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

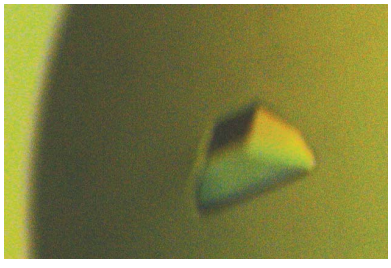
Crystallization and preliminary X-ray diffraction analysis of the arginine repressor of the hyperthermophile *Thermotoga neapolitana*

Jan Massant,[a]* Eveline Peeters,[a] Daniel Charlier[a] and Dominique Maes[b]

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

Correspondence e-mail: jan.massant@vub.ac.be

Received 26 September 2005
Accepted 28 November 2005
Online 16 December 2005



The arginine repressor of *Thermotoga neapolitana* (ArgRTnp) is a member of the family of multifunctional bacterial arginine repressors involved in the regulation of arginine metabolism. This hyperthermophilic repressor shows unique DNA-binding features that distinguish it from its homologues. ArgRTnp exists as a homotrimeric protein that assembles into hexamers at higher protein concentrations and/or in the presence of arginine. ArgRTnp was crystallized with and without its corepressor arginine using the hanging-drop vapour-diffusion method. Crystals of the aporepressor diffracted to a resolution of 2.1 Å and belong to the orthorhombic $P2_12_12_1$ space group, with unit-cell parameters $a = 117.73$, $b = 134.15$, $c = 139.31$ Å. Crystals of the repressor in the presence of its corepressor arginine diffracted to a resolution of 2.4 Å and belong to the same space group, with similar unit-cell parameters.

1. Introduction

The arginine repressor (ArgR) regulates the transcription of arginine biosynthetic genes in Gram-negative and Gram-positive bacteria from psychrophiles to hyperthermophiles (for recent surveys, see Charlier & Glansdorff, 2004; Charlier, 2004). Moreover, ArgR proteins play a multifunctional role in the bacterial cell. They inhibit biosynthetic promoters and are involved in the activation of several catabolic pathways that, depending on the organism and growth conditions, are applied to use arginine as a source of nitrogen, carbon or energy (Charlier & Glansdorff, 2004; Larsen *et al.*, 2005). The arginine repressor of *Streptomyces clavuligerus* participates in the regulation of clavulanic acid production (Ludovice *et al.*, 1992). Moreover, the *Escherichia coli* repressor is involved in the site-specific DNA-recombination mechanism that resolves multimeric forms of the ColE1 plasmid into its monomeric constituents (Stirling *et al.*, 1988; Colloms *et al.*, 1996).

The three-dimensional structures of the arginine repressors from *E. coli* (ArgREco; Van Duyne *et al.*, 1996; Sunnerhagen *et al.*, 1997), *Bacillus stearothermophilus* (ArgRBst; Ni *et al.*, 1999) and *B. subtilis* (AhrC; Dennis *et al.*, 2002) have been solved. Notwithstanding their low amino-acid sequence identity, they all have the same fold: an N-terminal DNA-binding domain is connected *via* a short linker to a C-terminal domain involved in the binding of arginine and oligomerization. The oligomerization domain has an α/β fold containing three α-helices packed against a four-stranded antiparallel β-sheet. Arginine repressors are hexamers composed of a dimer of trimers. Six arginine molecules bind to the trimer–trimer interfaces and act as a molecular glue consolidating the hexameric form. Whereas ArgREco and AhrC maintain their hexameric structures in the absence of arginine, ArgRBst exists as trimers that assemble into hexamers at higher protein concentrations and in the presence of arginine or DNA (Dion *et al.*, 1997). The N-terminal domain belongs to the winged helix–turn–helix (wHTH) type of DNA-binding domain. This fold includes a three-helix bundle flanked by wings formed by β-strands and loops (Brennan, 1993).

