Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Novel crystal form of the ColE1 Rom protein

Se Bok Jang,[a]* Mi Suk Jeong,[a] Richard J. Carter,[b] Elizabeth L. Holbrook,[b] Luis R. Comolli[b] and Stephen R. Holbrook[b]*

[a]Korea Nanobiotechnology Center, Pusan National University, Jangjeon-dong, Keumjeong-gu, Busan 609-735, South Korea, and [b]Structural Biology Department, Physical Biosciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA

Correspondence e-mail: sbjang@pusan.ac.kr, srholbrook@lbl.gov

Received 21 October 2005
Accepted 5 April 2006

PDB Reference: ColE1 Rom protein, 2ghy, r2ghysf.

The RNA I modulator protein (Rom) acts as a co-regulator of ColE1 plasmid copy number by binding to RNA kissing hairpins and stabilizing their interaction. The structure of Rom has been determined in a new crystal form from X-ray diffraction data to 2.5 Å resolution. In this structure, a dimer of the 57-amino-acid protein is found in the asymmetric unit. Each subunit consists almost entirely of two antiparallel α-helices joined by a short hairpin bend. The dimer contains a non-crystallographic twofold axis and forms a highly regular four-α-helical bundle. The structural packing in this novel crystal form is different from previously known Rom structures. The asymmetric unit contains one dimer, giving a crystal volume per protein weight (V_M) of 1.83 Å^3 Da^{-1} and a low solvent content of 30%. Strong packing interactions and low solvation are characteristic of the structure. The Rom protein was cocrystallized with the Tar–Tar* kissing hairpin RNA. Although the electron-density maps do not show bound RNA, altered conformations in the side chains of Rom that are known to be involved in RNA binding have been identified. These results provide additional information about Rom protein conformational flexibility and suggest that the presence of a highly charged polymer such as RNA can promote tight packing of an RNA-binding protein, even when the RNA itself is not observed in the crystal.

1. Introduction

The Rom (RNA I modulator) or Rop (repressor of primer) protein regulates ColE1 plasmid copy number (Itoh & Tomizawa, 1980; Polisky, 1988). It functions by binding to and stabilizing three kissing hairpin pairs that are formed between RNA I and RNA II (Tomizawa & Som, 1984). In the duplex form, RNA II is unable to fulfill its role as primer of DNA replication, plasmid replication is inhibited and copy number decreases (Tomizawa & Itoh, 1981; Tomizawa, 1984). It has been demonstrated that the hairpins alone can form stable pairs in isolation from the rest of RNA I and RNA II and that Rop is able to recognize and bind to these small complexes (Eguchi & Tomizawa, 1991).

The structure of the *Escherichia coli* ColE1 plasmid-encoded protein Rop is known (Banner *et al.*, 1987; Eberle *et al.*, 1990, 1991). Rop is a small homodimeric protein that forms a highly regular four-α-helical bundle. Each subunit consists of two antiparallel α-helices joined by a short loop. The bend region of Rop has attracted considerable interest, mainly because of the ongoing debate about the role of loops in the folding and stability of bundles and proteins in general, and as such it has been subject to numerous mutagenesis experiments (Chou *et al.*, 1992; Predki *et al.*, 1996; Nagi & Regan, 1997; Castagnoli *et al.*, 1994). In Rop$_{<2aa>}$ the discontinuity has been

Table 1
Data and refinement statistics.

Values in parentheses are for the highest resolution bin.

Space group	*C*2
Unit-cell parameters (Å, °)	a = 97.8, b = 39.1, c = 26.7, $\alpha = \gamma$ = 90.0, β = 96.2
Resolution range (Å)	30.0–2.5
Completeness (%)	97.5 (96.2)
Observed reflections	17623
Unique reflections	3577
$I/\sigma(I)$	2.1
$R_{\rm merge}$† (%)	11.7 (25.8)
Protein residues	114
Solvent molecules	33
$\langle B\rangle$, protein (Å^2)	9.9
$\langle B\rangle$, water (Å^2)	22.6
$R_{\rm cryst}$ ($R_{\rm free}$)‡ (%)	17.9 (26.3)
R.m.s.d. bond lengths (Å)	0.008
R.m.s.d. bond angles (°)	2.151
Ramachandran plot (%)	
Most favored regions	96.3
Additionally allowed regions	1.9
Generously allowed regions	0.9
Disallowed regions	0.9

† $R_{\rm merge} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where $I(h,i)$ is the intensity of the ith measurement of reflection h and $\langle I(h)\rangle$ is the mean value of $I(h,i)$ for all i measurements. ‡ $R_{\rm cryst} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$, where $F_{\rm obs}$ denotes the observed structure-factor amplitude and $F_{\rm calc}$ denotes the structure-factor amplitude calculated from the model; 10% of reflections were used to calculate $R_{\rm free}$.

eliminated by the insertion of Ala on both sides of bend residue Asp30 (Vlassi *et al.*, 1994). The two inserted Ala residues establish a continuous heptad pattern, resulting in crystals that diffract to ultrahigh (0.8 Å) resolution (Spyridaki *et al.*, 2000). Crystals of two other Rop mutants have been reported: an A31P mutant (Glykos *et al.*, 1999) and a Δ(30D–34Q) deletion mutant in the loop region (Papanikolau *et al.*, 2004).

Although the structures of many RNA protein-binding motifs have been identified, so far no structure has been reported for the Rop–kissing hairpin RNA complex. A previous attempt to cocrystallize a kissing hairpin complex resulted in crystals and the structure determination of an RNA duplex formed by the annealed kissing hairpins (Klosterman *et al.*, 1999).

