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

Crystallization and preliminary crystallographic studies of putative RNA 3′-terminal phosphate cyclase from the crenarchaeon *Sulfolobus tokodaii*

Satoru Shimizu,[a] Masanori Ohki,[a] Nami Okubo,[a] Kaoru Suzuki,[b] Masaru Tsunoda,[c] Takeshi Sekiguchi[b] and Akio Takénaka[a,c]*

[a]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 226-8501 Yokohama, Japan, [b]College of Science and Engineering, Iwaki-Meisei University, 970-8551 Fukushima, Japan, and [c]Faculty of Pharmacy, Iwaki-Meisei University, 970-8551 Fukushima, Japan

Correspondence e-mail: atakenak@iwakimu.ac.jp

Received 26 February 2009
Accepted 20 April 2009

RNA 3′-terminal phosphate cyclase (Rtc) is an enzyme involved in RNA splicing that converts the 3′-terminal hydroxyl group of truncated RNA to 2′,3′-cyclic phosphate, which is required just before its ligation. This reaction may occur in the following two steps: (i) Rtc + ATP → Rtc-AMP + PP_i and (ii) RNA-N3′p + Rtc-AMP → RNA-N>p + Rtc + AMP. In order to reveal the reaction mechanism, Rtc of *Sulfolobus tokodaii* (*St*-Rtc) overexpressed in *Escherichia coli* was purified and crystallized in the following states: *St*-Rtc, *St*-Rtc+Mn, *St*-Rtc+ATP, *St*-Rtc+AMP and *St*-Rtc-AMP. The crystals diffracted to 2.25–3.00 Å resolution and preliminary solutions of their structures have been obtained by molecular replacement using the structure of a selenomethionine-labelled *St*-Rtc crystal which was solved in advance using the MAD method as a model. These crystals grew in two different space groups ($P3_1$ and $P4_2$), with the former space group displaying two distinct packing modes.

1. Introduction

RNA 3′-terminal phosphate cyclase (Rtc) was initially identified in extracts of HeLa cells and *Xenopus* oocytes (Vicente & Filipowicz, 1988; Filipowicz & Shatkin, 1983). This enzyme is localized in the nucleoplasm, which is consistent with its postulated role in RNA processing. This enzyme catalyzes the ATP-dependent conversion of a 3′-phosphate group to a 2′,3′-terminal phosphodiester group at the 3′-end of various RNAs truncated by endonucleases. It has been proposed that the cyclization of the 3′-phosphate occurs in the following two steps:

(i) Rtc + ATP → Rtc-AMP + PP_i

(ii) RNA-N3′p + Rtc-AMP → RNA-N>p + Rtc + AMP.

In step (i), Rtc and ATP react with each other to form a covalently bonded Rtc-AMP intermediate in the presence of Mg^{2+} or Mn^{2+} ion. Step (ii) was proposed on detecting the formation of RNA-$N^{3'}pp^{5'}$A when the ribose at the RNA 3′-terminus was replaced by 2′-deoxy- or 2′-*O*-methylribose (Filipowicz *et al.*, 1985). The cyclization step probably occurs non-enzymatically (Genschik & Filipowicz, 1998; Genschik *et al.*, 1997; Palm *et al.*, 2000).

Rtc genes have been identified in eukaryotes, bacteria and archaea (Genschik *et al.*, 1997). The three-dimensional structure of Rtc of *Escherichia coli* has been determined by X-ray crystallography (Palm *et al.*, 2000). The structure is composed of two domains: large and small. The large domain consists of three repeats of the same folding module containing two α-helices and a four-stranded β-sheet. This motif of the module appears in several nucleic acid-binding proteins, for example the C-terminal domain of bacterial initiation factor 3 (IF3; Biou *et al.*, 1995). However, the catalytic mechanism and biological role of Rtc have not yet been clarified. The cellular RNA ligase identified in bacteria also requires 2′,3′-cyclic ends for ligation (Greer *et al.*, 1983; Arn & Abelson, 1996). Therefore, it has been speculated that Rtc catalyzes the cyclization of the 3′-monophosphate-terminal group of truncated RNAs which are formed by decyclization with phosphodiesterase or by hydrolysis of the product

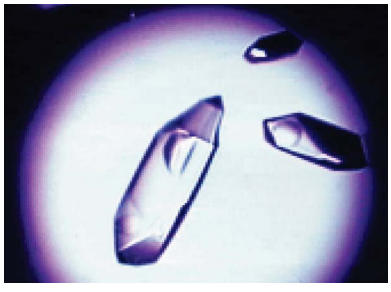

Table 1
Final crystallization conditions.

