STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Crystallization and crystallographic analysis of an *Arabidopsis* nuclear proteinaceous RNase P

Franziska Pinker,[a,b] Philippe Giegé[b] and Claude Sauter[a]*

[a]UPR 9002 du CNRS, IBMC, Centre National de la Recherche Scientifique, University of Strasbourg, 15 Rue René Descartes, 67084 Strasbourg, France, and [b]UPR 2357 du CNRS, IBMP, Centre National de la Recherche Scientifique, University of Strasbourg, 12 Rue du Général Zimmer, 67084 Strasbourg, France. *Correspondence e-mail: c.sauter@ibmc-cnrs.unistra.fr

Received 24 July 2015
Accepted 11 September 2015

Edited by T. Bergfors, Uppsala University, Sweden

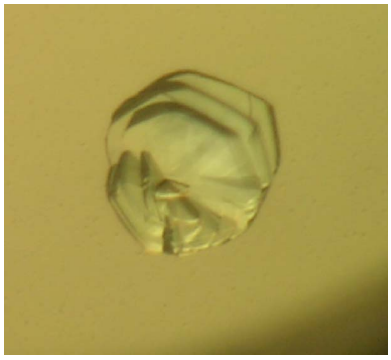
Keywords: RNase P; nuclear PRORP; PPR; tRNA maturation enzyme.

RNase P activity is ubiquitous and involves the 5′ maturation of precursor tRNAs. For a long time, it was thought that all RNases P were ribonucleoproteic enzymes. However, the characterization of RNase P in human mitochondria and in plants revealed a novel kind of RNase P composed of protein only, called PRORP for 'proteinaceous RNase P'. Whereas in human mitochondria PRORP has two partners that are required for RNase P activity, PRORP proteins are active as single-subunit enzymes in plants. Three paralogues of PRORP are found in *Arabidopsis thaliana*. PRORP1 is responsible for RNase P in mitochondria and chloroplasts, while PRORP2 and PRORP3 are nuclear enzymes. Here, the purification and crystallization of the *Arabidopsis* PRORP2 protein are reported. Optimization of the initial crystallization conditions led to crystals that diffracted to 3 Å resolution.

1. Introduction

RNase P is the universal enzyme catalyzing the maturation of precursor 5′-ends of transfer RNAs (tRNAs), as well as many other substrates such as ribosomal RNA, messenger RNA, transfer-messenger RNA and riboswitches (Hartmann & Hartmann, 2003; Lai *et al.*, 2010). It was first described in bacteria as a ribonucleoprotein (RNP) complex, in which the RNA component (P RNA) possesses the catalytic activity (Guerrier-Takada *et al.*, 1983). Since RNP RNase P was subsequently found in all three domains of life, *i.e.* bacteria, archaea and eukarya, it was believed to be a universally conserved RNP complex (Altman, 2007). However, early experiments in spinach chloroplasts and subsequently in human mitochondria suggested the existence of a proteinaceous type of RNase P that lacks an essential RNA component (Wang *et al.*, 1988; Rossmanith & Karwan, 1998; Pinker *et al.*, 2013). In 2008, Rossmanith and coworkers demonstrated by recombinant expression of human mitochondrial RNase P in *Escherichia coli* that it is a complex of three protein subunits with no RNA component (Holzmann *et al.*, 2008). In *Arabidopsis* this novel variant is composed of single proteins with RNase P activity that were named PRORP (for proteinaceous RNase P; Gobert *et al.*, 2013).

