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

Crystallization and preliminary X-ray crystallographic study of the wild type and two mutants of the CP1 hydrolytic domain from *Aquifex aeolicus* leucyl-tRNA synthetase

Vincent Cura,[a] Natacha Olieric,[a] Alexandre Guichard,[a] En-Duo Wang,[b] Dino Moras,[a] Gilbert Eriani[c] and Jean Cavarelli[a]*

[a]Département de Biologie et Génomique Structurales, UMR 7104, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP Strasbourg, 1 Rue Laurent Fries, 67404 Illkirch, France, [b]State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, The Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, People's Republic of China, and [c]Architecture et Réactivité de l'ARN, UPR 9002, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 Rue René Descartes, 67084 Strasbourg, France

Correspondence e-mail: cava@igbmc.u-strasbg.fr

Received 21 July 2005
Accepted 9 September 2005
Online 13 September 2005

The editing or hydrolytic CP1 domain of leucyl-tRNA synthetase (LeuRS) hydrolyses several misactivated amino acids. The CP1 domain of *Aquifex aeolicus* LeuRS was expressed, purified and crystallized by the hanging-drop vapour-diffusion method using ammonium sulfate as precipitant. Crystals belong to space group $P2_12_12_1$, with unit-cell parameters a = 38.8, b = 98.4, c = 116.7 Å. Crystals diffract to beyond 1.8 Å resolution and contain two monomers in the asymmetric unit. Two CP1 mutants in which a conserved threonine residue essential for the fidelity of the hydrolytic pathway is mutated to alanine or glutamic acid have also been expressed and crystallized. Crystals of the two CP1 mutants are isomorphs of the wild type and diffract to beyond 1.9 Å resolution. All structures were solved by molecular-replacement techniques.

1. Introduction

Aminoacyl-tRNA synthetases (AARSs) form a family of RNA-binding proteins which catalyse the attachment of specific amino acids to their cognate tRNA in a two-step reaction that is crucial for protein biosynthesis. The first step, amino-acid activation, leads to the formation of an intermediate, aminoacyl-adenylate, from amino acids, ATP and magnesium ions. In the second step, the aminoacyl moiety is transferred to one of the two hydroxyl groups of the 3′-terminal adenosine of the tRNA to form an aminoacyl-tRNA. In most organisms, there are 20 distinct AARSs, each responsible for the correct loading of a tRNA with its corresponding amino acid.

AARSs constitute a textbook example of multidomain proteins including insertion and terminal functional modules appended to one of the two class-specific active-site domains. The non-catalytic domains usually play distinct roles in the aminoacylation and are responsible for, or at least contribute to, the specificity of the tRNA recognition and/or the efficiency of aminoacylation (Martinis *et al.*, 1999). The accuracy of the aminoacylation reaction is essential for the fidelity of protein synthesis and several AARSs use two active sites for amino-acid discrimination by a double-sieve mechanism which prevents the production of tRNAs carrying non-cognate amino acids (Fersht, 1977). The 'synthetic' active site, in which the cognate amino acid is usually activated, converted to aminacyl-adenylate and transferred to the tRNA, adopts one of two class-specific topologies (Eriani *et al.*, 1990). Some class I and class II 'synthetic' sites are prone to errors which are corrected in the 'hydrolytic' active site, located in an additional domain, by a mechanism termed 'editing' (Jakubowski, 1978). Corrections can occur at the level of the mis-activated aminoacyl adenylates ('pre-transfer' editing) or on mis-acylated tRNA ('post-transfer' editing). To date, the editing function has been well documented for three class I synthetases [IleRS (Nureki *et al.*, 1998; Silvian *et al.*, 1999), LeuRS (Lincecum *et al.*, 2003) and ValRS (Fukai *et al.*, 2000)] and to a lesser extent for four class II synthetases [AlaRS (Beebe *et al.*, 2003), ThrRS (Dock-Bregeon *et al.*, 2000), PheRS (Roy *et al.*, 2004) and ProRS (Beuning & Musier-Forsyth, 2000)].