Crystal	*St*-Rtc*	*St*-Rtc	*St*-Rtc*+Mn	*St*-Rtc*+ATP	*St*-Rtc*-AMP	*St*-Rtc+AMP
Protein solution						
Buffer	0.02 *M* Tris–HCl† pH 8.0	0.02 *M* Tris–HCl pH 8.0	0.02 *M* Tris–HCl pH 8.0	0.02 *M* Tris–HCl pH 8.0	0.02 *M* Tris–HCl pH 8.0	0.02 *M* Tris–HCl pH 8.0
Concentration (mg ml^{-1})	10	5	5	5	10	5
Additive	—	—	2 m*M* $MnCl_2$	0.5 m*M* ATP 2 m*M* $MnCl_2$	0.5 m*M* ATP	0.5 m*M* ATP 2 m*M* $MnCl_2$
Reservoir solution						
Buffer	0.1 *M* Tris–HCl pH 8.5	0.1 *M* CH_3COONa–HCl pH 4.6‡	0.1 *M* CH_3COONa–HCl pH 4.6	0.1 *M* Tris–HCl pH 8.0	0.1 *M* Tris–HCl pH 8.0	0.1 *M* CH_3COONa–HCl pH 4.6
Salt	2.0 *M* $NH_4H_2PO_4$	2.0 *M* $(NH_3)_2SO_4$	2.0 *M* $(NH_3)_2SO_4$	2.1 *M* $(NH_3)_2SO_4$	2.0 *M* $(NH_3)_2SO_4$	2.0 *M* $(NH_3)_2SO_4$
pH of droplet	4.8§	5.7	5.7	7.9	8.2	5.7
Approximate size (µm)	200 × 100 × 50	200 × 100 × 50	100 × 50 × 50	200 × 50 × 50	400 × 50 × 50	200 × 100 × 100

† Tris–HCl: tris(hydroxymethyl)aminomethane titrated using HCl. ‡ CH_3COONa–HCl, sodium acetate titrated using HCl. § The high concentration of the acidic additive $NH_4H_2PO_4$ changed the buffer pH from basic to acidic.

of the above reactions, in order to regenerate the substrates in the subsequent ligation step.

In order to reveal the reaction mechanism of Rtc and its biological role by X-ray analyses, Rtc from *Sulfolobus tokodaii* (*St*-Rtc) and its selenomethione derivative (*St*-Rtc*[1]) have been overexpressed in *E. coli* and crystallized in the following states: apo forms of *St*-Rtc and *St*-Rtc* to resolve the phase problem and to determine the free state of the enzyme as well as the effect of selenium substitution on the structure; a complex form with Mn^{2+} ion (*St*-Rtc*+Mn) to identify the Mg^{2+} ion-binding site; a complex with ATP but no Mg^{2+} ion (*St*-Rtc*+ATP) to observe the ATP-bound state just before the first reaction; a complex with Mg^{2+} and ATP under basic conditions (*St*-Rtc*-AMP) and a complex with Mg^{2+} and ATP under acidic conditions (*St*-Rtc+AMP)[2]. The latter two structures will be of great interest in order to obtain insight into the effect of pH on the first reaction, since its optimal pH value has been reported to be 8–9 (Genschik *et al.*, 1997; Palm *et al.*, 2000) for the Rtcs from HeLa cells and *E. coli*.

2. Materials and methods

2.1. Overexpression and purification of *St*-Rtc and *St*-Rtc*

The gene encoding *St*-Rtc (gene ID ST0570, 339 amino acids, 37 505 Da, pI 9.89) was inserted into the pET-11a vector (Novagen). The recombinant plasmid was transformed in *E. coli* Rosetta-gami (DE3) (Novagen) and the cells were grown in LB culture at 310 K without induction. Following overnight incubation, the cells were harvested by centrifugation at 6000 rev min^{-1} for 10 min at 277 K and disrupted by sonication in 20 m*M* Tris–HCl pH 8.0 buffer. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18 000 rev min^{-1} for 20 min at 277 K, the supernatants were dialyzed against the same buffer containing 1.75 *M* ammonium sulfate and then applied onto a Resource ISO column (Amersham Bioscience). Bound proteins were eluted at a 2 ml min^{-1} flow rate with a linear gradient of 1.75–0 *M* ammonium sulfate in the same buffer. The pooled fractions were dialyzed against 50 m*M* MES (2-morpholinoethanesulfonic acid monohydrate titrated using NaOH) buffer pH 6.0 containing 5 m*M* β-mercaptoethanol and loaded onto a Resource S column (Amersham Bioscience); the bound proteins were eluted at a 4 ml min^{-1} flow rate with a 0–0.4 *M* NaCl linear gradient in the same dialysis buffer.

