Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, expression, purification, crystallization and preliminary X-ray crystallographic investigations of a unique editing domain from archaebacteria

Shweta Dwivedi, Shobha P. Kruparani and Rajan Sankaranarayanan*

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India

Correspondence e-mail: sankar@ccmb.res.in

Received 16 June 2004
Accepted 15 July 2004

Threonyl-tRNA synthetase (ThrRS) faces a crucial double-discrimination problem during the translation of genetic code. Most ThrRSs from the archaeal kingdom possess a unique editing domain that differs from those of eubacteria and eukaryotes. In order to understand the structural basis of the editing mechanism in archaea, the editing module of ThrRS from *Pyrococcus abyssi* comprising of the first 183 amino-acid residues was cloned, expressed, purified and crystallized. The crystals belong to the trigonal space group $P3_{1(2)}21$, with one molecule in the asymmetric unit.

1. Introduction

Aminoacyl-tRNA synthetases (aaRSs) establish the rules of the genetic code by attaching the correct amino acid to their cognate-tRNA (Carter, 1993; Meinnel *et al.*, 1995; Schimmel & Söll, 1979). They are partitioned into two classes of ten enzymes each based on their active-site topology (Eriani *et al.*, 1990). The topology is reflected in the function of these enzymes such as, for example, the binding of small substrates, tRNA binding and the enzymatic mechanism. Recent structural and biochemical work on both classes of enzymes has provided valuable insights into how very closely related amino-acid substrates are discriminated with very high accuracy in order to maintain high fidelity during translation of the genetic code. IleRS, which is specific for isoleucine, was the first case in which a clear structural basis for double-sieve discrimination of valine from isoleucine was shown (Nureki *et al.*, 1998). A CP1 domain has been shown to be responsible for removing the non-cognate substrate attached to the tRNA (Lin *et al.*, 1996). Furthermore, the tRNA molecule has been shown to switch the conformation of the CCA end between a bent and helical conformation to reach the active site and the editing site, respectively (Silvian *et al.*, 1999). Subsequent work on other class I enzymes such as ValRS and LeuRS have provided further enlightenment on the mechanism of the double-sieve process.

Most of the efforts towards understanding the discrimination between closely related amino acids have been focused on class I enzymes. ThrRS is a class II enzyme and faces a crucial double-discrimination problem of selecting threonine from valine and serine. The structural work on the enzyme from *Escherichia coli* provided valuable insights into the function of the molecule (Sankaranarayanan *et al.*, 1999). The enzyme is made up of four domains: catalytic, anticodon binding and two N-terminal modules (N1 and N2). The N-terminal modules are linked to the catalytic core *via* a linker helix. The enzyme uses a unique zinc ion in its active site that is neither catalytic nor structural and plays a crucial role in recognition by selecting the cognate threonine substrate and thus rejecting valine from being activated by the enzyme (Sankaranarayanan *et al.*, 2000). The enzyme also possesses tRNA-mediated editing activity in one of its N-terminal modules, by use of which the incorrectly charged serine is removed from $tRNA^{Thr}$ (Dock-Bregeon *et al.*, 2000). Unlike the case of the CP1 domain of the class I enzymes, which is completely conserved through evolution, the N-terminal module of ThrRS shows significant variation in crucial residues, number of domains and also primary structure. It has previously been noted that archaeal ThrRSs possess an N-terminal module that shows no sequence homology not only to eubacterial and eukaryotic ThrRSs but also to any other protein of known sequence or structure (Dock-Bregeon *et al.*, 2000). Recently, it has been shown that this domain is indeed involved in removing the non-cognate serine attached to $tRNA^{Thr}$ (Beebe *et al.*, 2004). In order to determine the three-dimensional structure of this unique sequence and to further understand the editing mechanism, a structural investigation of the editing module from *Pyrococcus abyssi* ThrRS, consisting of the first 183 amino acids with a calculated molecular weight of 21.097 kDa, has been undertaken.

2. Materials and methods

2.1. Cloning and expression

The gene for the N-terminal module of ThrRS from *P. abyssi* comprising of amino acids 1–183 was amplified by PCR from

genomic DNA using gene-specific primers and was cloned into the *Nde*I and *BamH*I sites of the pET-21b (Novagen) expression vector. The recombinant plasmid was transformed and overexpressed in *Escherichia coli* BL21(DE3) strain cells. Cells were grown in Luria–Bertani medium with 100 µg ml^{-1} ampicillin at 310 K until OD_{600} reached 0.6 absorbance units and were then induced at 298 K with 0.5 m*M* isopropyl D-thiogalactopyranoside; growth was continued for 16 h.