Aquifex aeolicus LeuRS is a heterodimer of $\alpha\beta$-type (Deckert *et al.*, 1998; Gouda *et al.*, 2002; Xu *et al.*, 2002). The α-subunit of *A. aeolicus* LeuRS (634 amino acids) contains most of the class I

active site (the so-called Rossmann fold) and a large insertion domain, termed the connective peptide 1 (CP1) domain, which is involved in the editing activity of LeuRS. The β-subunit of *A. aeolicus* LeuRS (289 amino acids) is responsible for the interaction with the anticodon loop and arm of the $tRNA^{Leu}$ substrate. The CP1 domain of IleRS, LeuRS and ValRS adopts a β-barrel topology reminiscent of that of the acid proteases (Nureki *et al.*, 1998; Silvian *et al.*, 1999; Fukai *et al.*, 2000; Lincecum *et al.*, 2003; Fukunaga *et al.*, 2004; Fukunaga & Yokoyama, 2005). LeuRS generally misactivates isoleucine, methionine, valine and several nonstandard metabolic intermediates such as norvaline and norleucine. Functional studies have shown that the isolated CP1 domain of *A. aeolicus* LeuRS catalyses the hydrolytic editing of both mischarged $tRNA^{Leu}$ and $minihelix^{Leu}$, an editing capability which is not shared with LeuRS-CP1 domains from other organisms (Zhao *et al.*, 2005). It has been suggested that a 20-amino-acid insertion peptide specific to the *A. aeolicus* CP1 domain might be responsible for the activity of the isolated CP1 domain (Zhao *et al.*, 2005). Therefore, the first aim of this work was to build a framework for a structural explanation of this unusual feature.

In the present study, the *A. aeolicus* CP1 domain (residues 225–443) was expressed in *Escherichia coli* cells, purified and crystallized. We also expressed, purified and crystallized two *A. aeolicus* CP1 mutants in which a highly conserved threonine residue in the 'T-rich' region (Mursinna *et al.*, 2001), Thr273 (Thr252 in *E. coli* numbering), was mutated to alanine (T273A) or glutamic acid (T273E). Extensive biochemical analyses on *E. coli* LeuRS and X-ray structures of *Thermus thermophilus* IleRS, ValRS and LeuRS have shown that the 'T-rich' region (residues 247–252 for *E. coli* LeuRS) is part of the 'hydrolytic' active site and is crucial for the discrimination of amino acids (Mursinna *et al.*, 2001, 2004; Mursinna & Martinis, 2002; Tang & Tirrell, 2002; Xu *et al.*, 2004). In *E. coli* LeuRS, the T273A mutation yielded an enzyme that efficiently hydrolyses the correct Leu-$tRNA^{Leu}$ product (Mursinna *et al.*, 2001), while the T273E mutations leads to mischarged $tRNA^{Leu}$ (Xu *et al.*, 2004). The second goal of this work was therefore to propose a detailed structural explanation for those effects on the editing activity of LeuRS.

2. Methods and results

2.1. Overexpression and purification of wild-type and mutant CP1 domains of *A. aeolicus* LeuRS

The isolated CP1 domain from *A. aeolicus* was prepared by cloning a gene fragment corresponding to residues 225–443 from LeuRS into the *Nde*I–*Bam*HI restriction sites of pET11c vector (Novagen). The gene fragment was excised from the native LeuRS α-subunit gene by PCR amplification. A DNA fragment encoding residues Gly225–Leu443 was amplified with two oligonucleotides (oligonucleotides 'NdeI', 5′-GGAACCATATGGGAC-GCTCCGAAGGAGCC, and 'stop', 5′-ATGTTCGGATCCTATAATCTGTAGGAGACTT). The oligonucleotides were designed to create an *Nde*I cloning site in front of Gly225 and a *Bam*HI site after the stop codon introduced at codon 444. The 0.66 kbp amplified DNA fragment was digested with *Nde*I and *Bam*HI and was ligated into the same sites of pET11c. Creation of an *Nde*I site resulted in the addition of a Met residue at the first position of the CP1 domain. The DNA sequence from an over-producing clone was entirely verified by sequencing. CP1 domains carrying mutations of residue 273 were constructed by PCR using the Quik Change Site Directed Mutagenesis kit from Stratagene. The oligonucleotides were the following: T273A, 5′-CGAGACCCGA-CGCGGTTTTCGGGGCT and 5′-AGCCCCGAAAACCGCGTCG-GGTCTCG; TE273, 5′-CGAGACCCGACGAAGTTTCGGGGCT and 5′-AGCCCCGAAAACTTCGTCGGGCTCG. The sequences of the entire domains were verified by DNA sequencing.