Rop binding to the wild-type ColE1 kissing hairpin and its variants has been extensively probed by biochemical methods. Elegant ribonuclease-cleavage studies have shown that Rop binds to the loop residues of the kissing complex in a pseudosymmetrical fashion (Eguchi & Tomizawa, 1990). The affinity constants derived from cleavage rates indicate that Rop can bind to 6, 7 or 8 nt loop complexes with similar affinity and recognize a specific structure rather than a particular sequence (Eguchi & Tomizawa, 1991). Alanine-scanning mutagenesis experiments on Rop binding to the wild-type ColE1 complex and two variants have implicated residues in helix H1 and its symmetrical counterpart H1′. Specifically, Asn10, Phe14, Gln18 and Lys25 abolish binding when substituted by Ala (Predki *et al.*, 1995). Lys3 may also play a role. Additional mutational experiments suggest that Phe14 residues in helices H1 and H1′ interact with loop residues in a pseudosymmetrical fashion (Predki *et al.*, 1995; Castagnoli *et al.*, 1989). These studies have also shown that helix H2 and its symmetry-related helix H2′ do not participate in binding. The mutagenesis studies have been confirmed by heteronuclear two-dimensional NMR spectroscopy mapping of the interface between the Rop protein and the Tar–Tar* kissing complex (Comolli *et al.*, 1998).

We describe the crystal structure of the Rop (referred to hereafter as Rom) protein from the *E. coli* plasmid ColE1 at 2.5 Å resolution and its structural packing in a novel crystal form. We have identified several residues of Rom with altered conformations that are implicated as RNA-binding sites.

2. Materials and methods

2.1. Expression and purification of the Rom protein

E. coli strain BL21(DE3) transformed with plasmid p2R was a gift from L. Regan (Yale University). The Rom protein produced by plasmid p2R contains an additional glycine residue compared with the wild-type protein (Predki *et al.*, 1995). Expression and purification were carried out as described in Predki *et al.* (1995) with minor modifications.

The cell culture was grown in LB medium with kanamycin (25 µg ml^{-1}) at 310 K. Rom was subsequently purified by reverse-phase HPLC using a C-18 column (Vydac) to eliminate residual RNAse activity. On average, between 20 and 30 mg purified protein were obtained per litre of medium. All buffers and solutions were prepared as described in Predki *et al.* (1995).

Electrospray mass spectra revealed two species that differed by the presence or absence of an N-terminal methionine residue. The relative amounts of each species

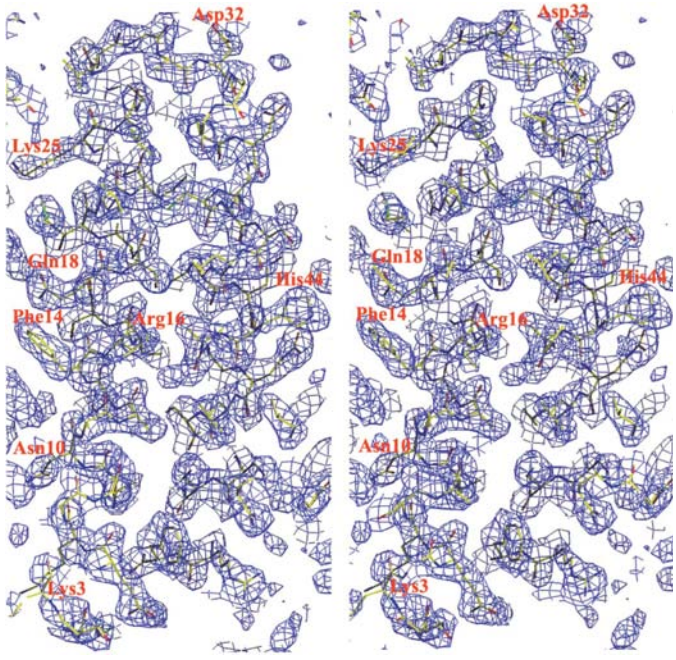

Figure 1
A typical example of the electron-density map of the Rom structure is shown. Stereoviews of $2F_{\rm o} - F_{\rm c}$ electron-density map of the disordered residues (red) contoured at 1σ.

Table 2
Refinements in various resolution ranges.

Resolution (Å)	R_{work} (%)	R_{free} (%)	*B* factors (Å^2)	
			Protein	Solvent
20–2.5	17.9	26.3	9.9	22.6
20–2.6	17.4	26.6	10.0	21.4
20–2.7	17.0	25.9	9.1	23.8
20–2.8	16.6	26.0	8.6	21.4
20–2.9	16.2	25.8	8.4	20.3

differed in different protein preparations and caused doubling of several peaks. Aside from the two Rom species, purity was judged to be better than 95% for all preparations. The protein concentrations were determined by UV spectroscopy using an extinction coefficient of 1716 M^{-1} cm^{-1} at 280 nm.

2.2. Synthesis and purification of Tar–Tar* RNA

All RNA molecules were synthesized *in vitro* using T7 RNA polymerase and a synthetic DNA template (Milligan & Uhlenbeck, 1989). The RNA species were purified using 20% denaturing polyacrylamide gel electrophoresis. RNA extinction coefficients at wavelengths of 260 and 280 nm were calculated using the program *Extinction* 92, which uses nearest-neighbor parameters (Puglisi & Tinoco, 1989). To form the Tar–Tar* kissing hairpin complex, each hairpin was first annealed separately by heating to 363 K followed by fast cooling in ice. They were combined afterwards in a 1:1 ratio. The formation of a bimolecular complex is evidenced by the strong concentration dependence of the melting temperature. Each hairpin alone and the kissing hairpin complex are all easily characterized by one-dimensional imino spectra. Monitoring the imino NMR resonances is an indirect way to confirm that the two hairpins have been combined in the proper ratio. The presence of the kissing hairpin complexes was further confirmed using the sum of extinction coefficients for the individual hairpins as a first approximation to their extinction coefficients. Hypochromicity arising from the loop–loop interaction was determined to affect the measurements by less than 2%.

2.3. UV melting

Optical absorbance melting studies were performed on a Gilford 250 spectrometer using a heating rate of 0.5 K min^{-1}. The RNA was dissolved in 10 m*M* sodium phosphate buffer pH 6.3. Optical absorbance melting studies of Rom binding to the kissing complex were carried out using RNA and protein homodimer concentrations of 10 µ*M* combined in a 1:1 ratio in sodium phosphate buffer plus various salt concentrations.