2.2. Purification

Cells were harvested by centrifugation at 4000*g* for 30 min at 277 K. The pellet was resuspended in ice-cold lysis buffer containing 50 m*M* Tris–HCl buffer pH 8.0, 20 m*M* NaCl, 1 m*M* PMSF, 0.2 m*M* benzimidine–HCl and 2 m*M* DTT and lysed by sonication on ice. The crude cell extract was subjected to nucleic acid precipitation with 1% streptomycin sulfate and 0.1% polyethyleneimine for 30 min at 277 K with constant stirring and was then centrifuged at 40 000*g* for 45 min at 277 K. Three chromatographic steps were performed in order to purify the protein from the supernatant. The first step involved passing the supernatant through a Q-Sepharose column (HiPrep 16/10 Q XL, Amersham Pharmacia) followed by overnight ammonium sulfate precipitation at 277 K. After centrifugation, the precipitated protein was resuspended in 50 m*M* Tris–HCl pH 8.0 and loaded onto a Phenyl Sepharose (Amersham Pharmacia) column equilibrated with 1.3 *M* ammonium sulfate and 50 m*M* Tris–HCl pH 8.0. Fractions containing the protein were pooled, concentrated and further purified to homogeneity by gel filtration with a Superdex-75 chromatography column. The eluted protein was concentrated by ultrafiltration by a 10 kDa cutoff (YM-10) Centricon device. The protein concentration was determined by a Bradford reagent assay.

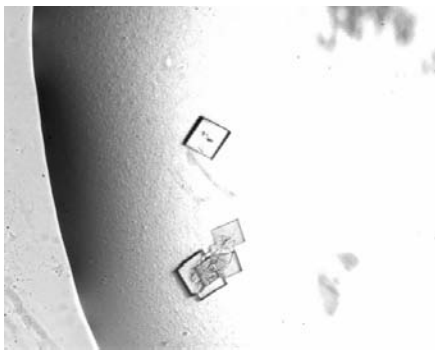

Figure 1
Crystals of the editing module of ThrRS from *P. abyssi*.

2.3. Crystallization and preliminary X-ray data

Crystallization trials were performed using the hanging-drop vapour-diffusion method, equilibrating varying drop volumes of protein solution against 750 µl reservoir solution. Initial crystallization attempts were performed at room temperature as well as at 277 K using Crystal Screens from Hampton Research and the crystallization conditions were refined with fine variations in pH, ionic strength and precipitant concentration.

Preliminary diffraction data were collected on an in-house MAR Research MAR-345dtb image-plate detector with Cu $K\alpha$ X-rays generated by a Rigaku RU-H3R rotating-anode generator equipped with an Osmic mirror system operated at 50 kV and 100 mA. Prior to flash-cooling in a liquid-nitrogen stream at 100 K for data collection, crystals were soaked in 15% glycerol along with the mother liquor for 5–20 s. X-ray data were processed using *DENZO* (Otwinowski & Minor, 1997) and subsequent scaling and merging of intensities was carried out using *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results

The recombinant plasmid was sequenced to confirm the correct sequence of the insert and the protein was expressed in soluble form in *E. coli*. A high level of expression facilitated purification using a three-step protocol. The cell lysate was loaded onto a Q-Sepharose column and the protein was eluted at ∼600 m*M* NaCl in a 20 m*M*–2*M* NaCl gradient. The fractions containing the protein were pooled and precipitated with ammonium sulfate at 90% saturation overnight and then centrifuged at 20 000*g* for 20 min at 277 K. After resuspension, the precipitated protein was loaded onto a Phenyl Sepharose column. Protein was eluted at around 125 m*M* ammonium sulfate in a linear gradient of 1.3–0 *M* ammonium sulfate. The fractions were concentrated to 1 ml by ultrafiltration then run through a Superdex-75 column equilibrated with buffer containing 50 m*M* Tris–HCl pH 8.0 and 50 m*M* NaCl. The purity of the protein eluted by the gel-filtration column was 90–95% as estimated by 15% SDS–PAGE, with a typical yield of about 3 mg of protein per litre of culture as estimated by a Bradford reagent assay. Fractions containing the purified protein in 50 m*M* Tris–HCl pH 8.0 and 50 m*M* NaCl were pooled, concentrated and exchanged with 50 m*M* Tris–HCl pH 7.5 by ultrafiltration using a 10 kDa cutoff Centricon. The N-terminal sequence of the concentrated and purified protein was verified using a Procise cLC sequencer (Applied Biosystems). The protein molecular weight was estimated by MALDI–MS and was found to be 21.060 kDa compared with the calculated value of 21.097 kDa.