Wild-type and mutant CP1 domains (T273A and T273E) have been expressed and purified with a similar protocol. Strains over-expressing wild-type and mutant CP1 domains were obtained by transformation of *E. coli* Rosetta (DE3) cells (Novagen). Cells were grown at 310 K in Luria–Bertani broth supplemented with 0.1 g l^{-1} ampicillin and 0.034 g l^{-1} chloramphenicol. CP1 expression was induced with 0.5 m*M* isopropyl β-D-thiogalactoside on exponentially growing culture. Cultures were harvested 3 h after induction. Cells were resuspended in 50 m*M* Tris–HCl, 20 m*M* sodium chloride, 1 m*M* magnesium chloride pH 8.5 and sonicated on ice. The lysate was clarified by centrifugation (25 000*g*, 40 min) and heated at 348 K for 30 min. Flocculated *E. coli* proteins were removed by a centrifugation step (25 000*g*, 10 min) and a Q-Sepharose (Amersham Bioscience) chromatography step was performed on the heated extract. A sodium chloride concentration gradient (20–300 m*M*) was loaded onto the column and CP1 was eluted at 100 m*M* sodium chloride. The fraction containing CP1 was mixed with two volumes of 4 *M* ammonium sulfate and applied onto a Phenyl Toyopearl 650S column (Tosoh) equilibrated in 1.3 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 7.5. A decreasing gradient of ammonium sulfate (1.3–0 *M*) was applied to the column and CP1 was eluted at 1.1 *M* ammonium sulfate. The purity of the protein was followed at each step by SDS–PAGE. Prior to crystallization trials, pure CP1 proteins were concentrated to 40 mg ml^{-1} using Centriprep YM-10 (Amicon). During the concentration step, the solution was progressively exchanged to 20 m*M* sodium chloride, 20 m*M* Tris–HCl pH 8.5. 8 mg of pure wild-type CP1, 6 mg of CP1 mutant T273A and 4 mg of CP1 mutant T273E were obtained from 1 l of culture.

2.2. Crystallization and X-ray data collection

Initial crystallization conditions for wild-type CP1 were found using the hanging-drop method by screening several precipitating agents such as ammonium sulfate, polyethylene glycol (PEG) and 2-methyl-2,4-pentanediol (MPD) at different concentrations and pH values and at two temperatures. The best crystals were obtained using ammonium sulfate as a precipitating agent and by mixing 2 µl protein solution (10 mg ml^{-1}) and 2 µl precipitant solution consisting of 1.9 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 8.5, 10 m*M* potas-

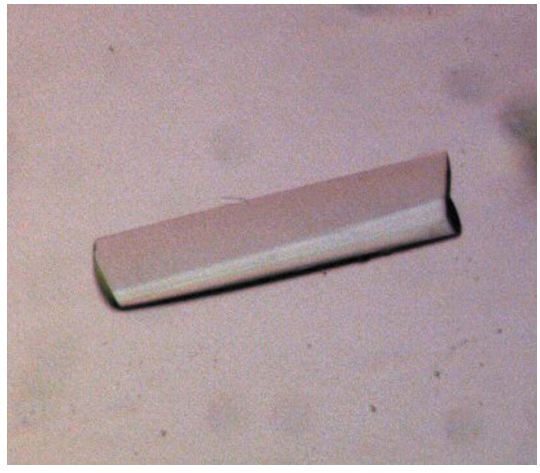

Figure 1
Wild-type crystals of *A. aeolicus* CP1 domain. Crystal dimensions are 350 × 100 × 50 µm.

Table 1
Data-collection statistics for wild-type (WT), T273E and T273A CP1 domains.

Values in parentheses are for the highest resolution shell.