After exchange of the buffer to 20 m*M* Tris–HCl pH 8.0 buffer containing 5 m*M* β-mercaptoethanol and 150 m*M* NaCl, the pooled fractions were applied onto a HiLoad 16/60 Superdex 75 column (Amersham Bioscience) and the target proteins were eluted at a 1 ml min^{-1} flow rate. After buffer exchange to 10 m*M* potassium phosphate pH 7.0 containing 5 m*M* β-mercaptoethanol, the pooled fractions were finally applied onto a BioScale CH51 column (Bio-Rad) and the proteins were eluted at a 2 ml min^{-1} flow rate using a linear 10–400 m*M* potassium phosphate concentration gradient at pH 7.0.

The proteins were concentrated to a final concentration of 5 mg ml^{-1} (A_{280} = 1.56) in 20 m*M* Tris–HCl pH 8.0 using centrifugal filter devices (Vivaspin 500, Sartorius Stedim Biotech). The fractions obtained in all the purification steps as well as the concentrated proteins were analyzed by SDS–PAGE. pH values and UV absorption spectra were measured using F-13 (Horiba Ltd, Japan) and BioSpec-mini (Shimadzu Corporation, Japan), respectively.

The selenomethionine derivative was expressed in *E. coli* B834 (DE3) (Novagen) using minimal M9 medium containing L-seleno-

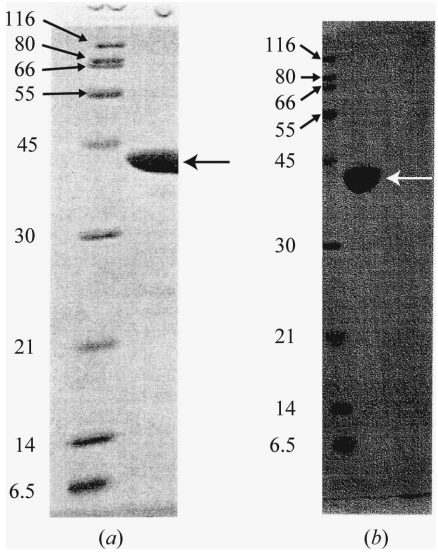

Figure 1
SDS–PAGE of *St*-Rtc (*a*) and *St*-Rtc* (*b*). The target proteins are indicated by arrows on the right. Values indicate the approximate molecular weights of the markers (kDa).

[1] The asterisk indicates selenomethionine derivative.
[2] Here, we assumed that the additive ATP reacted once to form an intermediate and was then hydrolyzed to AMP under acidic conditions.

methionine. The same procedures and conditions as described above for the case of *St*-Rtc were used for the cultivation, isolation and purification of *St*-Rtc*. The only difference was the order of column chromatography after the Resource S column. The eluted fractions were first applied onto a BioScale CH51 column (Bio-Rad) and the bound proteins were eluted at a 2 ml min^{-1} flow rate with a 10–400 m*M* potassium phosphate pH 7.0 linear gradient containing 5 m*M* β-mercaptoethanol. The eluted fractions were finally applied onto a HiLoad 16/60 Superdex 75 column (Amersham Bioscience) and the target proteins were eluted at a 1 ml min^{-1} flow rate using 20 m*M* Tris–HCl pH 8.0 buffer containing 5 m*M* β-mercaptoethanol and 150 m*M* NaCl. The protein concentration was adjusted to 10 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0 using the same method as described above. The purity was analyzed by SDS–PAGE.

