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

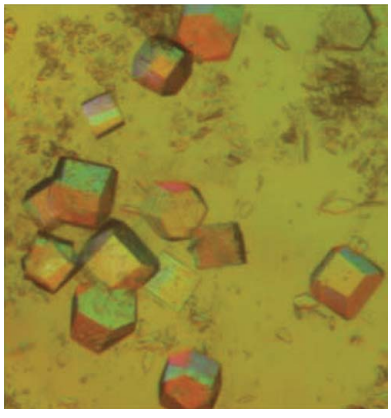
Large-scale purification and crystallization of the endoribonuclease XendoU: troubleshooting with His-tagged proteins

Fabiana Renzi,[a] Gianna Panetta,[a] Beatrice Vallone,[a] Maurizio Brunori,[a] Massimo Arceci,[c] Irene Bozzoni,[b,c,d] Pietro Laneve[c,d] and Elisa Caffarelli[c]*

[a]Dipartimento di Scienze Biochimiche, University of Rome 'La Sapienza', Piazzale Aldo Moro 5, 00185 Roma, Italy, [b]Istituto Pasteur–Fondazione Cenci Bolognetti, University of Rome 'La Sapienza', Piazzale Aldo Moro 5, 00185 Roma, Italy, [c]Istituto di Biologia e Patologia Molecolari CNR, University of Rome 'La Sapienza', Piazzale Aldo Moro 5, 00185 Roma, Italy, and [d]Dipartimento di Genetica e Biologia Molecolare, University of Rome 'La Sapienza', Piazalle Aldo Moro 5, 00185 Roma, Italy

Correspondence e-mail: elisa.caffarelli@uniroma1.it

Received 22 December 2005
Accepted 20 February 2006



XendoU is the first endoribonuclease described in higher eukaryotes as being involved in the endonucleolytic processing of intron-encoded small nucleolar RNAs. It is conserved among eukaryotes and its viral homologue is essential in SARS replication and transcription. The large-scale purification and crystallization of recombinant XendoU are reported. The tendency of the recombinant enzyme to aggregate could be reversed upon the addition of chelating agents (EDTA, imidazole): aggregation is a potential drawback when purifying and crystallizing His-tagged proteins, which are widely used, especially in high-throughput structural studies. Purified monodisperse XendoU crystallized in two different space groups: trigonal $P3_121$, diffracting to low resolution, and monoclinic $C2$, diffracting to higher resolution.

1. Introduction

Intron-encoded small nucleolar RNAs (snoRNAs) are non-coding RNAs that play essential roles in eukaryotes and archaea. They are transcribed as longer precursors and are generally released through splicing (reviewed by Filipowicz & Pogacic, 2002). Nevertheless, some snoRNAs can also be processed by endonucleolytic cleavage in the intron sequences flanking the snoRNA (Fragapane *et al.*, 1993). XendoU is the first endoribonuclease involved in the processing of snoRNAs in higher eukaryotes to be described and is involved in the processing of at least two snoRNAs in *Xenopus laevis* oocytes (Fragapane *et al.*, 1993; Caffarelli *et al.*, 1994; Renzi *et al.*, 2002; Laneve *et al.*, 2003). It is a U-specific enzyme that, uniquely among known endoribonucleases, releases 2′-3′ cyclic phosphate termini using Mn^{2+} as an essential cofactor. XendoU is conserved among higher eukaryotes and has no homology to any known cellular ribonuclease, although it shows significant homology to proteins tentatively annotated as serine proteases. A viral counterpart of XendoU, called NendoU, has been reported. It is a distinctive marker of the nidoviruses, which include the SARS coronavirus (Snijder *et al.*, 2003): the SARS homologue shows biochemical characteristics that are similar to those of XendoU, although it prefers double-stranded substrates, and is essential in virus replication and transcription (Ivanov *et al.*, 2004). XendoU was isolated from a cDNA bank, cloned with an extra 12 N-terminal amino acids including a non-cleavable His_6 tag, expressed in bacteria, affinity purified and tested for activity *in vitro* (Laneve *et al.*, 2003). Mutational studies allowed the identification of the residues (Glu161, Glu167, His162, His178 and Lys224) essential for RNA cleavage, while enzyme–RNA interaction assays led to the definition of the sequence required for Mn^{2+}-independent binding (Gioia *et al.*, 2005).