The discovery of PRORP enzymes led to the question whether the RNP and protein enzymes have evolved different modes of action to catalyze the same reaction. Biophysical and crystallographic studies of *Arabidopsis* PRORP enzymes have recently shown that they consist of two main domains: an N-terminal pentatricopeptide-repeat (PPR) domain and a C-terminal catalytic metallonuclease domain linked by a zinc-binding module (Fig. 1; Gobert *et al.*, 2010; Howard *et al.*,



2012). The PPR domain confers RNA binding and substrate specificity and consists of five PPR motifs folded into a pair of antiparallel α-helices (Giegé, 2013). The most conserved part of the PRORP enzymes is their C-terminal metallonuclease domain, which belongs to the large family of NYN (N4BP1, YacP-like nuclease) ribonucleases (Anantharaman & Aravind, 2006). The crystal structure of *A. thaliana* PRORP1 (*At*PRORP1) revealed the presence of two magnesium ions that are bound to conserved Asp residues of the NYN metallonuclease domain and are involved in catalysis. PRORP1 is an organellar enzyme that is active both in mitochondria and chloroplasts (Gobert *et al.*, 2010). *Arabidopsis* also expresses two closely related paralogues, *At*PRORP2 and *At*PRORP3, that are responsible for tRNA maturation in the nucleus.

In order to obtain deeper insights into the diversity of function and structural organization of PRORP enzymes, we initiated the crystallographic characterization of a nuclear

