Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Expression, purification, crystallization and preliminary diffraction analysis of RNase P protein from *Thermotoga maritima*

Angelika A. Krivenko,[a] Alexei V. Kazantsev,[a] Catherine Adamidi,[b] Daniel J. Harrington[c] and Norman R. Pace[a]*

[a]Department MCD Biology, University of Colorado, Boulder, CO 80309, USA, [b]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA, and [c]Stanford Synchrotron Radiation Laboratory, Stanford University, Menlo Park, CA 94025, USA

Correspondence e-mail: nrpace@colorado.edu

Received 15 March 2002
Accepted 29 April 2002

Ribonuclease P (RNase P), the ubiquitous endonuclease that catalyzes maturation of the 5′-end of tRNA in bacteria, is a ribonucleoprotein particle composed of one large RNA and one small protein. Two major structural types of bacterial RNase P RNA have been identified by phylogenetic comparative analysis: the A (ancestral) and B (*Bacillus*) types. The RNase P protein from *Thermotoga maritima*, a hyperthermophilic bacterium with an A-type RNase P RNA, has been expressed in *Escherichia coli*. A purification strategy was developed to obtain a protein preparation suitable for crystallization. Protein crystals suitable for diffraction studies were obtained and characterized.

1. Introduction

RNase P is the ribonucleoprotein responsible for maturation of the 5′-end of tRNA. The bacterial RNase P consists of a large (350–400 nucleotides) RNA and a small (~120 amino acids) protein subunit (Frank & Pace, 1998). Bacterial RNase P RNA alone is capable of processing pre-tRNA *in vitro* under high ionic strength conditions. However, the protein component is absolutely required for enzymatic activity *in vivo* (Frank & Pace, 1998).

Two major structural types of bacterial RNase P RNA have been identified by phylogenetic comparative analysis (Haas *et al.*, 1996). The B (or *Bacillus*) type of RNase P RNA is found exclusively among a relatively narrow phylogenetic group of low-G+C Gram-positive bacteria. RNase P RNA from all other bacterial groups studied so far belong to the A (or ancestral) type (Haas *et al.*, 1996). RNase P (holoenzyme and RNA alone) has so far eluded high-resolution structural studies. Structural differences between A-type and B-type RNase P RNAs are substantial and the sequence similarities among the RNase P proteins are low (typically ~35–40%). Nonetheless, A-type and B-type proteins can both activate the other type of RNA (Guerrier-Takada *et al.*, 1983). Commonalities in the structures of the two types of proteins are therefore expected to highlight the elements responsible for the activation.

Structural studies of RNase P proteins so far have focused on the B-type proteins, *e.g. Bacillus subtilis* (Stams *et al.*, 1998) and *Staphylococcus aureus* (Spitzfaden *et al.*, 2000).

We have overexpressed in *E. coli* the RNase P protein from *T. maritima*, a hyperthermophilic bacterium with an A-type RNase P RNA (Brown *et al.*, 1993). Despite recent reports of poor solubility of this protein (Paul *et al.*, 2001), we have developed an expression, purification and crystallization strategy allowing crystals of RNase P protein from *T. maritima* suitable for crystallographic studies to be obtained.

2. Protein expression and purification

Recombinant constructs for protein expression were produced using the *rnpA* gene (Gene Bank No. 3.1.26.5) amplified by PCR from genomic DNA from *T. maritima* (a kind gift from Dr James Brown, NCSU) with appropriate oligonucleotide primers. A number of commercially available expression systems were tested and are listed in Table 1. In the case of each expression system, a number of different *E. coli* strains were tested [BL21, BL21(DE3)pLysS, M15] as well as a number of different cell culture media (LB and TB; Sambrook *et al.*, 1989) and growth temperatures (296, 303 and 310 K). We found that in the systems tested the protein could be expressed in significant amounts only with glutathione S-transferase (GST) as an N-terminal fusion. The fusion polypeptide is cleaved by digestion with thrombin to furnish a *T. maritima* RNase P protein version that is missing the N-terminal methionine and has a Gly-Ser sequence instead.

The optimized expression protocol included growth of the cell culture [BL21(DE3)pLysS strain bearing the *rnpA* derivative of pGEX4Ta vector] in LB medium containing appropriate antibiotics at 310 K with vigorous shaking. When OD_{600} reached 0.4, expression was induced by addition of IPTG to 1 m*M*

Table 1
Expression systems used to overexpress RNase P protein from *T. maritima*.