Arginine repressors bind as hexamers to a tandem pair of ARG boxes, each consisting of a 18 bp imperfect palindromic sequence. The repressor makes contacts with two major-groove segments and the intermediate minor-groove segment of each box at the same face of the DNA helix (Wang *et al.*, 1998; Song *et al.*, 2002; Xu *et al.*, 2003;

Table 1
Crystallization conditions of the *T. neapolitana* aporepressor (ArgRTnp) and of the repressor in the presence of 10 m*M* L-arginine (ArgRTnp + L-Arg).

	Precipitant	Salt	Buffer	Time†
ArgRTnp				
a		0.8 *M* K,Na tartrate	0.1 *M* Na HEPES pH 7.5	9 weeks
b		2.0 *M* Na formate	0.1 *M* Na acetate pH 4.6	9 weeks
c	30% PEG 400	0.2 *M* $CaCl_2$	0.1 *M* Na HEPES pH 7.5	1 week
d	28% PEG 400	0.2 *M* $CaCl_2$	0.1 *M* Na HEPES pH 7.5	1 week
ArgRTnp + L-Arg				
e	30% PEG 400	0.2 *M* $CaCl_2$	0.1 *M* Na HEPES pH 7.5	1 week
f	30% Jeffamine M-600	0.05 *M* CsCl	0.1 *M* Na HEPES pH 7.0	2 weeks
g	30% Jeffamine M-600	0.1 *M* CsCl	0.1 *M* Na HEPES pH 7.0	2 weeks

† Time before first crystals appeared.

Morin *et al.*, 2003). ArgR proteins and their binding sites are remarkably well conserved among diverse bacteria, better than any other bacterial mechanism of regulation of amino-acid or nucleotide biosynthesis (Smith *et al.*, 1989; Miller *et al.*, 1997). Arginine-specific repression and activation provides a good model system to investigate the evolution of regulatory proteins and networks. In all investigated cases the ArgR-binding site overlaps the promoter of biosynthetic genes and operons. Therefore, it is proposed that repression acts by steric exclusion of the RNA polymerase by the bound ArgR, as demonstrated for *E. coli* ArgR at the *car*P2 promoter (Charlier *et al.*, 1988).

The arginine repressor of the Gram-negative hyperthermophile *Thermotoga neapolitana* (ArgRTnp) exhibits some unique features that clearly distinguish it from previously studied homologues (Dimova *et al.*, 2000; Song *et al.*, 2002; Morin *et al.*, 2003; Charlier, 2004). Its DNA-binding activity is nearly arginine-independent and shows a broad sequence specificity. ArgRTnp makes essentially strong contacts with only one ARG box of the operator and it has the remarkable capacity to also bind to heterologous operators and single ARG-box fragments. ArgRTnp was purified as a homotrimeric protein of 49.0 kDa that assembles into hexamers at higher protein concentrations and/or in the presence of arginine (Dimova *et al.*, 2000; Song *et al.*, 2002).

Here, we report the crystallization of ArgRTnp with and without its corepressor arginine. The structure of ArgRTnp is essential to elucidate the mechanisms of molecular regulation and thermostability of this hyperthermophilic transcription regulator. Comparison with the homologous repressors of mesophilic and psychrophilic bacteria will contribute to the understanding of the evolution of the arginine repressor in the bacterial domain and the adaptation of protein–DNA interactions to extremes of temperature.