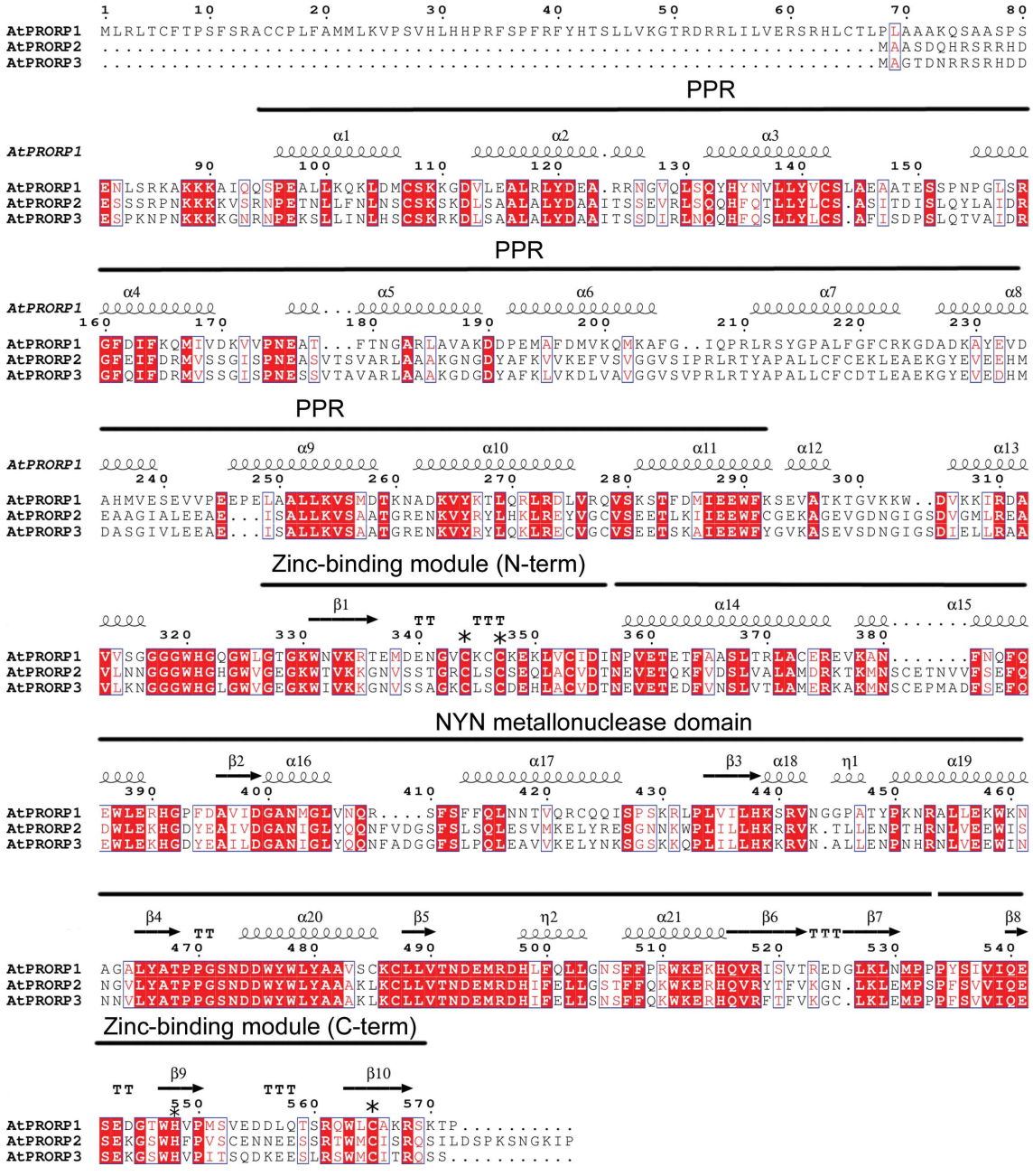


Figure 1
Alignment of the genomic sequences of PRORP1, PRORP2 and PRORP3 from *A. thaliana*. The organellar PRORP1 shares 47% identity with either of the two nuclear forms, while the nuclear PRORP2 and PRORP3 share 80% identity with each other. The secondary structures observed in the PRORP1 X-ray structure (PDB entry 4g26) are displayed. The crystallized PRORP1 construct was deprived of its N-terminal organellar targeting sequence and started at residue 92. The full-length PRORP2 construct used in this work started with MGA (instead of MA) and contained a C-terminal affinity tag (see §2). The PPR (RNA-binding) domain, the NYN metallonuclease (catalytic) domain and the bipartite zinc-binding module (Cys/His residues involved in zinc coordination are marked with asterisks) are indicated. This figure was prepared using *ESPript* (Robert & Gouet, 2014).

proteinaceous RNase P. The purification and crystallization of the nuclear *At*PRORP2 protein are reported. Optimization of the initial crystallization conditions led to crystals that diffracted to 3 Å resolution.

2. Materials and methods

2.1. Macromolecule production

The full length cDNA of *At*PRORP2 (NCBI reference sequence NP_179256) was cloned from the *Arabidopsis* transcriptome into pET-28b(+) vector (Novagen) coding for a C-terminal His_6 tag. Owing to cloning, one additional N-terminal amino acid (MGA instead of MA) and two additional amino acids (LE) upstream of the His tag were introduced, resulting in a protein of 537 amino acids (Fig. 1). *Escherichia coli* BL21 (DE3) cells transformed with pET-28-PRORP2 were grown at 310 K in LB medium containing 50 µg ml^{-1} kanamycin until the $OD_{600\ nm}$ reached 0.7. Protein was expressed overnight at 291 K in LB medium with a final concentration of 0.5 m*M* IPTG. The *E. coli* cells were harvested and the pellet was stored at 253 K until purification. 1 g of cells was lysed in 10 ml Ni–NTA buffer *A* [50 m*M* HEPES–Na pH 7.5, 250 m*M* NaCl, 5%(*w*/*v*) glycerol, 5 m*M* imidazole, 10 m*M* β-mercaptoethanol] and protease-inhibitor cocktail (Sigma, catalogue No. S8830; one tablet per 100 ml) with ten cycles of 30 s sonication (followed by 60 s on ice). The

Table 1
Macromolecule-production information.