Vector	Promoter	Fusion	Protease site	Vendor	Expression
pET11	T7 *lac*		No	Novagen	–
pET16	T7 *lac*	N-terminal His-tag	Factor Xa	Novagen	–
pTYB1	T7	C-terminal intein	No	NEB	–
pTYB11	T7	N-terminal intein	No†	NEB	–
pBAD-TOPO-TA	AraBAD		No†	Invitrogen	–
pQE-30	T5	N-terminal His-tag	No	Qiagen	+/–
pGEX4Ta	T7 *tac*	N-terminal GST	Thrombin	Amersham	+++

† Inducible self-cleavage activity of intein.

concentration and proceeded for 6–7 h at 310 K. The cells were harvested by centrifugation at 5000*g*, resuspended in an appropriate volume of lysis buffer [50 m*M* Tris–HCl pH 7.5, 4 m*M* EDTA, 10%(*v*/*v*) glycerine, 0.1%(*v*/*v*) NP-40 and Complete protease inhibitors (Roche) at the concentration recommended by the manufacturer]. The cell suspension was lysed by freeze–thaw followed by a single passage through a French Press at 110 MPa.

Initial attempts to purify the overexpressed GST-RNase P fusion polypeptide included affinity chromatography on Glutathione Sepharose 4B (Amersham) and ion-exchange chromatography under denaturing (5 *M* urea) or non-denaturing conditions. A major problem encountered was insolubility of the fusion protein. The fusion protein was soluble under denaturing conditions, but precipitated upon dialysis into non-denaturing conditions. We found that the fusion protein is initially expressed in a soluble form which precipitates upon dialysis from denaturing to native conditions or when the protein binds to a chromatography resin (affinity or ion-exchange). The protein also precipitates if left for several hours in the cellular lysate. However, thrombin digestion of the fusion protein performed directly in the cellular lysate immediately after cell lysis (600 NIH units of thrombin per 40 ml of the lysate obtained from 1 l expression culture) yielded *T. maritima* RNase P protein that was highly soluble under both denaturing and native conditions. Protease inhibitors added to the lysis buffer did not notably inhibit the proteolytic activity of the thrombin.

Purification of the digested RNase P protein from the cellular lysate was achieved by low-resolution ion-exchange chromatography under denaturing conditions (Tris-acryl M SP resin, LKB) followed by high-resolution ion-exchange chromatography (Mono S resin, Amersham) under non-denaturing conditions. The denaturing step was introduced to minimize the loss of RNase P protein owing to aggregation with nucleic acids in the lysate (Niranjanakumari *et al.*, 1998). In both cases, the column was eluted with a linear gradient of NaCl from 0.1 to 2.0 *M* in 50 m*M* Tris–HCl pH 7.5, 0.2 m*M* EDTA (plus 5 *M* urea for the denaturing run). Dry urea was added to the thrombin-digested lysate to 5 *M* final concentration prior to applying the lysate to the column. RNase P protein eluted between 0.75 and 1.20 *M* NaCl upon denaturing chromatography and between 1.25 and 1.50 *M* NaCl in the non-denaturing run. The purity of the protein preparation was monitored by SDS–PAGE at every step. Purified protein was extensively dialyzed against 10 m*M* Tris–HCl pH 7.5, 0.2 m*M* EDTA, concentrated to approximately 14 mg ml^{-1} using a Centricon 10 centrifugation device (10 kDa moelcular-weight cutoff; Millipore) and stored at 277 K. No precipitation or protein degradation was detected for at least one year.

The *T. maritima* RNase P protein, expressed and purified in the manner described above, can bind to the *T. maritima* RNase P RNA to form an active holoenzyme, as assayed by processing activity of *T. maritima* pre-tRNA^{Glu} at low salt concentration (data not shown).

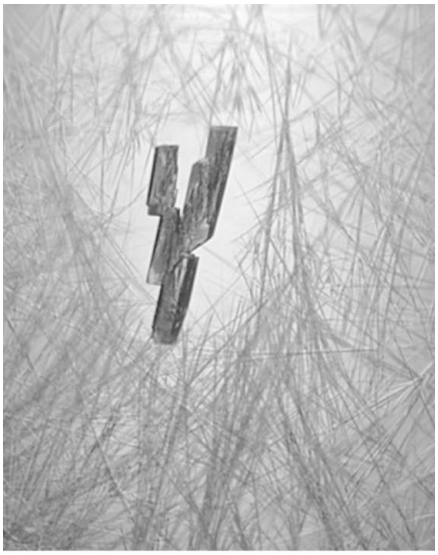
3. Crystallization

Crystallization of *T. maritima* RNase P protein was achieved using a vapor-diffusion technique. Initial conditions for crystallization were found with sparse-matrix screens (Cudney *et al.*, 1994; Jancarik & Kim, 1991). Screening of 96 precipitant formulations at a protein concentration of 14 mg ml^{-1} immediately furnished protein crystals of low quality that grew at ambient temperature (296 K) within one week. Subsequent optimizations of crystallization conditions included lowering protein concentration, variation of pH, precipitant identity and systematic variation of concentrations of all the components in the crystallization cocktail. The following trends were observed and used in obtaining RNase P protein crystals of good diffraction quality.