2. Materials and methods

2.1. Protein purification

ArgRTnp was expressed *via* the pET24a expression vector (Novagen) in an *E. coli* BL21 (DE3) RIL host (Novagen). A 5 l culture was grown at 303 K in rich medium supplemented with kanamycin and chloramphenicol. Overexpression of ArgRTnp was induced at a cell density of 10^9 cells ml^{-1} by adding 1.0 m*M* IPTG, followed by overnight growth at 293 K. Cells were collected by centrifugation, resuspended in 25 ml extraction buffer (20 m*M* Tris–HCl pH 8.0) and disrupted by sonication for 30 min with a Vibra cell sonciator (Bioblock Scientific). Cell debris was removed by centrifugation at 45 000*g* for 1 h. The cell-free extract was brought to 0.5%(*w*/*v*) in polyethyleneimine (PEI) by slow addition with gentle stirring of 0.05 ml 10% PEI per millilitre of extract. Precipitated nucleic acids were removed by centrifugation at 12 000*g* for 15 min. For the further purification of this hyperthermophilic repressor, advantage was taken of the thermal stability of the protein (Song *et al.*, 2002). A thermal treatment of 20 min at 348 K was followed by centrifugation of the denatured *E. coli* proteins. The supernatant was submitted to ion-exchange chromatography on a 6 ml Resource S column (Amersham Biosciences) and eluted with a 0–500 m*M* NaCl linear gradient. Pooled fractions were further purified by size-exclusion chromatography on a Superose P12 HR 10/30 column equilibrated with extraction buffer containing 200 m*M* NaCl. ArgRTnp eluted as a homotrimeric protein with an apparent molecular weight of 49.0 kDa (Song *et al.*, 2002). The polypeptide chain consists of 152 amino acids and migrates as a 16.5 kDa protein on SDS–PAGE (Song *et al.*, 2002). The purified protein was concentrated to 8 mg ml^{-1} in 0.05 *M* Tris–HCl, 200 m*M* NaCl or 0.05 *M* Tris–HCl, 200 m*M* NaCl, 10 m*M* L-arginine for crystallization of the repressor in the absence or presence of its corepressor, respectively.

2.2. Crystallization

The initial crystallization conditions were screened using the hanging-drop vapour-diffusion method with sparse-matrix crystallization kits (Crystal Screen and Crystal Screen 2; Hampton Research, Riverside, CA, USA). Immediately prior to setting up crystallization, the protein was centrifuged at 13 000 rev min^{-1} for 10 min in a standard tabletop centrifuge to remove precipitate. 2 µl protein solution (8 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 between one and nine weeks (Table 1).

2.3. X-ray diffraction data collection and processing

X-ray diffraction data were collected at 100 K using a MAR CCD detector (MAR Research, Norderstadt, Germany) at beamline BW7A at the DORIS storage ring, DESY, Hamburg. 400 images ($\Delta\varphi = 0.5^\circ$) and 300 images ($\Delta\varphi = 0.6^\circ$) were collected for a single crystal of the repressor with and without arginine, respectively, at a wavelength of 0.9116 Å. The images were processed using the *HKL* programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

2.4. Molecular replacement

Molecular-replacement searches were performed with the *PHASER* program from the *CCP*4 suite (McCoy *et al.*, 2005) using the structures of ArgREco (PDB codes 1xxb and 1xxc), ArgRBst (PDB code 1b4a) and AhrC (PDB code 1f9n) as search models. A model of ArgRTnp based on ArgRBst and AhrC was obtained using *SWISS-PDBVIEWER* and the *SWISS-MODEL* server (http://swissmodel.expasy.org/). Amino-acid sequence identity and similarity was determined using the *STRETCHER* program from the *wEMBOSS* suite of programs at BEN (Belgian EMBnet Node, http://www.be.embnet.org).

3. Results and discussion

ArgRTnp was crystallized in the absence and the presence of 10 m*M* L-arginine. Initial screening was performed with sparse-matrix crystallization kits. Microcrystals appeared within one week in several conditions of the sparse-matrix screen. After optimization, small (<0.2 mm) crystals of ArgRTnp formed in conditions *c* and *d* in the absence of arginine and in conditions *e*, *f* and *g* in the presence of arginine (Table 1). Large (>0.2 mm) crystals appeared in two

Table 2
Data-collection statistics.

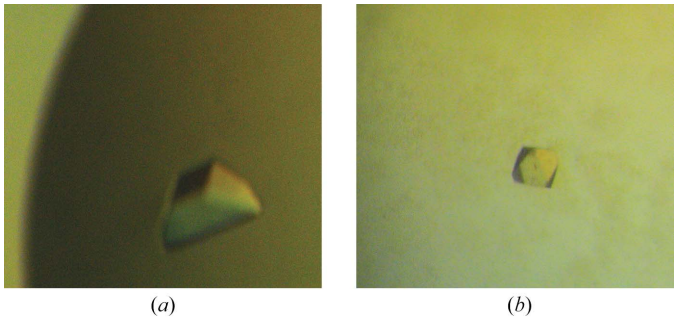
ArgRTnp, crystals of the aporepressor (Table 1 condition *a*); ArgRTnp + L-Arg, crystals of the repressor grown in the presence of arginine (Table 1 condition *e*). Values in parentheses correspond to the highest resolution shell.