2.2. Crystallization

Crystallization trials were carried out using the hanging-drop vapour-diffusion method by mixing equal volumes (1 µl) of the protein and reservoir solutions and equilibrating the mixed solutions against 700 µl reservoir solution in a 24-well plate (Stem Corporation, Japan) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt and by changing the pH of the buffer. The prepared crystals were (1) the apo form of *St*-Rtc*, (2) the apo form of *St*-Rtc, (3) a complex with Mn^{2+} ion (*St*-Rtc*+Mn), (4) a complex with ATP but not Mg^{2+} under basic conditions (*St*-Rtc*+ATP), (5) a complex with ATP and Mg^{2+} under basic conditions (*St*-Rtc*-AMP) and (6) a complex with ATP and Mg^{2+} under acidic conditions (*St*-Rtc+AMP). In the cases of (5) and (6), the protein solution was incubated for 30 s to allow the first reaction. The droplets were then mixed with the reservoir solutions, basic or acidic, to maintain or change the pH of the droplet, respectively. For crystallizations, 24-well plates (Stem Corporation, Japan), plain glass cover slides (diameter 18 mm circle, Matsunami Glass Ind. Ltd, Japan) coated with silicon (L-25, Fuji Systems Corporation, Japan) and high-vacuum grease (Dow Corning Toray Co. Ltd, Japan) were used. The optimized crystallization conditions for the six types of crystals are given in Table 1.

Table 2
Crystallographic data of the *St*-Rtc* crystal and statistics of the observed intensity data at four wavelengths.

Values in parentheses are for the highest resolution shell.

Data	Peak	Edge	High remote	Low remote
Wavelength (Å)	0.97898	0.97931	0.96408	0.98319
Resolution (Å)	50–2.00 (2.07–2.00)	50–2.10 (2.18–2.10)	50–2.20 (2.28–2.20)	50–2.20 (2.28–2.20)
Space group	$P3_1$	$P3_1$	$P3_1$	$P3_1$
Unit-cell parameters				
$a = b$ (Å)	84.0	84.1	84.2	84.2
c (Å)	109.2	109.2	109.3	109.3
Z†	2	2	2	2
Observed reflections	657599	584648	511916	512338
Unique reflections	58120	50413	43942	43942
Completeness (%)	99.7 (96.9)	100 (100)	100 (100)	100 (100)
Redundancy	5.7 (4.6)	5.8 (5.5)	5.8 (5.7)	5.8 (5.7)
$I/\sigma(I)$	31.4 (2.8)	34.4 (4.6)	34.3 (5.4)	36.8 (5.8)
$R_{\rm merge}$‡ (%)	6.5 (37.9)	6.3 (31.2)	6.2 (27.9)	6.0 (27.0)

† Number of protein molecules in the asymmetric unit. ‡ $R_{\rm merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

2.3. X-ray data collection and preliminary structure analysis

X-ray experiments with the *St*-Rtc* crystal were performed on beamline BL-17A of Photon Factory (PF; Tsukuba, Japan). A crystal was soaked for 30 s in reservoir solution containing 40% glycerol and mounted on a CryoLoop (Hampton Research). To apply multiple anomalous dispersion (MAD) techniques for phase estimation, four wavelengths (peak, edge, high remote and low remote) were selected

