Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary X-ray diffraction data of L7Ae sRNP core protein from *Pyrococcus abyssii*

Christophe Charron,[a] Xavier Manival,[b] Bruno Charpentier,[b] Christiane Branlant[b] and André Aubry[a]*

[a]Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques, UMR 7036 CNRS-UHP, Groupe Biocristallographie, Faculté des Sciences et Techniques, BP 239, 54506 Vandoeuvre-lès-Nancy, France, and [b]Laboratoire de Maturation des ARN et Enzymologie Moléculaire, UMR 7567 CNRS-UHP, Faculté des Sciences et Techniques, BP 239, 54506 Vandoeuvre-lès-Nancy, France

Correspondence e-mail: andre.aubry@lcm3b.uhp-nancy.fr

Received 13 August 2003
Accepted 7 October 2003

The L7Ae sRNP core protein from *Pyrococcus abyssii* was crystallized using the sitting-drop vapour-diffusion method. Crystals were obtained in the presence of $MgCl_2$, PEG 2000 MME and acetate buffer at pH 4.0. A native data set has been collected at 2.9 Å resolution using a rotating-anode generator at room temperature. Crystals belong to the orthorhombic space group $P2_12_12$, with unit-cell parameters a = 70.7, b = 112.9, c = 34.8 Å. There are two monomers of MW 14 200 Da per asymmetric unit and the packing density V_M is 2.45 Å^3 Da^{-1}. A molecular-replacement analysis gave solutions for the rotation and translation functions.

1. Introduction

It was recently shown that as in eukarya, 2′-O-methylations and pseudouridylations of archaeal ribosomal RNAs are RNA-guided processes. The C/D box sRNAs guide 2′-O-methylations, while the H/ACA sRNAs guide uridine-to-pseudouridine conversions (for a review, see Bachellerie *et al.*, 2002; Omer *et al.*, 2003). The eukaryal snoRNAs associate with four core proteins, *i.e.* 15.5 kDa protein (human) or Snu13p (in yeast), Nop56p, Nop58p and the Nop1p/fibrillarin 2′-*O*-methyltransferase (Lafontaine & Tollervey, 2000; Newman *et al.*, 2000; Watkins *et al.*, 2000). The 15.5 kDa/Snu13p protein plays an essential role in the assembly process (Watkins *et al.*, 2000, 2002). Surprisingly, the archaeal homologue of the 15.5 kDa/Snu13p protein is the ribosomal L7Ae protein (Watkins *et al.*, 2000; Kuhn *et al.*, 2002), which binds both 23S rRNA (Ban *et al.*, 2000) and sRNAs (Omer *et al.*, 2002). Indeed, active C/D box sRNP particles have recently been reconstituted by incubation of an *in vitro* transcribed C/D sRNA with the protein L7Ae, the archaeal counterpart of the Nop56p and Nop58p proteins (protein aNOP56/58 or aNOP5) and the archaeal homologue of Nop1p (aNOP1) (Omer *et al.*, 2002). Thus, protein L7Ae is a multifunctional protein in archaea. It has even been shown recently that protein L7Ae is a component of the archaeal H/ACA sRNP (Rozhdestvensky *et al.*, 2003).

The 15.5 kDa/Snu13p protein also has two distinct functions in vertebrates. It binds both C/D box snoRNAs and the spliceosomal U4 snRNA (Nottrott *et al.*, 1999). The three-dimensional structure of the 15.5 kDa–U4 snRNA complex has been solved at 2.9 Å resolution by X-ray analysis (Vidovic *et al.*, 2000). In this complex, the RNA adopts a K-turn structure, with two G·A sheared pairs and a protruding U residue bound into a protein pocket. Several K-turn structures have been found in ribosomal RNAs (Klein *et al.*, 2001). According to the three-dimensional structure of *Haloarcula marismortui* 50S subunit (Ban *et al.*, 2000), the 23S rRNA KT-15 K-turn is recognized by both the L15e and L7Ae ribosomal proteins. The C/D motifs of snoRNAs and sRNAs are also expected to form a K-turn structure (Marmier-Gourrier *et al.*, 2003; Omer *et al.*, 2002; Watkins *et al.*, 2000), which are respectively recognized by the 15.5 kDa/Snu13p and L7Ae proteins. However, no three-dimensional structure of a C/D motif either free or in association with protein has yet been established. Only the crystal structure of an aNOP56/58–NOP1 heterodimer has recently been solved (Aittaleb *et al.*, 2003).