Table 1
Essential crystallographic data.

Values in parentheses are for the highest resolution shell

Space group	$P3_{1(2)}21$
Unit-cell parameters	
$a = b$ (Å)	61.75
c (Å)	64.89
Resolution (Å)	25.0–1.95 (2.02–1.95)
Observations	34227 (2477)
Unique reflections	10459 (928)
Completeness (%)	96.7 (87.2)
Redundancy	3.3 (2.7)
R_{merge}† (%)	4.1 (31.4)
$I/\sigma(I)$	25.28 (2.25)
Solvent content (%)	27.7
V_M (Å^3 Da^{-1})	1.70
Monomers per AU	1

† $R_{merge} = \sum|I(h) - \langle I(h)\rangle|/\sum I(h)$, where $I(h)$ is the observed intensity and $\langle I(h)\rangle$ is the mean intensity of reflection h over all measurements of $I(h)$.

Initially, microcrystals appeared in 20%(*w*/*v*) PEG 10 000 and 0.1 *M* HEPES pH 7.5 at 277 K. Conditions were optimized from the first screen results and well diffracting crystals (Fig. 1) were obtained by mixing 2 µl protein solution at 3.25 mg ml^{-1} in 50 m*M* Tris–HCl pH 7.5 with an equal volume of reservoir solution [25%(*w*/*v*) PEG 8000 and 0.1 *M* HEPES pH 7.0] at 277 K. The crystals grew to maximum dimensions of 30 × 10 × 10 µm in 2–3 weeks. Despite the small size of the crystals, they diffracted X-rays to a maximum resolution of 1.95 Å, presumably because of the low solvent content of the crystals (27.7%). Indexing of the diffraction pattern indicated that the crystals belonged to the trigonal space group $P3_{1(2)}21$, with unit-cell parameters $a = b = 61.75$, $c = 64.89$ Å, $\alpha = \beta = 90.0$, $\gamma = 120.0^\circ$. The X-ray data were collected by rotating the crystal through a total of 56.5° with 0.5° oscillation and a crystal-to-detector distance of 155 mm. The mosaicity of the crystal was found to be 0.68. The crystal data-collection statistics are shown in Table 1. The overall completeness and R_{merge} are 96.7% and 4.1%, respectively. Efforts are under way to prepare heavy-atom derivatives and selenomethionine-substituted protein in order to solve the structure.

SD thanks the Council of Scientific and Industrial Research (CSIR), India for a senior research fellowship. We thank V. M. Shanmugam for help during X-ray data

collection. RS is an International Senior Research Fellow (ISRF) of The Wellcome Trust, UK in India.

References

Beebe, K., Merriman, E., Pouplana, L. R & Schimmel, P. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 5958–5963.
Carter, C. W. Jr (1993). *Annu. Rev. Biochem.* **62**, 715–748.
Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Calliet, J., Springer, M., Rees, B., Francklyn, C. S. & Moras, D. (2000). *Cell*, **103**, 877–884.
Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). *Nature (London)*, **347**, 203–206.
Lin, L., Hale, S. P. & Schimmel, P. (1996). *Nature (London)*, **384**, 33–34.
Meinnel, T., Mechulam, Y. & Blanquet, S. (1995). *tRNA: Structure, Biosynthesis and Function*, edited by D. Söll & U. L. RajBhandary, pp. 251–291. Washington DC: ASM Press.
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.
Sankaranarayanan, R., Dock-Bregeon, A. C., Rees, B., Bovee, M., Caillet, J., Romby, P., Francklyn, C. S. & Moras, D. (2000). *Nature Struct. Biol.* **7**, 461–465.
Sankaranarayanan, R., Dock-Bregeon, A. C., Romby, P., Caillet, J., Springer, M., Rees, B., Ehresmann, C., Ehresmann, B. & Moras, D. (1999). *Cell*, **97**, 371–381.
Schimmel, P. & Söll, D. (1979). *Annu. Rev. Biochem.* **48**, 601–648.
Silvian, L. F., Wang, J. & Steitz, T. A. (1999). *Science*, **285**, 1074–1077.