3. Protein–RNA titrations

Aliquots of the kissing hairpins were prepared by lyophilizing the appropriate volume of stock solution containing the RNA complex in 10 m*M* sodium phosphate pH 6.3. Titration of Rom with Tar–Tar* was performed by dissolving each aliquot of dry RNA with the protein solution in buffer. In this way the volume, sodium chloride or magnesium chloride and DTT

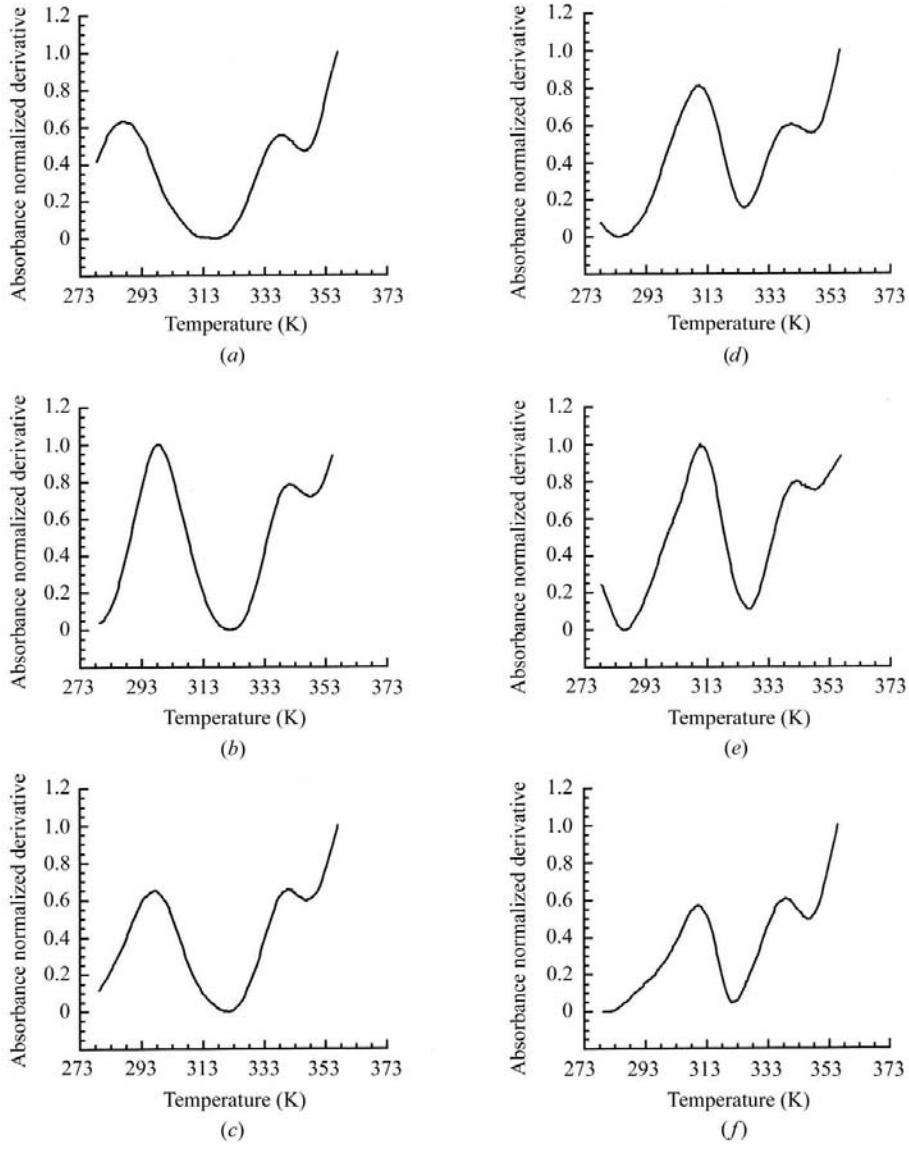

Figure 2
Tar–Tar* UV melting curves in the presence and absence of Rom. The experiments were performed in 10 m*M* sodium phosphate pH 6.3 and 0.1 m*M* EDTA for the experiments without magnesium. All concentrations of RNA and Rom are 10 m*M*. (*a*) Tar–Tar* in 50 m*M* NaCl, (*b*) Tar–Tar* in 150 m*M* NaCl and (*c*) Tar–Tar* in 150 µ*M* $MgCl_2$. (*d*), (*e*) and (*f*) are the same as (*a*), (*b*) and (*c*), respectively, in the presence of a 1:1 ratio of Rom.

concentration of the NMR sample remained constant throughout the experiment, while the concentration of sodium phosphate varied from 10 to 30 m*M*. Control experiments showed that variations of sodium phosphate concentration within this range produce no changes in the spectra. Formation of the kissing hairpin complex was corroborated by comparison of one-dimensional imino spectra with data for the free RNA. Similarly, titration of Tar–Tar* with Rom was performed by dissolving lyophilized aliquots of the protein in sodium phosphate buffer and 10 m*M* sodium chloride with a solution containing Tar–Tar* dissolved in the final buffer. Typical concentrations of the complexes at a 1:1 Rom homodimer to Tar–Tar* stoichiometry ranged from 1 to 3 m*M*. The correct 1:1 protein:RNA ratio was determined by UV absorbance melting.

Table 3
Root-mean-square differences between Rom and Rop (1rop).

Overall differences (all atoms) and the largest backbone differences except for the N- and C-termini are given in angstroms.