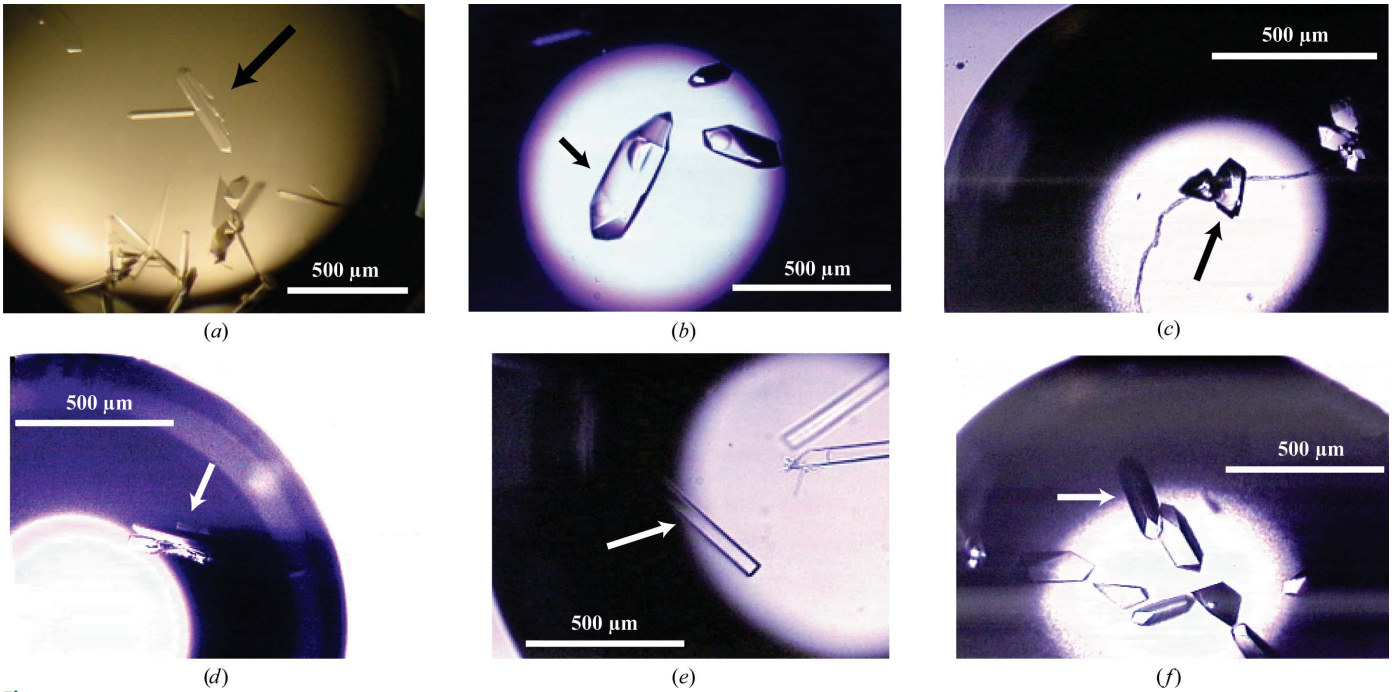

Figure 2
Crystals of *St*-Rtc* (*a*), *St*-Rtc (*b*), *St*-Rtc*+Mn (*c*), *St*-Rtc*+ATP (*d*), *St*-Rtc*-AMP (*e*) and *St*-Rtc+AMP (*f*). Arrows indicate the crystals used for X-ray diffraction experiments.

based on X-ray absorption fine-structure (XAFS) measurement. The diffraction patterns at 100 K were recorded on an ADSC Quantum 4R CCD detector positioned at 210.8, 210.7, 209.3 and 215.0 mm away from the crystal for the four wavelengths, respectively; each frame was taken with a 10 s exposure time and with a 2° oscillation step over the full range of 360°. Bragg spots were indexed and their intensities were estimated by integrating around them. Intensity data were then merged and scaled between the frames with the *HKL*-2000 package (Otwinowski & Minor, 1997). These data were further converted to structure-factor amplitudes using *TRUNCATE* from the *CCP*4 suite (Collaboration Computational Project, Number 4, 1994).

The remaining five crystals, *St*-Rtc, *St*-Rtc*+Mn, *St*-Rtc*+ATP, *St*-Rtc*-AMP and *St*-Rtc+AMP, were mounted on CryoLoops (Hampton Research) with 30–40% glycerol. Their diffraction patterns were taken in the oscillation range 0–180° on the BL-6A and NW-12 beamlines of Photon Factory and were processed using the methods described above.

The MAD method was used to solve the phase problem of the *St*-Rtc* crystal. The programs *SHARP*/*autoSHARP* (Vonrhein *et al.*, 2007), in combination with *SHELXD* (Sheldrick, 2008), *SOLOMON* (Abrahams & Leslie, 1996) and *ARP*/*wARP* (Perrakis *et al.*, 1999), gave a unique plausible solution. The *St*-Rtc* structure was used as a probe in molecular replacement for the other crystals. The programs used were *AMoRe* (Navaza, 1994) and *Phaser* (McCoy *et al.*, 2007).

3. Results and discussion

After the final purification, the protein purity and molecular weights of the target proteins were estimated by SDS–PAGE (Fig. 1). The target proteins appeared as main bands at positions corresponding to the molecular weights (*St*-Rtc, 37 505 Da; *St*-Rtc*, 37 786 Da) estimated from the amino-acid sequence. Thus, it was considered that the two proteins were highly purified for crystallization. Table 1 shows the conditions which produced crystals of *St*-Rtc*, *St*-Rtc and their complexes, while Fig. 2 shows images of the crystals and Fig. 3 shows examples of their diffraction patterns. The crystallographic data and statistics of the observed intensity data are given in Tables 2 and 3. In every crystal, the acceptable Z value was always estimated to be 2;