	WT	WT + norvaline	T273E	T273A + Leu
Beamline station (ESRF)	ID14-1	ID29	ID29	ID14-3
Wavelength (Å)	0.934	0.9762	0.9762	0.931
Resolution range (Å)	1.77	2.2	1.85	1.9
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters				
a (Å)	38.7	38.7	38.6	38.6
b (Å)	98.2	98.0	98.2	98.0
c (Å)	116.7	117.1	117.1	117.1
Measured reflections	171518	127164	207516	228560
Unique reflections	43101	23152	38143	35735
Completeness (%)	99.3 (99.3)	98.7 (94.6)	97.0 (94.6)	99.2 (96.4)
$I/\sigma(I)$	40.1 (11.5)	21.8 (6.5)	26.7 (3.8)	28.1 (3.0)
Redundancy	4.0 (3.8)	5.5 (5.0)	5.4 (4.5)	6.4 (4.8)
$R_{\rm merge}$† (%)	3.3 (8.4)	6.0 (24.1)	5.6 (29.2)	5.2 (31.1)

† $R_{\rm merge}$ indicates the agreement of individual reflections over the set of unique averaged reflections. $R_{\rm merge} = \sum_h \sum_i |\langle I(h)\rangle - I_{h,i}| / \sum_h \sum_i I_{h,i}$, where $I_{h,i}$ is the ith observed intensity of a measured reflection of Miller index h and $\langle I_h\rangle$ is the average intensity for this unique reflection.

sium chloride, 1 m*M* calcium chloride, 1% PEG 400 and equilibrating against reservoirs containing 1 ml precipitant solution. Crystals grew to 350 × 100 × 50 µm within two weeks at 297 K (Fig. 1). Crystals of the T273A and T273E mutants were obtained using the wild-type conditions and reached maximum size (200 × 100 × 30 µm) after three weeks at 297 K. Prior to data collection, crystals were transferred into a cryoprotection solution consisting of 15%(*v*/*v*) glycerol, 2.1 *M* ammonium sulfate, 100 m*M* Tris–HCl pH 8.5, 10 m*M* potassium chloride, 1 m*M* calcium chloride, 1% PEG 400 for 30 s and then flash-cooled into liquid ethane.

X-ray diffraction data for wild-type CP1 were collected using a single cryocooled (100 K) crystal on beamline ID14-1 at the European Synchrotron Radiation Facility (ESRF). Crystals of wild-type CP1 diffracted to beyond 1.7 Å and a complete data set was obtained between 50 and 1.77 Å from 200 frames of 0.5° oscillation (exposure time 12 s per oscillation; crystal-to-detector distance 160 mm). The data were indexed and scaled with *HKL*2000 (Otwinowski & Minor, 1997). The crystals belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 38.8, b = 98.4, c = 116.7 Å (Table 1). The asymmetric unit contains two molecules of CP1, with a corresponding crystal volume per protein weight of 2.2 Å^3 Da^{-1} and a solvent content of 44.8% (assuming a partial specific volume of 0.74 cm^3 g^{-1}). X-ray diffraction data for cocrystals of the wild-type CP1 in complex with norvaline, obtained under similar crystallization conditions, were collected between 50 and 2.2 Å from 360 frames of 0.5° oscillation (exposure time 4 s per oscillation with an attenuated beam; crystal-to-detector distance 190 mm) on beamline ID29 at the ESRF.

X-ray diffraction data for the T273E mutant were collected using a single cryocooled (100 K) crystal at beamline ID29 at the ESRF. A complete data set was obtained between 50 and 1.85 Å from 360 frames of 0.5° oscillation (exposure time 3 s per oscillation with an attenuated beam; crystal-to-detector distance 145 mm).

X-ray diffraction data for the T273A mutant (in the presence of leucine) were collected using a single cryocooled (100 K) crystal at beamline ID14-3 at the ESRF. A complete data set was obtained between 50 and 1.9 Å from 360 frames of 0.5° oscillation (exposure time 18 s per oscillation; crystal-to-detector distance 135 mm). All the protein crystals (wild type, T273E, T273A) were isomorphs.