2. Methods

2.1. Expression and purification

The XendoU coding sequence was cloned downstream to the His_6-coding region of the pQE30 vector (Qiagen; Laneve *et al.*, 2003). In order to obtain large amounts of protein for crystallization screening, XendoU was overexpressed in bacteria [*Escherichia coli* M15 (pREP4) strain] in LB medium at 303 K to an optical density

Table 1
Data-collection details, scaling statistics and main crystallographic parameters of trigonal and monoclinic XendoU crystals.

Values in parentheses refer to the highest resolution shell.

	Trigonal	Monoclinic form *A*	Monoclinic form *B*
X-ray source	XRD1, Elettra	XRD1, Elettra	ID14-EH1, ESRF
Wavelength (Å)	1.2	1.2	0.934
Resolution (Å)	30.0–3.3	30.0–3.2	30.0–2.2
Space group	$P3_121$	$C2$	$C2$
Unit-cell parameters (Å, °)	$a = b = 83.36$, $c = 313.00$	$a = 168.92$, $b = 53.48$, $c = 137.49$, $\beta = 119.24$	$a = 164.45$, $b = 53.20$, $c = 133.47$, $\beta = 121.86$
No. of observations	150023	116331	276703
No. of unique reflections	16906	17997	64530
Average $I/\sigma(I)$	12.0	10.4	8.8
Completeness (%)	92.0 (91.7)	99.2 (97.9)	96.0 (92.8)
R_{merge}†	0.066 (0.196)	0.089 (0.33)	0.17 (0.39)
Unit-cell volume (Å^3)	1887187	1083882	964575
Molecules in ASU	3	3	3
Solvent content (%)	57.43	50.59	46.01
Mosaicity (°)	0.40	0.66	1.04

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

($OD_{600\,nm}$) of 0.55 and was induced with 0.1 m*M* isopropyl β-D-thiogalactoside (IPTG) for 4 h with continuous agitation and good aeration. Subsequent manipulations were carried out at 277 K. Harvested cells were resuspended in lysis buffer (50 m*M* Na/phosphate pH 7.8, 300 m*M* NaCl), 10 m*M* imidazole and 1 mg ml^{-1} lysozyme while stirring for 2 h. After sonication, the cell lysate was centrifuged for 30 min at 18 000*g* and the clear supernatant was filtered (0.2 µm filter, Sartorius) and applied onto a 5 ml Ni–NTA column (Amersham) previously equilibrated with lysis buffer. The column was washed with 1 l washing buffer 1 (10 m*M* HEPES pH 7.8, 300 m*M* NaCl, 10 m*M* imidazole) and 30 ml buffer 2 (10 m*M* HEPES pH 7.8, 50 m*M* NaCl, 75 m*M* imidazole). The protein was eluted in 10 m*M* HEPES pH 7.8, 50 m*M* NaCl, 100 m*M* imidazole. Repeated steps of ultrafiltration and dilution across 10 kDa cutoff filters (Vivascience) eliminated the imidazole. The His-tagged protein was purified by HPLC gel filtration using a polymeric resin gel (PWXL-G3000 60 cm × 17 mm column; Tosoh Biosciences) and concentrated to 0.5 m*M* (17 mg ml^{-1}) in 10 m*M* HEPES pH 7.8, 50 m*M* NaCl, 20 m*M* EDTA. SDS–PAGE analysis showed a protein purity of greater than 95%.

XendoU was also cloned into pGEX (Amersham Biosciences) to create a GST-XendoU fusion protein. The resulting expression plasmid was transformed into *E. coli* BL21 (DE3). Cells expressing the cleavable GST-tagged protein were grown and lysed following the same protocol used for the His-tagged XendoU. The supernatant from the cell lysate was applied onto a glutathione-derivatized column (FFT-GST; Amersham Biosciences) equilibrated with lysis buffer (10 m*M* Tris–HCl pH 7.8, 50 m*M* NaCl). The column was washed with 30 volumes of washing buffer 1 (10 m*M* Tris–HCl pH 7.8, 300 m*M* NaCl) and ten volumes of washing buffer 2 (10 m*M* Tris–HCl pH 7.8, 50 m*M* NaCl) and subsequently incubated with bovine thrombin (Sigma; 0.5 U per milligram of protein) at 277 K overnight to allow tag proteolysis on the column. A FFT-PABA-derivatized resin column (Amersham) that binds proteases using the protease inhibitor *para*-aminobenzoic acid (PABA) was connected in line with the FFT-GST column. A single elution step allowed the thrombin to be retained on the PABA column and the untagged protein to be recovered in the flowthrough. The untagged XendoU was HPLC purified by gel filtration (PWXL-G3000 60 cm × 17 mm column; Tosoh Biosciences) and concentrated to 0.5 m*M* (17 mg ml^{-1}) in 10 m*M* HEPES pH 7.8, 50 m*M* NaCl. SDS–PAGE analysis showed a protein purity of greater than 95%.

