-21a(+) and purified the protein as described in §2. The purified protein showed over 99% homogeneity, as verified by SDS–PAGE (Fig. 1). N-terminal sequencing of the purified protein confirmed that it was the expected protein (unpublished data). We crystallized the ST1022 protein in the absence and presence of the ligand using the sitting-drop vapour-diffusion method (Fig. 2) as described in §2 and the crystals grew to dimensions of $0.1 \times 0.1 \times 0.05$ and $0.3 \times 0.3 \times 0.05$ mm, respectively. Complete data sets were obtained for these crystals, as shown in Table 1. The asymmetric units of both the native and putative complex crystals of ST1022 gave specific volumes (V_M) of 2.82 and 2.85 Å^3 Da^{-1} with approximate solvent contents of 56 and 57%, respectively; these V_M values were well within the observed ranges for protein crystals (Matthews, 1968). The native ST1022 structure and the putative ligand-complex structure were determined by the molecular-replacement method using the structure of FL11 (PDB code 1ri7) from *Pyrococcus* sp. OT3 as a search model (Koike *et al.*, 2004). The solution was found using *AUTO-MOLREP* from the *CCP*4 program suite and refinement was carried out using *CNS* (Brünger *et al.*, 1998). A total of 5% of the reflections were used for the calculation of R_{free} (Brünger, 1992). The R and R_{free} values of the partially refined structures were 36.6% and 37.5% for native ST1022 and 38.1% and 40.0% for the putative ligand complex, respectively. Interestingly, in the putative ligand-complex structure we could observe clear density for the bound Gln and further structural refinement with the Gln is in progress (Kumarevel *et al.*, unpublished data).

The authors would like to thank C. Kuroishi for cloning. This work was supported by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Bell, S. D. & Jackson, S. P. (2001). *Curr. Opin. Microbiol.* **4**, 208–213.
Brinkman, A. B., Dahlke, I., Tuininga, J. E., Lammers, T., Dumay, V., de Heus, E., Lebbink, J. H. G., Thomm, M., de Vos, W. M. & van der Oost, J. (2000). *J. Biol. Chem.* **275**, 38160–38169.
Brinkman, A. B., Ettema, T. J., de Vos, W. M. & van der Oost, J. (2003). *Mol. Microbiol.* **48**, 287–294.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Calvo, J. M. & Matthews, R. G. (1994). *Microbiol. Rev.* **58**, 466–490.
Chen, S. & Calvo, J. M. (2002). *J. Mol. Biol.* **318**, 1031–1042.
Chen, S., Rosner, M. H. & Calvo, J. M. (2001). *J. Mol. Biol.* **312**, 625–635.
Corpet, F. (1998). *Nucleic Acids Res.* **16**, 10881–10890.
Geiduschek, E. P. & Ouhammouch, M. (2005). *Mol. Microbiol.* **56**, 1397–1407.
Jafri, S., Evoy, S., Cho, K., Craighead, H. G. & Winans, S. C. (1999). *J. Mol. Biol.* **288**, 811–824.
Koike, H., Ishijima, S. A., Clowney, L. & Suzuki, M. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 2840–2845.
Leonard, P. M., Smits, S. H., 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.
McPherson, A. (1990). *Eur. J. Biochem.* **189**, 1–23.
Madhusudhan, K. T., Huang, N. & Sokatch, J. R. (1995). *J. Bacteriol.* **177**, 636–641.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Newman, E. B. & Lin, R. (1995). *Annu. Rev. Microbiol.* **49**, 747–775.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Ouhammouch, M. (2004). *Curr. Opin. Genet. Dev.* **14**, 133–138.
Suzuki, M. (2003). *Proc. Jpn Acad. B*, **79**, 274–289.
Thaw, P., Sedelnikova, S. E., Muranova, T., Wiese, S., Sylvia, A., Alonso, J. C., Brinkman, A. B., Akerboom, J., van der Oost, J. & Rafferty, J. B. (2006). *Nucleic Acids Res.* **34**, 1439–1449.
Willins, D. A., Ryan, C. W., Platko, J. V. & Calvo, J. M. (1991). *J. Biol. Chem.* **266**, 10768–10774.