In an attempt to study the C/D box sRNA–L7Ae interaction, we initiated a crystallographic study of the L7Ae protein. Here, we describe the purification and the crystallization of L7Ae and the preliminary analysis of the X-ray diffraction data showing that the crystals are suitable for structure determination.

2. Materials and methods

2.1. Cloning procedures, expression and purification

For preparation of recombinant protein, the L7Ae ORF was PCR-amplified from the genomic DNA of the *Pyrococcus abyssii* GE5 strain (a generous gift from Dr J. Myllykallio, University of Paris-Orsay). We used two oligonucleotides, the gene-specific primers L7Ae-5′ (5′-**GGATCC**ATGGAGGGATGGATGATGG-3′) and L7Ae-3′ (5′-**CTCGAG**TCACTTCATGAGCTCCCTA-3′), which gener-

ated a DNA fragment bordered by a *Bam*H1 and a *Xho*I restriction site (in bold). The PCR product was cloned into the GST-fusion expression vector pGEXT-4T-1 (Amersham Biosciences) and the insert of the selected clone was sequenced on both strands.

Escherichia coli strain BL21-CodonPlus-RIL (Stratagene) cells transformed with the pGEX-4T-1-L7Ae plasmid were grown at 310 K to an OD_{600} of 0.8 in 4 l of LB medium (Miller, 1972) containing 100 mg l^{-1} ampicillin. Expression of the GST-L7Ae ORF was induced by the addition of 0.2 m*M* isopropyl-β-D-thiogalactopyranoside (MBI-Fermentas) and growth was continued for 6 h at 303 K. SDS–PAGE analysis of the total cell lysate demonstrated the solubility of the recombinant protein.

For purification, cells were broken by two passages through a disruption system at 138 MPa (Constant Cell) in buffer *A* (50 m*M* Tris–HCl pH 8.0, 1 m*M* EDTA, 1 m*M* DTT, 150 m*M* NaCl) containing 0.5 mg ml^{-1} benzamidine protease inhibitor (Sigma). After 1 h centrifugation at 277 K and 141 000*g*, nucleic acids from the supernatant were precipitated by addition of 0.025% polyethylenimine (Sigma). A second centrifugation was performed for 20 min at 10 000*g* and the supernatant was directly incubated with glutathione-Sepharose beads (Amersham Biosciences). Incubation was performed overnight under shaking at room temperature. The beads were then washed with 100 ml of saline buffer *B* (50 m*M* Tris–HCl pH 8, 1 m*M* EDTA, 1 m*M* DTT, 2 *M* NaCl). Elimination of the GST moiety of the fusion protein by proteolytic cleavage was directly performed on the beads in buffer *A* for 12 h at room temperature, using 10 U of bovine thrombin (Calbiochem) per milligram of protein. The eluate was then incubated for 20 min at 338 K. Under these conditions, only the archaeal L7Ae protein remained soluble and traces of nucleic acids were eliminated by a second precipitation with PEI (at 0.05%). After centrifugation at 10 000*g* for 20 min at room temperature, the protein was concentrated using a Centriprep (Millipore, 3 kDa cutoff) to a concentration of about 10 mg ml^{-1}. The concentration of soluble protein was determined from the absorbance at 280 nm using an extinction coefficient of 0.69 cm^{-1} mg ml^{-1}.