The structure of wild-type CP1 was solved by molecular-replacement methods using the CP1 domain of *T. thermophilus* LeuRS (Cusack *et al.*, 2000) as a probe with *MOLREP* (Collaborative Computational Project, Number 4, 1994). Structures of the two mutants and of the complex of wild-type CP1 with norvaline were obtained in a straightforward manner from the wild-type structure.

We thank members of the ESRF–EMBL Joint Structural Biology Group for use of the ESRF beamline facilities. We thank David Rose for technical assistance with cell cultures and Dr Andrew Atkinson for careful reading of the manuscript. The work described here was supported by funds from the Génopôle Alsace–Lorraine and the European Commission as SPINE contract No. QLG2-CT-2002-00988 under the Public Research, Technology and Development program 'Quality of Life and Management of Living Resources'.

References

Beebe, K., Merriman, E. & Schimmel, P. (2003). *J. Biol. Chem.* **278**, 45056–45061.
Beuning, P. J. & Musier-Forsyth, K. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 8916–8920.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cusack, S., Yaremchuk, A. & Tukalo, M. (2000). *EMBO J.* **19**, 2351–2361.
Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J. & Swanson, R. V. (1998). *Nature (London)*, **392**, 353–358.
Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Caillet, J., Springer, M., Rees, B., Francklyn, C. S., Ehresmann, C. & Moras, D. (2000). *Cell*, **103**, 877–884.
Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). *Nature (London)*, **347**, 203–206.
Fersht, A. R. (1977). *Biochemistry*, **16**, 1025–1030.
Fukai, S., Nureki, O., Sekine, S., Shimada, A., Tao, J., Vassylyev, D. G. & Yokoyama, S. (2000). *Cell*, **103**, 793–803.
Fukunaga, R., Fukai, S., Ishitani, R., Nureki, O. & Yokoyama, S. (2004). *J. Biol. Chem.* **279**, 8396–8402.
Fukunaga, R. & Yokoyama, S. (2005). *J. Mol. Biol.* **346**, 57–71.
Gouda, M., Yokogawa, T., Asahara, H. & Nishikawa, K. (2002). *FEBS Lett.* **518**, 139–143.
Jakubowski, H. (1978). *Biochim. Biophys. Acta*, **518**, 345–350.
Lincecum, T. L. Jr, Tukalo, M., Yaremchuk, A., Mursinna, R. S., Williams, A. M., Sproat, B. S., Van Den Eynde, W., Link, A., Van Calenbergh, S., Grotli, M., Martinis, S. A. & Cusack, S. (2003). *Mol. Cell*, **11**, 951–963.
Martinis, S. A., Plateau, P., Cavarelli, J. & Florentz, C. (1999). *Biochimie*, **81**, 683–700.
Mursinna, R. S., Lee, K. W., Briggs, J. M. & Martinis, S. A. (2004). *Biochemistry*, **43**, 155–165.
Mursinna, R. S., Lincecum, T. L. Jr & Martinis, S. A. (2001). *Biochemistry*, **40**, 5376–5381.
Mursinna, R. S. & Martinis, S. A. (2002). *J. Am. Chem. Soc.* **124**, 7286–7287.
Nureki, O., Vassylyev, D. G., Tateno, M., Shimada, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T. L., Schimmel, P. & Yokoyama, S. (1998). *Science*, **280**, 578–582.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Roy, H., Ling, J., Irnov, M. & Ibba, M. (2004). *EMBO J.* **23**, 4639–4648.
Silvian, L. F., Wang, J. & Steitz, T. A. (1999). *Science*, **285**, 1074–1077.
Tang, Y. & Tirrell, D. A. (2002). *Biochemistry*, **41**, 10635–10645.
Xu, M. G., Chen, J. F., Martin, F., Zhao, M. W., Eriani, G. & Wang, E. D. (2002). *J. Biol. Chem.* **277**, 41590–41596.
Xu, M. G., Li, J., Du, X. & Wang, E. D. (2004). *Biochem. Biophys. Res. Commun.* **318**, 11–16.
Zhao, M. W., Zhu, B., Hao, R., Xu, M. G., Eriani, G. & Wang, E. D. (2005). *EMBO J.* **24**, 1430–1439.