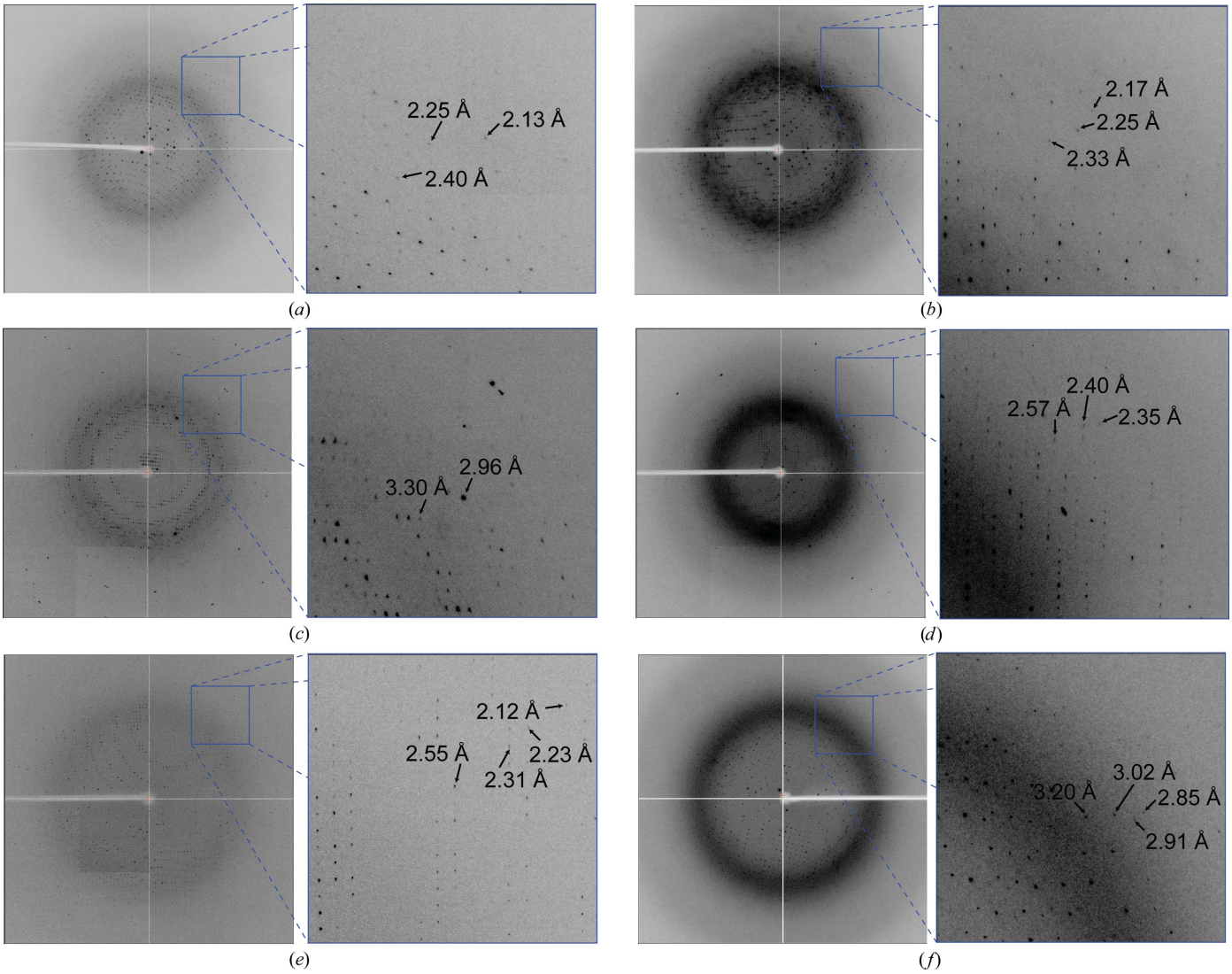

Figure 3
X-ray diffraction images of *St*-Rtc* (*a*), *St*-Rtc (*b*), *St*-Rtc*+Mn (*c*), *St*-Rtc*+ATP (*d*), *St*-Rtc*-AMP (*e*) and *St*-Rtc+AMP (*f*) crystals, each taken with 1° oscillation (left), and an enlargement of the image in the box (right). Arrows indicate diffraction spots at the resolutions indicated.