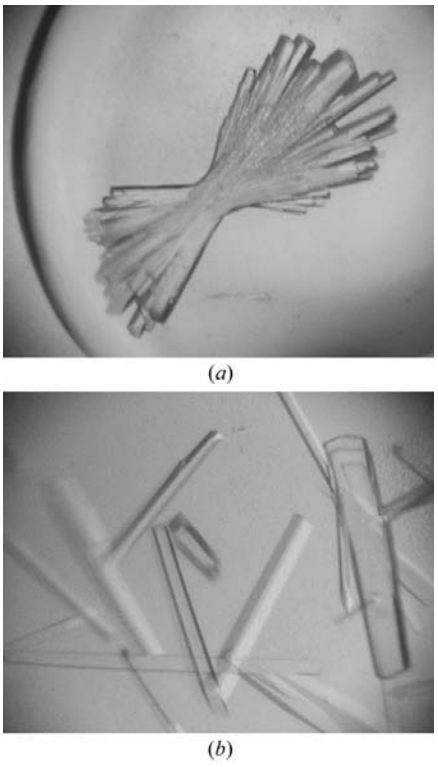

Figure 1
Crystals of protein L7Ae from *P. abyssii*. (*a*) A bunch of crystals obtained using the sitting-drop vapour-diffusion method in the presence of $MgCl_2$, PEG 2000 MME and acetate buffer pH 4.0. (*b*) Single crystals suitable for X-ray diffraction analysis obtained by microseeding.

2.2. Crystallization

Initial crystallization conditions were searched by sparse-matrix screening (Jancarik & Kim, 1991) using the vapour-diffusion technique in sitting drops at 293 K. The 3 µl drops were equilibrated over 200 µl reservoirs. A total of 582 conditions were covered with nine screens (Hampton Research Screens I and II, Hampton Research PEG/Ion Screen, Hampton Research Clear Strategy Screens I and II, Molecular Dimensions Limited 3D Structure Screen I and II and deCODE Genetics Wizard Screens I and II). In each condition, 2 µl of 10 mg l^{-1} L7Ae solution was mixed with 1 µl of reservoir solution. Single crystals suitable for X-ray diffraction analysis were obtained by microseeding (Stura, 1999).

2.3. Diffraction measurements

X-ray data were collected in our laboratory at room temperature on a DIP2030 Enraf–Nonius area detector using a rotating-anode generator. The wavelength of the incident radiation was 1.542 Å and the crystal-to-detector distance was 210 mm. A total range of 100° was covered, with a 1.0° oscillation per image. Data were processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1996) and the indexed intensities were converted to structure factors using *TRUNCATE* from the *CCP*4 program suite (Collaborative Computational Project, Number 1994) without any σ cutoff.

Table 1
Statistics of X-ray data measurement for *P. abyssii* L7Ae crystals.

Values in parentheses correspond to the last resolution shell.

Space group	$P2_12_12$
Unit-cell parameters (Å)	a = 70.7, b = 112.9, c = 34.8
Resolution range (Å)	20–2.9 (3.0–2.9)
No. unique reflections	5958
R_{sym} (%)	8.4 (23.1)
Completeness (%)	91.0 (77.5)
$\langle I/\sigma(I)\rangle$	10.5 (3.2)
Multiplicity	3.2

3. Results and discussion

Of the 582 crystallization conditions tested, only one gave crystals, namely condition No. 3 from Hampton Research Clear Strategy Screen I kit containing 25%(*v*/*v*) PEG 2000 MME, 200 m*M* $MgCl_2$ and buffer No. 1 (100 m*M* sodium acetate buffer pH 4.5). Bunches of crystals grew in about one week (Fig. 1*a*). Single crystals suitable for X-ray diffraction analysis (Fig. 1*b*) were obtained by microseeding using microcrystals and drops having compositions close to that of the equilibrium drops. The drops were produced by mixing 1 µl of a 10 mg ml^{-1} protein solution in 300 m*M* NaCl and 20 m*M* Tris–HCl pH 8.0 and 4 µl of a solution containing 20%(*v*/*v*) PEG 2000 MME, 160 m*M* $MgCl_2$ and 100 m*M* sodium acetate buffer pH 4.0.

A total of 18 984 reflections in the resolution range 20–2.9 Å were collected at room temperature in our laboratory. They were reduced to 5958 unique reflections. Overall, the data set had an R_{sym} of 8.4% on intensities ($R_{sym} = \sum|I - \langle I\rangle|/\sum I$) and was 91.0% complete (Table 1).