Protein chain	Rom chain *B*	Rop (monomer)	Rop (dimer)
Rom chain *A*			
Overall r.m.s.d.	0.9	0.7	
Max. backbone r.m.s.d.s	Thr7*A*–Thr7*B*, 0.9	Glu24*A*–Glu24C, 0.9	
	Phe14*A*–Phe14*B*, 0.8	Lys25*A*–Lys25*C*, 0.9	
	Glu24*A*–Glu24*B*, 0.8	Leu29*A*–Leu29*C*, 0.9	
		Glu33*A*–Glu33*C*, 1.0	
Rom chain *B*			
Overall r.m.s.d.		1.0	
Max. backbone r.m.s.d.s		Leu29*B*–Leu29*D*, 0.9	
		Gly30*B*–Asp30*D*, 0.9	
		Ala31*B*–Ala31*D*, 0.9	
		Asp32*B*–Asp32*D*, 0.8	
		Asp43*B*–Asp43*D*, 0.9	
Rom dimer (*AB*)			
Overall r.m.s.d.			1.0
Max. backbone r.m.s.d.s			Glu24*A*–Glu24*C*, 1.0
			Glu33*A*–Glu33*C*, 0.9
			Asp36*A*–Asp36*C*, 1.0
			Phe14*B*–Phe14*D*, 0.8
			His42*B*–His42*D*, 0.8

3.1. Crystallization and data collection

The Rom protein (0.75 m*M* dimer in 20 m*M* Tris–HCl pH 7, 1 m*M* EDTA, 50 m*M* NaCl) was mixed with 0.75 m*M* Tar–Tar* kissing hairpin [3′-CUCGG-(Tar:AGGGUC)-UGUCG-5′, 5′-GAG-GAGCC-(Tar*:UCCCAG)-ACAGC-3′] in 0.1 m*M* EDTA. This solution produced rod-shaped crystals in hanging drops when mixed with a precipitant containing 30% PEG MME 2000, 0.1 *M* sodium acetate pH 4.6 and 0.2 *M* ammonium sulfate and equilibrated at 295 K against the same solution. Under these conditions, the crystals appeared after about 3–7 d and the rods grew to maximum dimensions of 360 × 50 µm within four weeks. Diffraction data of the protein were collected from a flash-cooled crystal at 100 K. X-ray data were collected on a Rigaku R-AXIS IIC

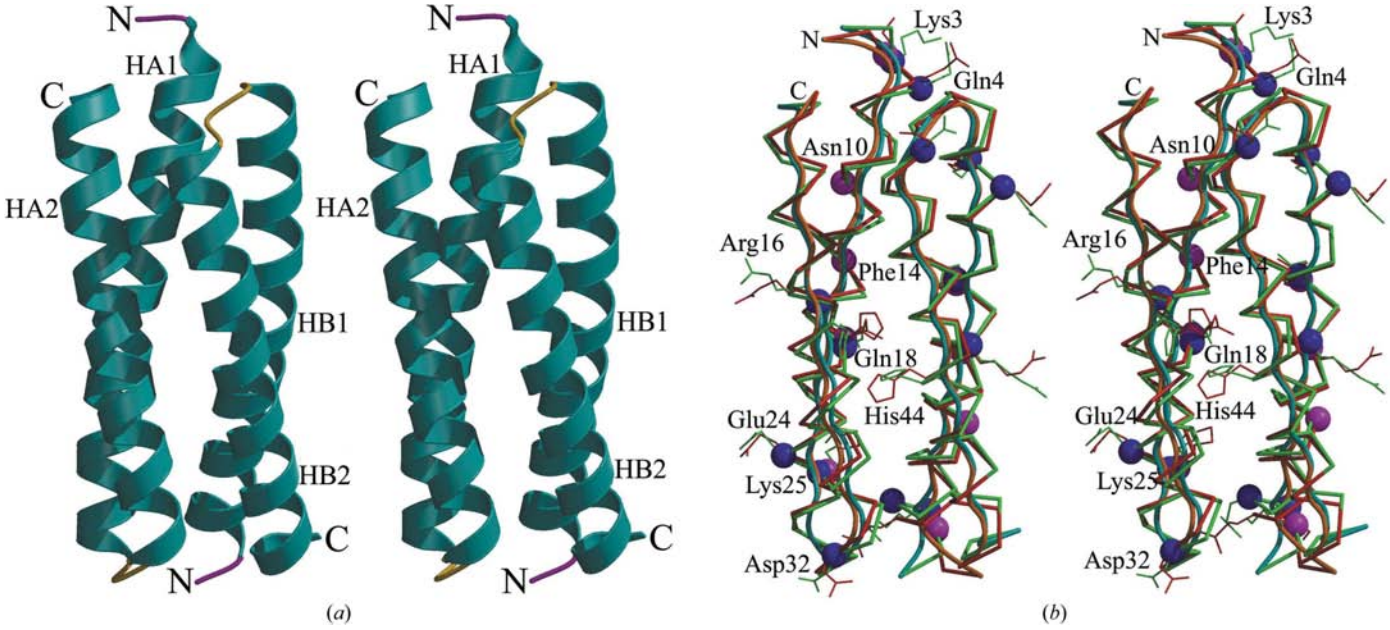

Figure 3
Overall structure of the Rom protein. (*a*) Backbones are shown as ribbons, with α-helices in cyan and loops in orange. The Rom protein contains one dimer per asymmetric unit. Helices HA1 and HA2 of the first monomer and HB1 and HB2 of the second are indicated, as are the N- and C- termini of each chain. (*b*) Structural comparisons of Rom and Rop (1rop). Superpositions of the C^{α} traces are shown in green (Rom) and red (Rop) and coil traces in cyan (Rom) and yellow (Rop) (residues *A*1–*A*56 and *B*1–*B*57) (green) on Rom (residues *A*1–*A*57 and *B*1–*B*57). The RNA-binding residues are indicated by blue (Rom) and magenta (Rop) circles and labeled. The Rom structure is similar to Rop (1rop) other than the loops, C- and N- termini and some of the side-chain conformations.