Source organism	*A. thaliana*
DNA source	At2g16650
Forward primer	CACGGTCAATGGCAAGGTAT
Reverse primer	TGAAAATGTTGCTGGCTGAGA
Expression vector	pET-28b(+), Novagen
Expression system	*E. coli* BL21 (DE3)
Complete amino-acid sequence of the PRORP2 construct	MGAASDQHRSRRHDESSSRPNKKKKVSRNPETNL-LFNLNSCSKSKDLSAALALYDAAITSSEVRLS-QQHFQTLLYLCSASITDISLQYLAIDRGFEIF-DRMVSSGISPNEASVTSVARLAAAKGNGDYAF-KVVKEFVSVGGVSIPRLRTYAPALLCFCEKLE-AEKGYEVEEHMEAAGIALEEAEISALLKVSAA-TGRENKVYRYLHKLREYVGCVSEETLKIIEEW-FCGEKAGEVGDNGIGSDVGMLREAVLNNGGGW-HGHGWVGEGKWTVKKGNVSSTGRCLSCSEQLA-CVDTNEVETQKFVDSLVALAMDRKTKMNSCET-NVVFSEFQDWLEKHGDYEAIVDGANIGLYQQN-FVDGSFSLSQLESVMKELYRESGNNKWPLILL-HKRRVKTLLENPTHRNLVEEWISNGVLYATPP-GSNDDWYWLYAAAKLKCLLVTNDEMRDHIFEL-LGSTFFQKWKERHQVRYTFVKGNLKLEMPSPF-SVVIQESEKGSWHFPVSCENNEESSRTWMCIS-RQSILDSPKSNGKIPLEHHHHHH

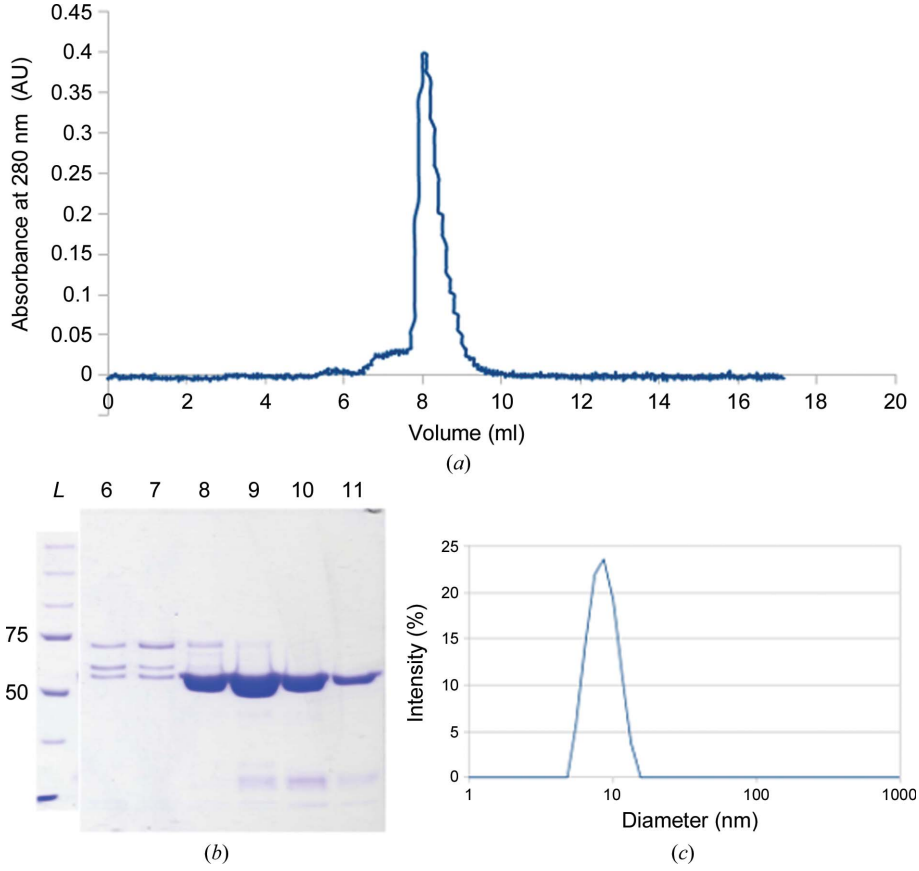


Figure 2
PRORP2 purification and quality control. (*a*) Elution profile of full-length PRORP2 during size-exclusion chromatography (SEC). (*b*) Coomassie-stained 8% SDS–PAGE of PRORP2. Lane *L*, molecular-weight ladder (labelled in kDa); lanes 6–11, fractions from SEC. Fractions 9 and 10 were pooled to perform crystallization assays. (*c*) Monodisperse population of a pure ultracentrifuged PRORP2 sample observed by DLS in the crystallization buffer containing 5%(*w*/*v*) glycerol.