Table 3
Crystallographic data and statistics of the observed intensity data of the *St*-Rtc, *St*-Rtc*+Mn, *St*-Rtc*+ATP, *St*-Rtc*-AMP and *St*-Rtc+AMP crystals.

Values in parentheses are for the highest resolution shell.

Crystal	*St*-Rtc	*St*-Rtc*+Mn	*St*-Rtc*+ATP	*St*-Rtc*-AMP	*St*-Rtc+AMP
X-ray source	NW12	NW12	NW12	NW12	BL-6A
Wavelength (Å)	1.00	1.00	0.98	1.00	0.98
Resolution (Å)	50–2.25 (2.33–2.25)	50–3.20 (3.31–3.20)	50–2.40 (2.49–2.40)	50–2.25 (2.33–2.25)	50–2.90 (3.00–2.90)
Space group	$P3_1$	$P3_1$	$P4_2$	$P4_2$	$P3_1$
Unit-cell parameters					
$a = b$ (Å)	84.4	83.5	91.1	90.8	84.4
c (Å)	109.4	109.7	89.9	90.3	108.4
Z†	2	2	2	2	2
Observed reflections	239382	82064	211829	265661	110711
Unique reflections	41221	14154	28513	34783	19249
Completeness (%)	99.9 (99.6)	99.7 (99.5)	99.8 (99.8)	99.7 (99.1)	100 (100)
Redundancy	5.8 (5.8)	5.8 (5.8)	7.4 (7.1)	7.6 (7.4)	5.8 (5.7)
$I/\sigma(I)$	43.2 (6.3)	35.3 (11.9)‡	36.6 (7.6)	23.1 (2.8)	30.1 (7.2)
$R_{\rm merge}$§ (%)	5.5 (30.2)	6.5 (17.3)	8.7 (25.8)	8.0 (28.5)	8.1 (31.8)

† Number of protein molecules in the asymmetric unit. ‡ Although the crystal diffracted to 3.0 Å resolution, it was difficult to integrate beyond 3.2 Å resolution because the spots were highly deformed in the outer region owing to damage on flash-freezing. § $R_{\rm merge} = 100 \times \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$.