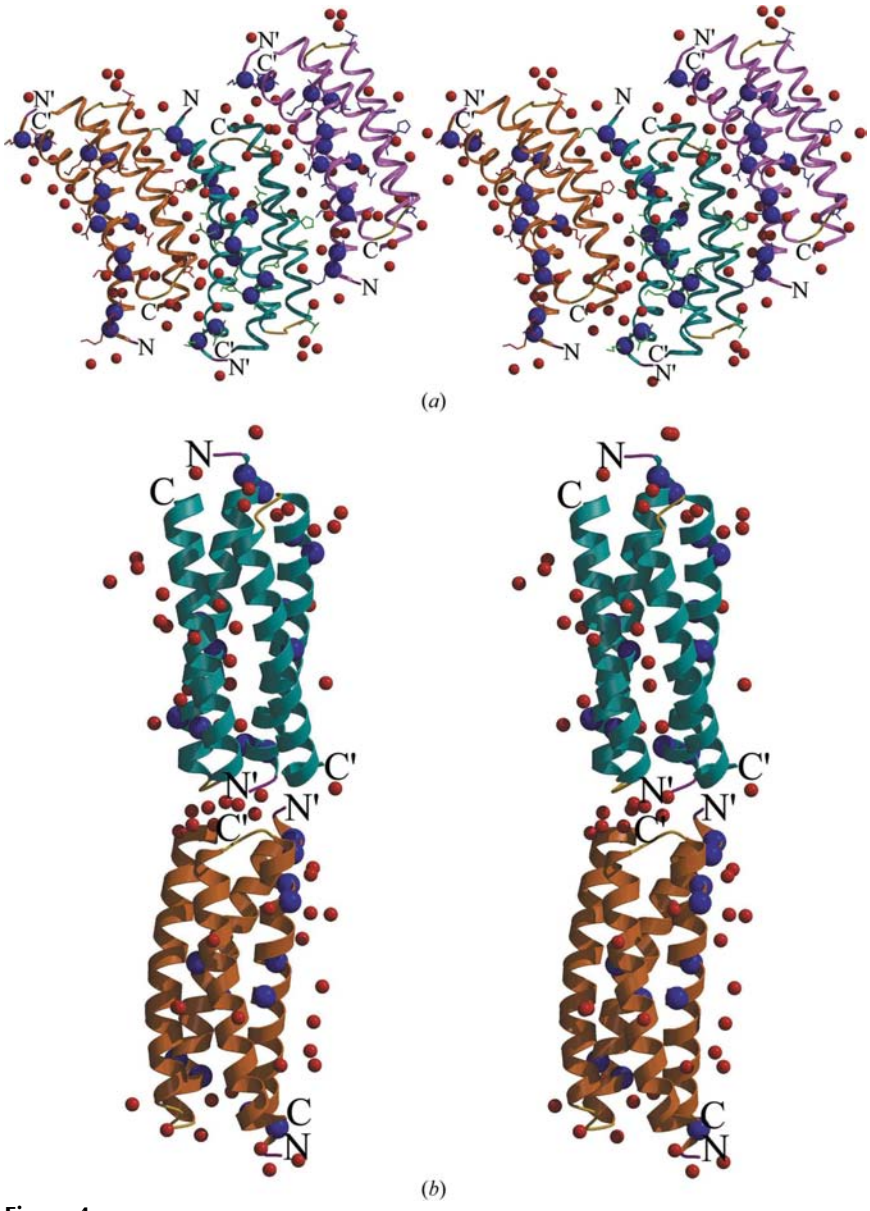

Figure 4
Stereoviews of crystal-packing interactions. (*a*) Water molecules are shown as red spheres. The residues implicated in RNA binding are shown as blue spheres. Most water molecules are located around the RNA-binding sites. (*b*) The N- and C-termini are shown to be involved in head-to-head packing.

image-plate system using Cu $K\alpha$ radiation (λ = 1.5418 Å) and Φ scans with a scan width of 2.0°. Diffraction data (Table 1) were obtained and processed using the programs *DENZO* and *SCALEPACK* (Otwinowski, 1993). The crystals belong to the monoclinic space group *C*2 (unit-cell parameters a = 97.8, b = 39.1, c = 26.7 Å, β = 96.2°). A total of 17 623 reflections were measured, which reduced to 3577 unique reflections, representing a 97.5% complete data set. $R_{\rm merge}$ for this data was 11.7% (96.2% completeness in the highest resolution shell). The complete data statistics are summarized in Table 1.

3.2. Structure determination

The structure was determined by the molecular-replacement method using the *EPMR* program (Kissinger *et al.*, 1999). The coordinates of the crystal structure of Rop Asp30Gly were obtained from the Protein Data Bank (PDB code 1gto) and used as the search model (Predki *et al.*, 1996). We used the dimer model (chains *B* and *C*) without any modification. The search was carried out with data between 10.0 and 4.0 Å and the best solution had a correlation coefficient of 56.1% and an R factor of 47.8%.

3.3. Model refinement

During refinement, bulk-solvent correction allowed the inclusion of all low-resolution reflections and ∼10% of the data were randomly selected for cross-validation (Brünger, 1992*a*). Initial positional refinement of the rigid-body refined model using 12.0–2.5 Å data followed by group *B*-factor refinement lowered $R_{\rm free}$ to 39.4%. The model was improved by iterative model building using the program *O* (Jones *et al.*, 1991). The refinement was continued using *X-PLOR* (Brünger, 1992*b*) with an annealing temperature of 4000 K. Restraints were placed on bond lengths, bond angles, non-bonded contacts and temperature factors of neighboring atoms. $2F_{\rm o} - F_{\rm c}$ maps as well as omit maps were calculated at regular intervals to allow manual rebuilding of side chains with different rotamers from the search model. Solvent molecules (all regarded as water) were added conservatively with due regard to their environment. The standard crystallographic R factor for data between 20 and 2.5 Å was 17.9% and the $R_{\rm free}$ value was 26.3% (in Table 1). The disordered residues are clearly visible and well resolved in the electron-density maps. An example of the electron-density map is shown in Fig. 1. The model exhibits good geometry, with root-mean-square deviations (r.m.s.d.s) from ideal bond lengths and angles of 0.008 Å and 2.15°, respectively. The Ramachandran plots indicate that all amino-acid residues are in the generously allowed regions except residue Lys3*A* (packing interface region; Cowtan & Main, 1993). The structural representations in Figs. 1 and 3–7 were generated with *MOLSCRIPT* (Kraulis, 1991), *GRASP* (Nichols & Yanofsky, 1979) and *RASTER3D* (Merritt & Murphy, 1994).