Table 2
Crystallization.

Method	Batch
Plate type	Terazaki microplates (Greiner Bio-One, catalogue No. 653 102)
Temperature (K)	277
Protein stock concentration (mg ml^{-1})	2.5
Buffer composition of protein solution	50 m*M* HEPES–Na pH 7.5, 250 m*M* NaCl, 5%(*w*/*v*) glycerol, 1 m*M* TCEP
Crystallant solution	200 m*M* sodium malonate pH 6, 20% PEG 3350
Drop volume and mixing ratio	2 µl, 1:1

Ni–NTA resin (Sigma, catalogue No. H0537) was equilibrated with buffer *A*. The protein was loaded at 1 ml min^{-1} and was sequentially washed with ten column volumes at 2 ml min^{-1} with increasing imidazole concentrations (steps at 5, 10 and 15 m*M*). The protein was eluted with 2.5 column volumes of buffer *A* supplemented with 250 m*M* imidazole at 2 ml min^{-1}. Fractions containing the protein were pooled, concentrated in SEC buffer [50 m*M* HEPES–Na pH 7.5, 250 m*M* NaCl, 15%(*w*/*v*) glycerol, 1 m*M* TCEP] on a cellulose membrane (30K MWCO, Amicon Ultra, Millipore) and further purified using size-exclusion chromatography at 0.5 ml min^{-1} on a Superdex 200 10/300 GL (GE Healthcare) column equilibrated with SEC buffer (Fig. 2*a*). Fractions containing PRORP2 were concentrated, ultracentrifuged and stored at 277 K until use. The average purification yield was about 1 mg of pure *At*PRORP2 per gram of host cells. Sample purity and homogeneity were assessed by Coomassie-stained SDS–PAGE (Fig. 2*b*) and dynamic light scattering at 293 K (Fig. 2*c*) after ultracentrifugation (1 h, $10^5 g$, 277 K) using a Zetasizer NanoSeries Nano-S instrument (Malvern). Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

The search for crystallization conditions was carried out by vapour diffusion using a Mosquito Crystal pipetting robot (TTP Labtech, UK) and 96-well sitting-drop plates (CrystalEx microplate, conical flat bottom, five subwells, Corning). Initial screening was performed with PRORP2 at 5–10 mg ml^{-1} in SEC buffer. This concentration turned out to be excessive because the protein precipitated in many drops. Further experiments were performed at 2–5 mg ml^{-1} and the SEC buffer was modified to contain only 5%(*w*/*v*) glycerol instead of 15%(*w*/*v*). Different screens were tested by mixing 150 nl protein solution with 150 nl crystallant solution and equilibrating at 277 and 293 K over a reservoir containing 35 µl crystallant solution. The tested screens were JCSG+, Clear Strategy Screen 1, Morpheus (Molecular Dimensions), Magic1 and Magic2 (MPI, Martinsried, Germany), PEG/Ion (Hampton Research) and Wizard I and II (Emerald Bio). The experiments for optimization were set up manually as micro-batch experiments in Terazaki microplates (Greiner Bio-One, catalogue No. 653 102) or as hanging drops in 24-well Linbro plates (Hampton Research) by mixing the protein and crystallant solutions in ratios of 1:1 or 2:1 with total volumes of 2 and 3 µl, respectively. Various crystallant and protein concentrations, as well as additives and different temperatures (277 and 293 K), were tested. Crystallization information is summarized in Table 2.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Crystal	1	2
Diffraction source	X06DA, SLS	X06DA, SLS
Wavelength (Å)	1.0	1.0
Temperature (K)	100	100
Detector	PILATUS 2M	PILATUS 2M
Crystal-to-detector distance (mm)	300	300
Rotation range per image (°)	0.25	0.25
Total rotation range (°)	360	360
Exposure time per image (s)	0.25	0.25
Space group	*P*1	*P*1
a, *b*, *c* (Å)	70.5, 72.8, 80.3	68.1, 72.5, 80.4
α, β, γ (°)	63.1, 72.3, 78.4	63.1, 71.8, 77.3
Mosaicity (°)	0.62	0.29
Resolution range (Å)	50–3.0 (3.19–3.00)	50–3.2 (3.39–3.20)
Total No. of reflections	88462 (15571)	74751 (12431)
No. of unique reflections	25055 (4252)	21035 (3395)
Completeness (%)	92.4 (96.7)	97.7 (98.2)
Multiplicity	3.5 (3.7)	3.2 (3.3)
$\langle I/\sigma(I)\rangle$	8.2 (0.8)	8.8 (1.0)
$R_{\rm meas}$ (%)	14.4 (215)	15.0 (162)
$CC_{1/2}$ (%)	99.8 (36.0)	99.7 (45.6)
Overall *B* factor from Wilson plot (Å^2)	90.3	89.0
Matthews coefficient (Å^3 Da^{-1})	2.9	2.8
Solvent content (%)	57.4	55.6
PRORP2 monomers per asymmetric unit	2	2