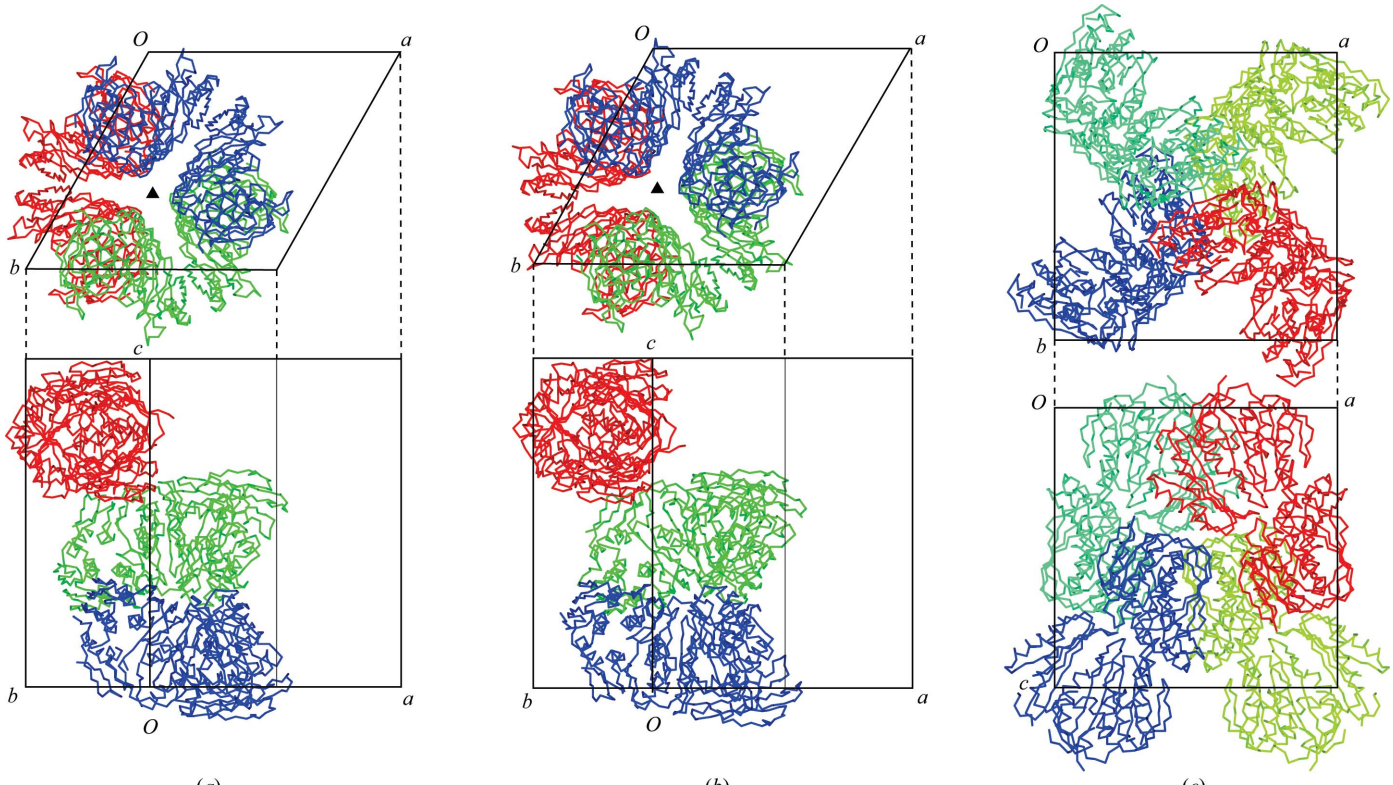

Figure 4
The molecular packing in the unit cells of the *St*-Rtc*, *St*-Rtc*+Mn and *St*-Rtc+AMP crystals (*a*), of the *St*-Rct crystal (*b*) and of the *St*-Rtc*-AMP and *St*-Rtc+AMP crystals (*c*). The two proteins forming a dimer are drawn with C^{α} chains in the same colour.

the $V_{\rm M}$ values were in the range 2.44–2.99 Å^3 Da^{-1} and the solvent content was in the range 49.6–58.9%.

A preliminary crystal structure has been successfully solved by the MAD method and initial refinement yielded a correlation coefficient in $|E^2|$ of 0.822. This structure was then used to solve the remaining five crystal structures by molecular replacement. This confirmed the presence of a noncrystallographic dimer in each crystal form. Interestingly, the crystals grown under acidic conditions (*St*-Rtc, *St*-Rtc*+Mn and *St*-Rtc+AMP) belonged to space group $P3_1$, while the crystals grown under basic conditions (*St*-Rtc*+ATP and *St*-Rtc*-AMP) belonged to space group $P4_2$. This suggests that pH plays a significant role in selecting alternate molecular packings. Moreover, a closer examination of the $P3_1$ structures showed that their packing is not truly isomorphous: the protein molecules stacked along the 3_1 screw axis in the *St*-Rtc crystals are rotated by about 15° around the *c* axis compared with the other $P3_1$ crystal forms. Diagrams of the three crystal-packing arrangements are shown in Fig. 4; these also show that the proteins are reasonably packed with no molecular collisions. Full refinement and structure analysis is in progress and will be reported elsewhere.

This work was supported in part by a Grants-in-Aid for the Protein3000 Research Program from the Ministry of Education,

Culture, Sports, Science and Technology of Japan. We thank S. Kuramitsu for organizing the research group in the program and N. Igarashi and S. Wakatsuki for facilities and help during data collection.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Arn, E. A. & Abelson, J. N. (1996). *J. Biol. Chem.* **269**, 31145–31153.
Biou, V., Shu, F. & Ramakrishnan, V. (1995). *EMBO J.* **14**, 4056–4064.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Filipowicz, W. & Shatkin, A. J. (1983). *Cell*, **32**, 547–557.
Filipowicz, W., Strugala, K., Konarska, M. & Shatkin, A. J. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 1316–1320.
Genschik, K. D. & Filipowicz, W. (1998). *J. Biol. Chem.* **73**, 25516–25526.
Genschik, P., Billy, E., Swianiewicz, M. & Filipowicz, W. (1997). *EMBO J.* **16**, 2955–2967.
Greer, C. L., Javor, B. & Abelson, J. (1983). *Cell*, **33**, 899–906.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Palm, G. J., Billy, E., Filipowicz, W. & Wlodawer, A. (2000). *Structure*, **8**, 13–23.
Perrakis, A., Morris, R. M. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Vicente, O. & Filipowicz, W. (1988). *Eur. J. Biochem.* **176**, 431–439.
Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.