4. Results and discussion

Fig. 2 shows the stabilizing effect of sodium chloride and magnesium chloride on the kissing hairpin complex in the absence and presence of Rom protein. The addition of Rom in

a 1:1 ratio significantly increases the melting temperature of the complex. Changing the sodium chloride concentration from 50 to 150 m*M* does not change the melting temperature in the presence of Rom (Figs. 2*d* and 2*e*). The effect of magnesium is approximately the same as sodium, but at a thousand-fold lower concentration (Fig. 2*f*). The complex was expected to be stabilized under the crystallization conditions (0.1 *M* sodium acetate pH 4.6 and 0.2 *M* ammonium sulfate).

The final structure, refined against 2.5 Å resolution data, consists of 114 residues and 33 water molecules. The asymmetric unit of the crystal contains a dimer with a twofold noncrystallographic axis of symmetry. A four-helix bundle is formed by the association of two identical helix–turn–helix monomers, each composed of 57 amino acids (Fig. 3*a*). Association of the amphiphilic α-helices appears to be stabilized primarily through hydrophobic interactions, with little contribution from electrostatic interactions. The six amino acids at the C-terminus of Rop, which extend beyond the end of helices A2 and B2, have been shown to be unstructured in both X-ray crystallographic and nuclear magnetic resonance studies. In the Rom structure described here, the residues Gly57 and Gly57′ are observed in the HA2 and HB2 helices. The remaining five amino acids could not be seen in the electron-density map and were presumably disordered as reported in the previously determined structures. The average temperature factor for the main-chain atoms is 9.9 Å^2 and for the water molecules is 22.6 Å^2. The *B*-factor plot of the main-chain atoms (not shown) clearly indicates the flexibility of the loop region as well as the N- and C-terminal tails of the protein. The N-terminus of the subunits is very flexible even though the N- and C-termini are located close to the N- and C-terminal ends of an adjacent subunit (head-to-head packing interactions; Figs. 4*a* and 4*b*).

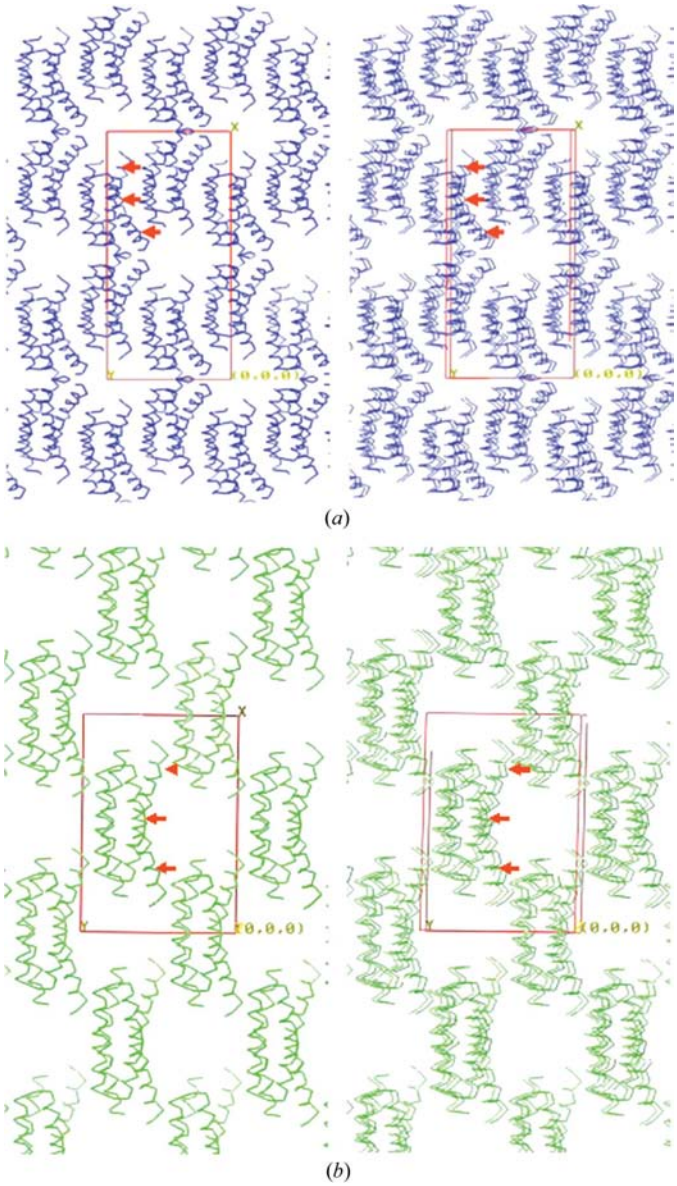

Figure 5
Stereoviews of the layers of Rom and Rop (1rop). (*a*) Strong crystal-packing interactions and low solvation are shown in the Rom structure. (*b*) Relatively weak crystal-packing interactions are shown in the Rop structure. The regions of potential RNA binding are represented in the figure as serving as an opening (Rop) and closing (Rom) gateway. Putative RNA-binding sites are indicated by red arrows.

Table 4
Selected distances of packing interactions (Rom and Rop).

Charged residues are shown in bold. An asterisk indicates a residue implicated in RNA binding. S indicates a solvent (water) molecule.