The purity and monodispersity of the protein were analyzed by HPLC gel filtration both on methacrylate resin gel, which is resistant over a wide range of pH, and silica resin gel, which is inert toward analytes (PWXL and SWXL-G3000 columns, respectively, 30 cm × 0.7 mm; Tosoh Biosciences) (see Fig. 1).

2.2. Crystallization

The His-tagged XendoU used in crystallization screening was solubilized in buffers containing 20 m*M* EDTA. The hanging-drop vapour-diffusion method and Index Screen (Hampton Research) were used for initial crystallization trials, with a drop volume of 1 + 1 µl and a reservoir volume of 0.5 ml.

Initial crystals grew in hanging drops with 30%(*w*/*v*) PEG 4000, 0.2 *M* sodium citrate pH 5–6 at 293 K; these crystals were thin and fragile and did not diffract even at a synchrotron source. Subsequent screening resulted in the growth of crystals (300 × 300 × 300 µm) in 40% $(NH_4)_2SO_4$, 0.2 *M* sodium citrate pH 6 at 293 K which belonged to space group $P3_121$ but diffracted poorly; the resolution was improved by cocrystallization with 50 m*M* uridine 5′-monophosphate (5′-UMP), which may bind to the active site, possibly inducing the stabilization of flexible regions (Fig. 2*a*).

Crystals diffracting to higher resolution in space group *C*2 were obtained by the substitution of sodium citrate by 0.2 *M* phosphate pH 5.5; these crystals were larger but were not single (Fig. 2*c*, top). Single crystals were obtained by tuning the growth speed and nucleation by (i) depositing silicon oil above the precipitating solution to slow down phase equilibrium and (ii) using macroseeding and microseeding techniques. This led to the growth of large platelet-shaped single crystals (500 × 500 × 30 µm; Fig. 2*c*, bottom) which diffracted to higher resolution (2.2 Å).

2.3. Data collection

Crystals grown in 40% $(NH_4)_2SO_4$, 0.2 *M* sodium citrate pH 6.0 were cryoprotected by soaking in mother liquor with the addition of 25% glycerol and tested for diffraction at 100 K. Although they did not diffract when exposed on a Cu $K\alpha$ rotating-anode X-ray generator, they diffracted to 3.3 Å resolution at a synchrotron source and proved to belong to the trigonal lattice, space group $P3_121$ (Fig. 2*b*). To overcome the spot overlap arising from the long *c* axis,

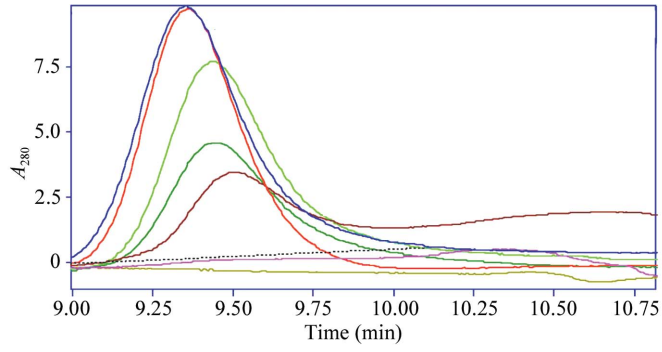


Figure 1
Effect of chelating agents in the solubilization of aggregated His-tagged protein. HPLC gel-filtration analysis shows that 250 m*M* (red) and 100 m*M* (blue) imidazole yield monodisperse His-tagged protein. 20 m*M* (light green), 10 m*M* (dark green) and 5 m*M* (purple) EDTA also improve monodispersity in a concentration-dependent fashion. Addition of 20 m*M* Mn^{2+} salt with and without EDTA (brown and pink) causes the protein to aggregate again. Elution profiles were analyzed using the *CSW* program.

the radiation wavelength was increased to 1.2 Å, the crystal-to-detector distance was increased to 210.00 mm and diffraction images were collected with a thin oscillation range ($\Delta\varphi$ = 0.4°). Crystals grown in 40% $(NH_4)_2SO_4$, 0.2 *M* phosphate pH 5.5 diffracted to 2.2 Å resolution, although not in all crystal orientations (Fig. 2*d*); this is probably correlated to the thin third dimension and to imperfect or loose packing as suggested by the tendency of crystals to split into several thin platelets upon manipulation and to be heterogeneous in unit-cell parameters. Monoclinic crystals in fact displayed at least two different unit cells (*A* and *B* forms). Different protocols of dehydration and iterated cryocooling techniques did not improve either the anisotropy or the resolution limit (Samygina *et al.*, 2000).