L7Ae crystals belong to a primitive orthorhombic space group, with unit-cell parameters $a = 70.7$, $b = 112.9$, $c = 34.8$ Å and $Z = 4$. Systematic extinctions [$I/\sigma(I) < 3$] of $h00$ (with $h = 2n + 1$) and $0k0$ (with $k = 2n + 1$) reflections indicate the space group to be $P2_12_12$. L7Ae is a monomer with MW 14 200 Da; assuming the presence of two monomers in the asymmetric unit, the packing density V_M is 2.45 $Å^3$ Da^{-1} and the solvent content is 49.8%. These values are in good agreement with those for other proteins (Matthews, 1968) and strongly support the presence of two monomers per asymmetric unit.

Initial attempts to solve the crystal structure of L7Ae were performed by molecular

replacement with the program *AMoRe* (Navaza, 1994) in the resolution range 10–3.5 Å. The monomeric L7Ae protein from *H. marismortui* (Klein *et al.*, 2001), which shows 55% sequence identity to the *P. abyssi* L7Ae protein, served as a structural model. The rotation function was calculated with a 22 Å integration radius. Subsequently, one- and two-body translation searches allowed the positioning of two L7Ae molecules in the asymmetric unit. Indeed, the calculation of the translation function gave one major peak with a correlation coefficient of 48.3%, while the next solution exhibited a correlation of 38.8%. Rigid-body refinement of the two molecules yielded a correlation coefficient of 57.7% in the resolution range 10–3.5 Å. Model building, refinement and structure analysis of L7Ae will be published elsewhere. Coordinates have been deposited in the PDB with code 1pxw.

This work was supported by the Centre National de la Recherche Scientifique (CNRS), the French Ministère de la Recherche et des Nouvelles Technologies, the Région Lorraine (Pôle Régional Bioingénérie) and the Action Concertée Incitative Microbiologie of the French Ministère de la Recherche et des Nouvelles Technologies (2001).

References

Aittaleb, M., Rashid, R., Chen, Q., Palmer, J. R., Daniels, C. J. & Li, H. (2003). *Nature Struct. Biol.* **10**, 256–263.
Bachellerie, J. P., Cavaille, J. & Huttenhofer, A. (2002). *Biochimie*, **84**, 775–790.
Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000). *Science*, **289**, 905–920.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Klein, D. J., Schmeing, T. M., Moore, P. B. & Steitz, T. A. (2001). *EMBO J.* **20**, 4214–4221.
Kuhn, J. F., Tran, E. J. & Maxwell, E. S. (2002). *Nucleic Acids Res.* **30**, 931–941.
Lafontaine, D. L. & Tollervey, D. (2000). *Mol. Cell Biol.* **20**, 2650–2659.
Marmier-Gourrier, N., Clery, A., Senty-Segault, V., Charpentier, B., Schlotter, F., Leclerc, F., Fournier, R. & Branlant, C. (2003). *RNA*, **9**, 821–838.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Miller, J. H. (1972). *Experiments in Molecular Genetics*. New York: Cold Spring Harbor Laboratory Press.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Newman, D. R., Kuhn, J. F., Shanab, G. M. & Maxwell, E. S. (2000). *RNA*, **6**, 861–879.
Nottrott, S., Hartmuth, K., Fabrizio, P., Urlaub, H., Vidovic, I., Ficner, R. & Luhrmann, R. (1999). *EMBO J.* **18**, 6119–6133.
Omer, A. D., Ziesche, S., Decatur, W. A., Fournier, M. J. & Dennis, P. P. (2003). *Mol. Microbiol.* **48**, 617–629.
Omer, A. D., Ziesche, S., Ebhardt, H. & Dennis, P. P. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 5289–5294.
Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.
Stura, E. A. (1999). *Crystallization of Nucleic Acids and Proteins. A Practical Approach*, 2nd ed., edited by A. Ducruix & R. Giegé, pp. 177–208. Oxford University Press.
Rozhdestvensky, T. S., Tang, T. H., Tchirkova, I. V., Brosius, J., Bachellerie, J. P. & Huttenhofer, A. (2003). *Nucleic Acids Res.* **31**, 869–877.
Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R. & Ficner, R. (2000). *Mol. Cell*, **6**, 1331–1342.
Watkins, N. J., Dickmanns, A. & Luhrmann, R. (2002). *Mol. Cell Biol.* **22**, 8342–8352.
Watkins, N. J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C. & Luhrmann, R. (2000). *Cell*, **103**, 457–466.