Interaction	Distance (Å)
Rom	
α-Helix	
***Lys3*A* N$^{\zeta}$—Asn27*A* O$^{\delta1}$**	**3.0**
***Lys3*A* N$^{\zeta}$—S5 O**	**3.0**
S5 O—Glu39*A* O$^{\varepsilon2}$	**2.9**
S5 O—Glu39*B* O$^{\varepsilon1}$	**2.8**
Gln4*A* N$^{\varepsilon2}$—Thr2*A* O	2.8
***Asn10*A* O$^{\delta1}$—Arg13*A* N$^{\gamma1}$**	**2.9**
***Asn10*A* N$^{\delta1}$—Asp43*A* O$^{\delta1}$**	**3.0**
Arg13*A* N$^{\gamma2}$—Tyr49 O$^{\gamma}$	**3.4**
Arg13*A* O—S15 O	**2.7**
S15 O—Asn27*B* O$^{\delta1}$	3.1
Ser17*A* O$^{\gamma}$—Arg13*A* O	3.3
Ser17*A* O$^{\gamma}$— S15 O	3.0
***Gln18*A* N$^{\varepsilon2}$—Glu33*B* O$^{\varepsilon2}$**	**3.1**
Gln18*A* O$^{\varepsilon1}$—S3 O	3.0
S3 O—Ser51*A* O$^{\gamma}$	2.6
Glu24*A* O$^{\varepsilon2}$—His4*B* N$^{\delta1}$	**3.0**
Glu24*A* O$^{\varepsilon2}$—Glu39*B* O$^{\varepsilon1}$	**3.1**
***Lys25*A* N$^{\zeta}$—S28 O**	**3.3**
S28 O—Glu24*B* O$^{\varepsilon2}$	**3.2**
Loop (Å)	
Asn27*A* O$^{\delta1}$—Lys3*A* N$^{\zeta}$	**3.0**
Glu28*A* O—Arg50*B* N$^{\gamma2}$	**3.2**
Glu28*A* O$^{\varepsilon2}$—His42*B* N$^{\varepsilon2}$	**2.9**
Gly30*A* O—Arg55*B* N$^{\gamma1}$	**2.8**
Asp32*A* O$^{\delta2}$—S25 O	**3.3**
S25 O—Arg55*B* N$^{\gamma1}$	**3.1**
Glu33*A* OE2—S9 O	**3.2**
S22 O—Arg55*B* N$^{\gamma1}$	**3.3**
Gln34*A* N$^{\varepsilon1}$—Arg55*B* N$^{\gamma1}$	**2.8**
Gln34*A* N$^{\varepsilon2}$—Ser51*B* O	2.9
Gln34*A* N$^{\varepsilon2}$—Ser51*B* O$^{\gamma}$	2.8
Rop (1rop)	
N-terminus	
***Lys3*A* N—Asp36*B* O$^{\delta1}$**	**2.9**
Gln4*A* N—Asp36*B* O$^{\delta2}$	**3.0**
Gln4*A* N$^{\varepsilon2}$—Asp36*B* O$^{\delta2}$	**2.8**
Thr7*A* O$^{\gamma}$—Asp32*B* O$^{\delta2}$	**3.0**

The final model has been refined to 2.5 Å resolution to a current R_{work} of 17.9% and R_{free} of 26.3%; R_{free} was 8.4% higher than R_{work}. These results suggest that the difference between R_{work} and R_{free} is relevant to the high R_{merge} value. The model has been refined to the resolution range of the data (Table 2). With one dimer in the asymmetric unit, non-crystallographic symmetry (NCS) could be exploited during refinement by restraining chemically identical chains of the individual Rom protein dimers.

Structural comparisons (Table 3) indicate that Rom chain *A* and the starting model (Rop subunit; 1rop) are very similar overall, with a root-mean-square deviation of 0.7 Å for the main-chain atoms of all residues, while the average r.m.s.d. between Rom chain *B* and the Rop subunit is significantly higher (1.0 Å). This difference in chain conformations is also reflected in the 0.9 Å r.m.s.d. between chains *A* and *B*. Thus, there is a greater difference between chains *A* and *B* than between chain *A* and the Rop subunit. Superposition of Rom and the Rop dimer also indicated that there are (1.0 Å r.m.s.d. overall) conformational changes induced by the different crystallization conditions and packing.

From the superposition of the two structures, the residues with backbone r.m.s.d. values exceeding 0.8 Å (Table 3) are located primarily in the HA1 and loop regions (residues 14, 24, 25, 29–33) and in the N- and C-terminal regions (residues 1, 2, 54, 55, 56 and 57). Exceptions are His42*B* and Asp43*B*, with r.m.s.d.s of 0.8 and 1.0 Å, respectively. In the Rom crystal, the loops between helices HA1 and HA2 are exposed to the solvent and flexible, but the electron density of the loop is clearly visible. The differences in the structures and the packing of Rom and Rop (Table 3; Figs. 5*a* and 5*b*) suggest that the presence of the polyelectrolyte kissing hairpin RNA in the crystallization solution can induce small changes in the backbone and side-chain conformations and allow tighter packing by changing the ionic strength of the mother liquor. We attempted to crystallize the Rom protein without RNA and obtained only precipitate using identical crystallization conditions to those used to successfully grow Rom crystals in the presence of the kissing hairpin RNA.