	ArgRTnp	ArgRTnp + L-Arg
Unit-cell parameters (Å, °)	a = 117.7, b = 134.2, c = 139.3, $\alpha = \beta = \gamma = 90$	a = 117.4, b = 134.1, c = 139.0, $\alpha = \beta = \gamma = 90$
Space group	$P2_12_12_1$	$P2_12_12_1$
X-ray source	DESY BW7A	DESY BW7A
Wavelength (Å)	0.9116	0.9116
Resolution (Å)	30–2.1 (2.18–2.10)	30–2.4 (2.49–2.40)
Temperature (K)	100	100
No. of observations	633540	723316
No. of unique reflections	124003 (11509)	86327 (8500)
Completeness (%)	96.0 (89.9)	100.0 (99.8)
Redundancy	5.1 (4.5)	8.4 (8.3)
Average $I/\sigma(I)$	11.2 (4.4)	14.0 (7.5)
R_{merge} (%)	7.9 (28.4)	8.7 (23.5)
Mosaicity (°)	0.8	0.4

conditions of the sparse-matrix screen (Table 1, conditions *a* and *b*) after nine weeks.

A data set to a resolution of 2.1 Å was collected from a crystal of the aporepressor grown from condition *a* (Table 1; Fig. 1*a*). 30% glycerol was used as a cryoprotectant. The crystals belong to the orthorhombic $P2_12_12_1$ space group, with unit-cell parameters a = 117.7, b = 134.2, c = 139.3 Å. Data-collection information and statistics are given in Table 2. The value of the Matthews coefficient V_M is 2.8 Å^3 Da^{-1} assuming the presnece of two hexamers or four trimers in the asymmetric unit. The corresponding value of the solvent content is 55.8%. Crystals of the repressor in the presence of its corepressor arginine in condition *e* (Table 1; Fig. 1*b*) appeared faster than those for the aporepressor but had smaller dimensions. A data set to a resolution of 2.4 Å was collected from a single crystal. The crystals belong to the same space group as above, with unit-cell parameters a = 117.4, b = 134.1, c = 139.0 Å. Data-collection information and statistics are given in Table 2. Comparison of both data sets resulted in an R_{merge} of 31% (41.1% in the 2.49–2.40 Å resolution shell). This suggests that the corepressor is effectively bound to the aporepressor in this crystal.

Notwithstanding the low sequence identity between the arginine repressors of different bacteria, all the solved structures of arginine repressors have a similar fold. The amino-acid sequence of ArgRTnp is 36.2% identical (58.6% similar) to ArgRBst, 38.2% identical (58.6% similar) to ArgRBsu and 33.1% identical (50.6% similar) to ArgREco. Slightly higher sequence identities are found for the separate oligomerization domains. Attempts at molecular replacement using the arginine repressors of *B. stearothermophilus*, *B. subtilis* and *E. coli* as models (Van Duyne *et al.*, 1996; Sunnerhagen *et al.*, 1997; Ni *et al.*, 1999; Dennis *et al.*, 2002) have not been successful. Notwithstanding the 70% sequence identity between AhrC and ArgRBst, the structure of the DNA-binding domains of AhrC had to be solved using heavy-metal derivatives (Dennis *et al.*, 2002). The structure of the oligomerization core was solved by molecular replacement using the structure of ArgRBst as a model. It turns out that the DNA-binding domains of AhrC adopt slightly different positions around the oligomerization core and deviate from strict NCS. In order to solve the structure of ArgRTnp, a selenomethionine derivative is being prepared. Because the ArgRTnp polypeptide chain contains only one methionine, bromide soaking will also be applied for phasing.