2.3. Data collection and processing

For diffraction experiments, crystals were mounted in nylon loops with a diameter of 200 µm (Hampton Research, USA), flash-cooled in the mother liquor and stored in liquid nitrogen. Two diffraction data sets were collected from independent crystals at 100 K with a wavelength of 1 Å using a PILATUS 2M detector at a crystal-to-detector distance of 300 mm on the X06DA beamline at the SLS, PSI, Villigen, Switzerland. Diffraction images were processed using the *XDS* package (Kasbch, 2010) and data-collection statistics are summarized in Table 3. The solvent content was estimated using the *CCP*4 package and the self-rotation function was analyzed with *GLRF* (Winn *et al.*, 2011; Tong & Rossmann, 1997).

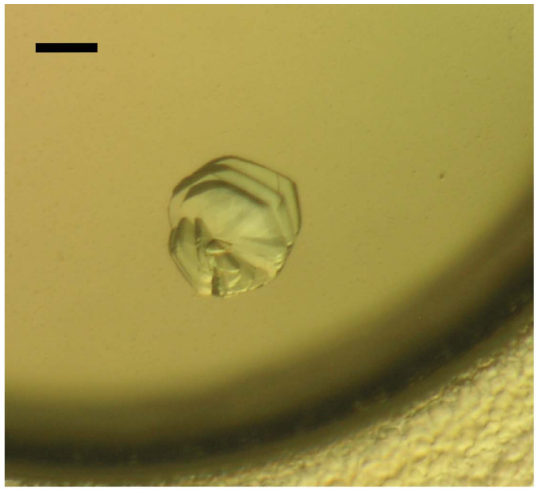
3. Results and discussion

Full-length PRORP2 including its N-terminal nuclear localization sequence and a C-terminal His tag (Fig. 1) was expressed in *E. coli* and purified to homogeneity using affinity and size-exclusion chromatography. Analysis of pure PRORP2 by DLS indicated that the protein was stable, homogeneous and monodisperse in solution, with a hydrodynamic radius of 4.5 nm (Fig. 2).