All diffraction data were indexed and scaled with the *HKL* package (Otwinowski, 1993), *MOSFLM* and *SCALA* (from the *CCP*4 package; Collaborative Computational Project, Number 4, 1994). Data-collection details, scaling statistics and main crystallographic parameters are summarized in Table 1.

3. Results and discussion

Large-scale expression of His-tagged XendoU resulted in soluble protein that was heterogeneously aggregated, a condition that affects crystallization (Wilson, 2003). To obtain monodisperse protein, several conditions were explored. Only the addition of chelating agents (100 m*M* imidazole or 20 m*M* EDTA) yielded homogeneously disperse protein. We attributed the polydispersity to the formation of intermolecular bridges between His tags mediated by divalent metal ions present as contaminants. Consistently, the addition of Mn^{2+} to a

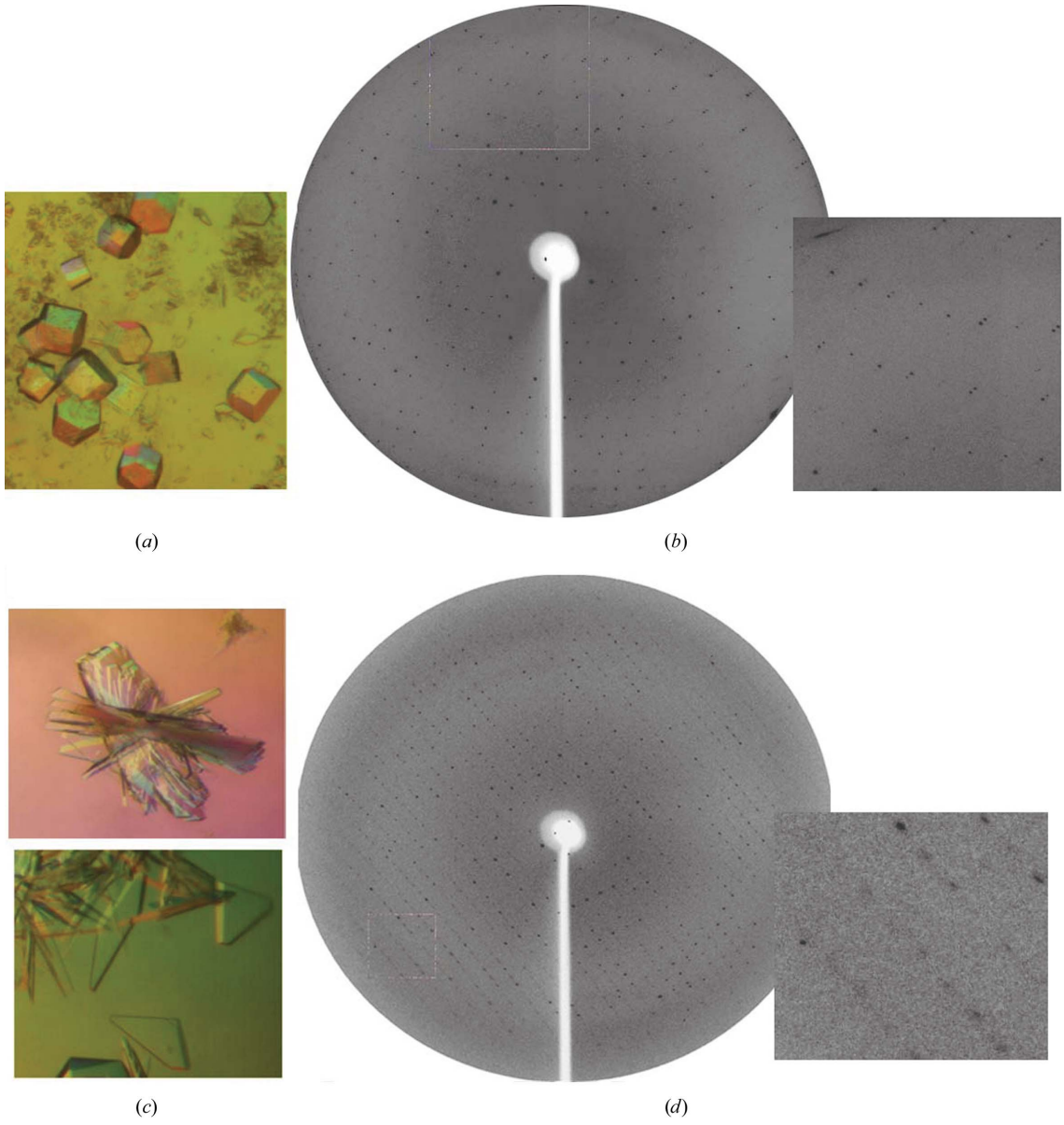


Figure 2
(*a*) Trigonal crystals in 40% $(NH_4)_2SO_4$, 0.2 *M* sodium citrate pH 6, 17 mg ml^{-1} protein, 50 m*M* UMP at 293 K and (*b*) their diffraction image with an enlargement. (*c*) Monoclinic crystals in 40% NH_4SO_4, 0.2 *M* sodium phosphate pH 5.5, 17 mg ml^{-1} protein, 50 m*M* UMP at 293 K improved by seeding from aggregated (top) to single crystals (bottom) and (*d*) their diffraction image with an enlargement.

protein preparation which had been made monomeric using EDTA was sufficient to induce aggregation again (Fig. 1).

To confirm the hypothesis of intermolecular bridges involving His tags *via* divalent metals and to exclude that possibility that these interactions could involve a site on the protein other than the His tag, we expressed a cleavable GST-fused XendoU under the same conditions as the His-tagged enzyme. Analytic HPLC gel filtration both on methacrylate and on silica gel showed the protein to be completely monodisperse even after addition of Mn^{2+}, thus demonstrating that aggregation was dependent on the His tag and not on other metal-binding sites on XendoU.