(*a*) (*b*)

Figure 1
Crystals used for X-ray diffraction: (*a*) crystal of ArgRTnp grown in condition *a* (Table 1), (*b*) crystal of ArgRTnp grown in the presence of 10 m*M* L-arginine in condition *e* (Table 1).

JM is Postdoctoraal Onderzoeker of the FWO-Vlaanderen. This research was financed by a PhD grant from IWT-Vlaanderen to EP, grants from the FWO-Vlaanderen, the Research Council of the Vrije Universiteit Brussel, from the Flanders Interuniversity Institute for Biotechnology (VIB) and the Vlaamse Gemeenschapscommissie. The authors gratefully acknowledge the use of synchrotron beamtime at the EMBL beamline BW7A at the DORIS storage ring, DESY, Hamburg.

References

Brennan, R. G. (1993). *Cell*, **74**, 773–776.
Charlier, D. (2004). *Biochem. Soc. Trans.* **32**, 310–313.
Charlier, D. & Glansdorff, N. (2004). *Biosynthesis of Arginine and Polyamines.* In *EcoSal: Escherichia coli and Salmonella: Cellular and Molecular Biology*, edited by R. Curtiss III. Washington, DC: ASM Press. http://www.ecosal.org.
Charlier, D., Weyens, G., Roovers, M., Piette, J., Bocquet, C., Piérard, A. & Glansdorff, N. (1988). *J. Mol. Biol.* **204**, 867–877.
Colloms, S., McCulloch, R., Grant, K., Neilson, L. & Sherratt, D. (1996). *EMBO J.* **15**, 1172–1181.
Dennis, C. A., Glykos, N. M., Parsons, M. R. & Philips, S. E. V. (2002). *Acta Cryst.* D**58**, 421–430.
Dimova, D., Weigel, P., Takahashi, M., Marc, F., Van Duyne, G. D. & Sakanyan, V. (2000). *Mol. Gen. Genet.* **263**, 119–130.
Dion, M., Charlier, D., Wang, H., Gigot, D., Savchenko, A., Hallet, J.-N., Glansdorff, N. & Sakanyan, V. (1997). *Mol. Microbiol.* **25**, 385–398.
Larsen, R., Kok, J. & Kuipers, O. P. (2005). *J. Biol. Chem.* **280**, 19319–19330.
Ludovice, M., Martin, J. F., Carrachas, P. & Liras, P. (1992). *J. Bacteriol.* **174**, 4606–4613.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Miller, C. M., Baumburg, S. & Stockley, P. G. (1997). *Mol. Microbiol.* **26**, 37–48.
Morin, A., Huysveld, N., Braun, F., Dimova, D., Sakanyan, V. & Charlier, D. (2003). *J. Mol. Biol.* **332**, 537–553.
Ni, J., Sakanyan, V., Charlier, D., Glansdorff, N. & Van Duyne, G. D. (1999). *Nature Struct. Biol.* **6**, 427–432.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Smith, M. C. M., Czaplewski, L., North, A. K., Baumberg, S. & Stockley, P. G. (1989). *Mol. Microbiol.* **3**, 23–28.
Song, H., Wang, H., Gigot, D., Dimova, D., Sakanyan, V., Glansdorff, N. & Charlier, D. (2002). *J. Mol. Biol.* **315**, 255–274.
Stirling, C. J., Szatmari, G., Stewart, G., Smith, M. C. M. & Sherratt, D. (1988). *EMBO J.* **7**, 4389–4395.
Sunnerhagen, M., Nilges, M., Otting, G. & Carey, J. (1997). *Nature Struct. Biol.* **4**, 819–826.
Van Duyne, G. D., Ghosh, G., Maas, W. K. & Sigler, P. B. (1996). *J. Mol. Biol.* **256**, 377–391.
Wang, H., Glansdorff, N. & Charlier, D. (1998). *J. Mol. Biol.* **277**, 805–824.
Xu, Y., Sun, Y., Huysveld, N., Gigot, D., Glansdorff, N. & Charlier, D. (2003). *J. Mol. Biol.* **326**, 353–369.