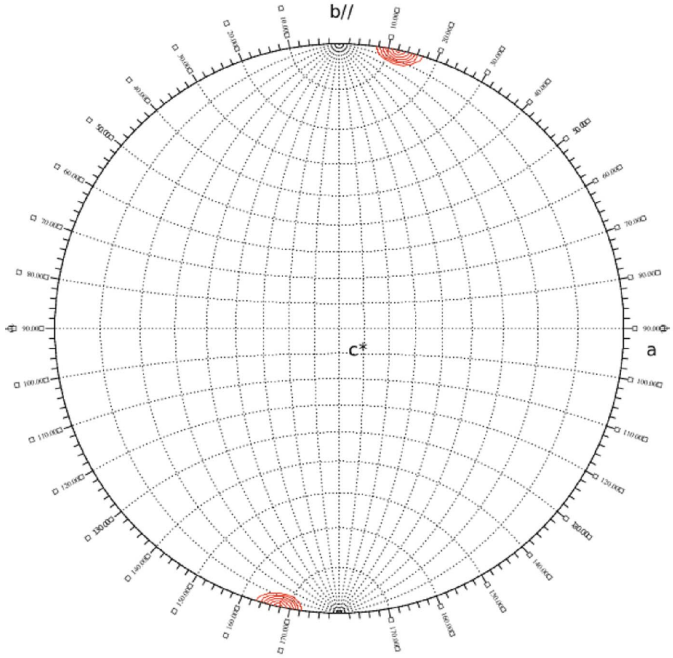
Initial crystallization screening was carried out with PRORP2 (5–10 mg ml^{-1}) in SEC buffer at 293 K, and yielded many precipitated drops but no crystal-like hits. We therefore

changed the protein storage buffer and reduced the glycerol concentration from 15 to 5%(*w*/*v*) to favour crystallization. We also set up all screens at two temperatures: 277 and 293 K. Small intertwined plate-like crystals were obtained after several days at 277 K in condition E6 of the PEG/Ion screen (Hampton Research). The presence of PRORP2 in the crystals was assessed using SDS–PAGE and mass spectrometry (result not shown). The initial crystals diffracted to 7 Å resolution, but the data could not be integrated. This condition, consisting of 200 m*M* sodium malonate pH 6, 20%(*w*/*v*) PEG 3350, was used as a starting point for further optimization.

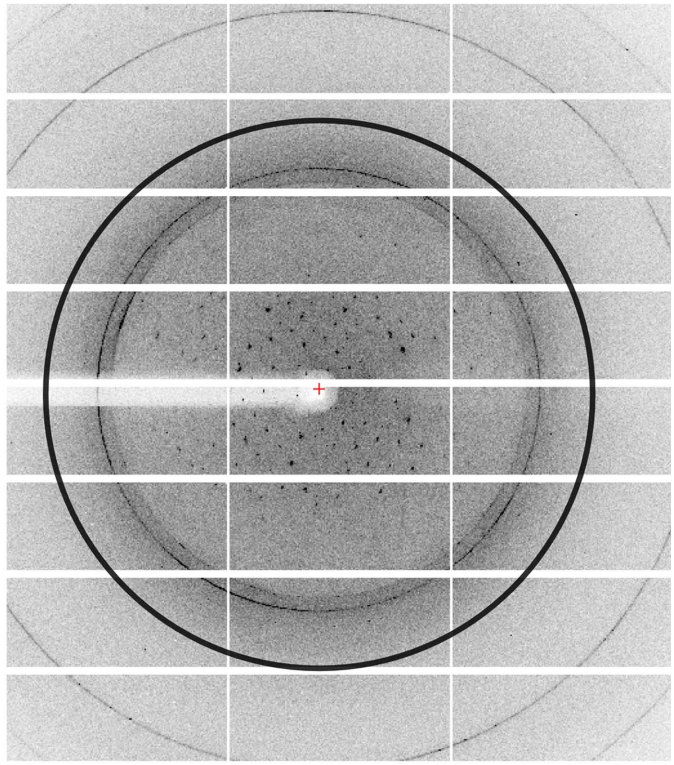
To decrease the number of crystals per drop and to favour their growth in three dimensions, we performed grid screening: the protein concentration was varied from 2 to 10 mg ml^{-1}, the PEG concentration from 16 to 25%(*w*/*v*) PEG 3350 and the salt concentration from 100 to 200 m*M*. Various additives were tested using Additive Screen (Hampton Research). The optimized conditions, which were initially defined by vapour diffusion, were adapted to the batch method to speed up crystal production. This did not change the plate-like morphology of the crystals or their strong tendency to grow in stacks (Fig. 3*a*). However, the largest specimens could be separated using a cryoloop for flash-cooling in liquid nitrogen and diffracted to 3 Å resolution. Using the crystals shown in Fig. 3, with average dimensions of 200 × 100 × 10 µm, two complete monoclinic data sets were collected. No twinning was detected in either data set. They could not be merged owing to non-isomorphism. Estimation of the solvent content (Table 3) and self-rotation function (Fig. 3) analysis suggested the presence of two molecules in the asymmetric unit related by a twofold noncrystallographic symmetry axis. Molecular replacement with homology models derived from the crystal structure of *A. thaliana* PRORP1 (PDB entry 4g26; Howard *et al.*, 2012) is in progress.