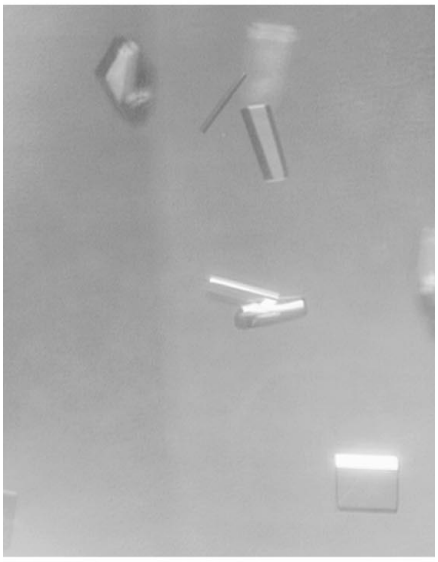
(i) Protein concentration could be reduced to 2–3 mg ml^{-1}.

(ii) Protein crystals grew only in the presence of sulfate (Na^+, K^+ or NH_4^+).

(iii) Two crystal forms, plates and thin long needles, could coexist in the same crystallization condition (Fig. 1*a*). High concentrations of the precipitant (PEG 1500) and low pH tend to favor formation of needles, while low concentrations of precipitant and higher pH favor the formation of plates. Utilization of low-molecular-weight precipitants (PEG 200–400, 2-propanol) suppresses formation of plates; higher molecular-weight precipitants (PEG 4000–



(*a*)



(*b*)

Figure 1
Crystals of RNase P protein from *T. maritima*. (*a*) Two crystal forms found in the crystallization drops. (*b*) Plate-like crystals obtained by microseeding with a cat whisker that were used for the data collection. The crystal dimensions are approximately 150 × 150 × 45 µm.

8000) suppress formation of needles, but the quality of the plates decreases significantly.

(iv) Crystal growth is limited by slow nucleation. In the conditions optimal for the plate-like crystal growth (100 m*M* NaOAc pH 4.8–5.2, 12–18% PEG 1500, 200 m*M* K_2SO_4 and 3 mg ml^{-1} protein), crystals often failed to grow. Addition of crystal seeds obtained from the low-quality plate-like crystals to a freshly set crystallization drop furnished a reasonable amount of crystals that appeared overnight with high reproducibility and grew to a reasonable size within 3 d (Fig. 1*b*). The nucleating ability of the needle-like crystal form was not investigated.

4. Data collection and processing

Prior to data collection, the protein crystals were soaked in a cryoprotectant solution (35% PEG 1500, 100 m*M* NaOAc pH 5.2, 20 m*M* K_2SO_4) for at least 5 min and flash-cooled in nitrogen gas at 100 K. Data from the crystals were collected at cryogenic temperature (100 K) using an R-AXIS IIC image-plate system mounted on a Rigaku RU-H2R rotating-copper-anode generator equipped with focusing mirrors and an X-Stream cryocooler (Molecular Structure Corp., The Woodlands, Texas, USA). 180° of data were collected in frames of 1°, with 10 min exposure time per frame. The raw data were indexed, processed and scaled using the *HKL* software package (Otwinowski & Minor, 1997). Data-collection statistics are outlined in Table 2. The protein crystals diffract to 2.0 Å and belong to the monoclinic space group $P2_1$, with unit-cell parameters $a = 56.29$, $b = 64.41$, $c = 68.19$ Å, $\beta = 101.94°$.

Table 2
Data-collection statistics.

Values for the highest resolution shell (2.07–2.00 Å) are given in parentheses (ten shells).

Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 56.29$, $b = 64.41$, $c = 68.19$, $\beta = 101.94$
Resolution (Å)	2.00 (2.07–2.00)
Observed reflections	109492 (8199)
Unique reflections	29573 (2404)
Data completeness (%)	91.4 (74.5)
$R_{\rm merge}$† (%)	7.5 (33.2)
$I/\sigma(I)$	17.1 (3.3)
Reflections with $I/\sigma(I) > 3$ (%)	80.0 (57.5)

† $R_{\rm merge} = \sum |I - \langle I \rangle| / \sum I$.

This work was funded by the National Institutes of Health.

References

Brown, J. W., Haas, E. S. & Pace, N. R. (1993). *Nucleic Acids Res.* **21**, 671–679.
Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.
Frank, D. N. & Pace, N. R. (1998). *Annu. Rev. Biochem.* **67**, 153–180.
Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983). *Cell*, **35**, 849–857.
Haas, E. S., Banta, A. B., Harris, J. K., Pace, N. R. & Brown, J. W. (1996). *Nucleic Acids Res.* **24**, 4775–4782.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Niranjanakumari, S., Kurz, J. C. & Fierke, C. A. (1998). *Nucleic Acids Res.* **26**, 3090–3096.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276,** 307–326.
Paul, R., Lazarev, D. & Altman, S. (2001). *Nucleic Acids Res.* **29**, 880–885.
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. New York: Cold Spring Harbor Laboratory.
Spitzfaden, C., Nicholson, N., Jones, J. J., Guth, S., Lehr, R., Prescott, C. D., Hegg, L. A. & Eggleston, D. S. (2000). *J. Mol. Biol.* **295**, 105–115.
Stams, T., Niranjanakumari, S., Fierke, C. A. & Christianson, D. W. (1998). *Science*, **280**, 752–755.