Screening for crystallization of non-tagged XendoU after GST cleavage only yielded crystals using the same conditions as for the tagged proteins, but with poorer or unimproved diffraction. In summary, it was possible to reverse the aggregation mediated by the His tag through the addition of EDTA or imidazole, allowing the growth of crystals suitable for structural analysis. Use of chelating agents would be a less expensive and time-consuming alternative to specific proteolytic cleavage of the His-tagged protein, which in the case of XendoU was not effective for crystal improvement. This information may be valuable especially in high-throughput structural projects, where a large number of proteins are expressed, purified and screened for crystallization.

Data were collected at the Elettra (Trieste, Italy) and ESRF (Grenoble, France) synchrotron facilities. This work was partially supported by MIUR of Italy (FIRB/RBLA03B3KC-004, FIRB/RBNE015MPB and PRIN/RBNE01KXC9 and Centro di Eccellenza BEMM).

References

Caffarelli, E., Arese, M., Santoro, B., Fragapane, P. & Bozzoni, I. (1994). *Mol. Cell. Biol.* **14**, 2966–2974.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Filipowicz, W. & Pogacic, V. (2002). *Curr. Opin. Cell Biol.* **14**, 319–327.
Fragapane, P., Prislei, S., Michienzi, A., Caffarelli, E. & Bozzoni, I. (1993). *EMBO J.* **12**, 2921–2928.
Gioia, U., Laneve, P., Dlakic, M., Arceci, M., Bozzoni, I. & Caffarelli, E. (2005). *J. Biol. Chem.* **280**, 18996–19002.
Ivanov, K. A., Hertzig, T., Rozanov, M., Bayer, S., Thiel, V., Gorbalenya, A. E. & Ziebuhr, J. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 12694–12699.
Laneve, P., Altieri, F., Fiori, M. E., Scaloni, A., Bozzoni, I. & Caffarelli, E. (2003). *J. Biol. Chem.* **278**, 13026–13032.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Renzi, F., Filippini, D., Loreni, F., Bozzoni, I. & Caffarelli, E. (2002). *Biochem. Biophys. Acta*, **1575**, 26–30.
Samygina, V. R., Antonyuk, S. V., Lamzin, V. S. & Popov, A. N. (2000). *Acta Cryst.* D**56**, 595–603.
Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L. M., Guan, Y., Rozanov, M., Spaan, W. J. M. & Gorbalenya, A. E. (2003). *J. Mol. Biol.* **331**, 991–1004.
Wilson, W. W. (2003). *J. Struct. Biol.* **142**, 56–65.