(*a*)



(*b*)



(*c*)

Figure 3
Characterization of PRORP2 crystals. (*a*) Stack of three crystals of *Arabidopsis* PRORP2 grown by batch in 100 m*M* sodium malonate pH 6, 10%(*w*/*v*) PEG 3350 (scale bar = 0.1 mm). Crystals were separated using a cryoloop prior to cryocooling and data collection. (*b*) Representative X-ray diffraction image collected on the X06DA beamline at the SLS synchrotron, Switzerland. The black circle corresponds to the 3 Å resolution limit. (*c*) Projection of the self-rotation function (κ = 180°, resolution = 4–8 Å) along the *c** direction. The red marker indicates the orientation of the twofold non-crystallographic rotation axis.

Acknowledgements

This work was supported by the French Centre National de la Recherche Scientifique, by the University of Strasbourg, by an

ANR Blanc research grant 'PRO-RNase P' (ANR-11-BSV8-008-01/02) to PG and CS and by the LabEx consortium 'MitoCross' in the frame of the French National Programme 'Investissement d'Avenir' (ANR-11-LABX-0057_MITOCROSS). FP was supported by an IdEx PhD grant from the University of Strasbourg.

References

Altman, S. (2007). *Mol. Biosyst.* **3**, 604–607.
Anantharaman, V. & Aravind, L. (2006). *RNA Biol.* **3**, 18–27.
Giegé, P. (2013). *RNA Biol.* **10**, 1417–1418.
Gobert, A., Gutmann, B., Taschner, A., Gössringer, M., Holzmann, J., Hartmann, R. K., Rossmanith, W. & Giegé, P. (2010). *Nature Struct. Mol. Biol.* **17**, 740–744.
Gobert, A., Pinker, F., Fuchsbauer, O., Gutmann, B., Boutin, R., Roblin, P., Sauter, C. & Giegé, P. (2013). *Nature Commun.* **4**, 1353.
Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983). *Cell*, **35**, 849–857.
Hartmann, E. & Hartmann, R. K. (2003). *Trends Genet.* **19**, 561–569.
Holzmann, J., Frank, P., Löffler, E., Bennett, K. L., Gerner, C. & Rossmanith, W. (2008). *Cell*, **135**, 462–474.
Howard, M. J., Lim, W. H., Fierke, C. A. & Koutmos, M. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 16149–16154.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Lai, L. B., Vioque, A., Kirsebom, L. A. & Gopalan, V. (2010). *FEBS Lett.* **584**, 287–296.
Pinker, F., Bonnard, G., Gobert, A., Gutmann, B., Hammani, K., Sauter, C., Gegenheimer, P. A. & Giegé, P. (2013). *RNA Biol.* **10**, 1457–1468.
Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
Rossmanith, W. & Karwan, R. M. (1998). *Biochem. Biophys. Res. Commun.* **247**, 234–241.
Tong, L. & Rossmann, M. G. (1997). *Methods Enzymol.* **276**, 594–611.
Wang, M. J., Davis, N. W. & Gegenheimer, P. (1988). *EMBO J.* **7**, 1567–1574.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.