The structure of Rop (1rop) was originally determined at 1.7 Å resolution (Banner *et al.*, 1987). The helices pack antiparallel to their neighbors to form a four-helix coiled-coil bundle. In Rop, the hydrophilic residues that are involved in RNA binding (Lys3, Asn10, Gln18 and Lys25) surround Phe14 and

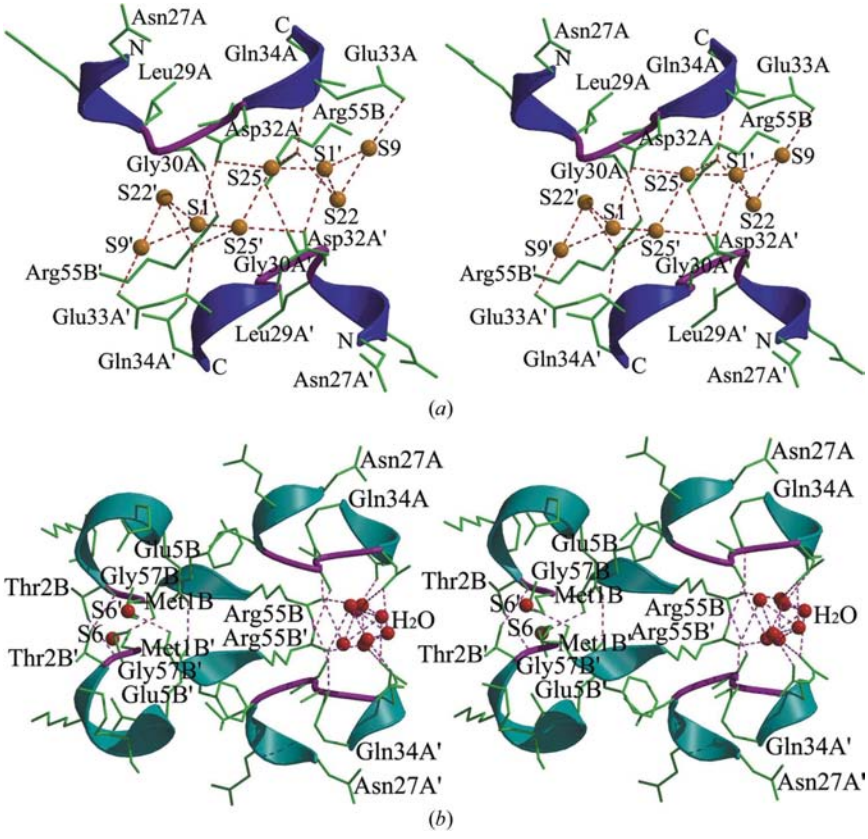

Figure 6
Detailed views of packing between the loop regions of Rom. (*a*) The residues and solvent (water) molecules in the hydrogen-bonding network are labeled. The charged hydrophilic residues of each loop interact through solvent between chains. (*b*) Close-up view of the linker regions of the dimers. The amino acids of the C-terminus interact with those of the N-terminus through hydrogen bonding with the water molecules (solvents S6 and S6′).

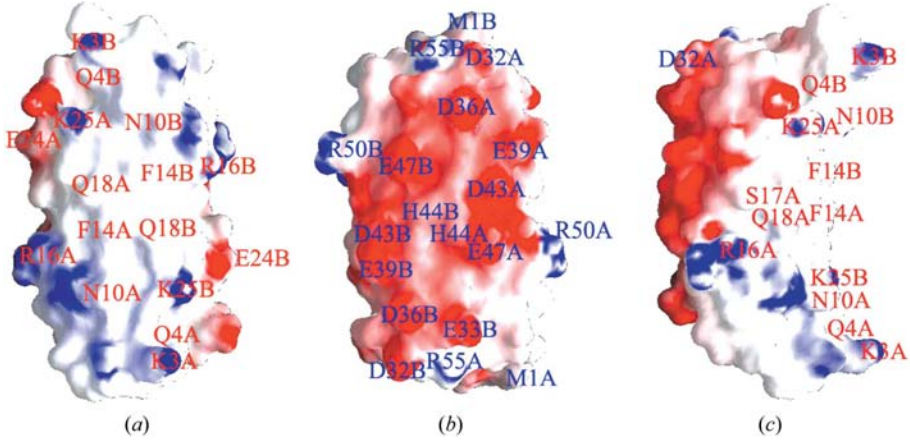

Figure 7
Surface representation of the Rom homodimer. Three views are shown: (*a*) the face of the RNA-binding sites, (*b*) the dimer interface and (*c*) the monomer interface. The relative distribution of surface charge is shown with acidic regions in red, basic regions in blue and neutral regions in white.

are likely to make contacts to the ribose-phosphate backbone. Here, in the Rom structure, we have identified differences in the side chains of a number of residues (Lys3, Gln4, Arg16, Gln18, Glu24, Lys25, Asp32 and His44) compared with the Rop structure. We also superimposed 1gto on the Rom structure and found the same side-chain mobility/RNA-contact residues as the results of the 1rop structure. The side chains of Rom that have large deviations from 1gto are Lys3, Gln4, Asn10, Phe14, Arg16, Gln18, Glu24, Lys25, Asp32, Glu33 and His44. Thus, the flexible side chains could be interaction motifs with RNA. The RNA-binding site is conserved except for the residues Gln4, Arg16 and Glu24. Glu24 is located in helix HA1, which is a part of the RNA-binding site. As shown in Table 4, many of the packing interactions involve charged residues and the putative binding residues in the α-helix and loop, forming polar interactions with the side chains of amino acids or water molecules.

The amino acids Lys3, Gln4, Arg16, Gln18, Glu24 and Lys25 are highly conserved (Fig. 3*b*) in Rop structures. These residues are part of HA1 and HB1, with Lys3 and Gln4 sitting atop the loop between H1 and H2. Three polar amino acids on this segment, Lys3, Gln18 and Lys25, are involved in crystal-packing interactions, but would be solvent-exposed in solution. The additional amino acids potentially involved in RNA binding, Gln4, Arg16 and Gln24, are accessible.

In our crystal structure, the Rom protein tightly packs with a symmetry-related Rom molecule. Even though the packing is tighter than the previously known Rop structure, the structure is less ordered than Rop, resulting in a lower resolution of the diffraction data. The RNA-binding residues of the Rom protein are actually less accessible in the Rom crystal than that of the previous Rop structure. We speculate that the presence of the kissing hairpin in the crystallization solution may promote its affinity for the symmetry-related Rom protein (Figs. 6*a* and 6*b*) by increasing the ionic strength of the solution, allowing tighter packing.

The Rom crystal structure presents compact packing of a homodimer, making a contact surface greater than previously known structures (Fig. 5*a*). The RNA-binding sites are predicted in Fig. 3(*b*). The binding site is an exposed surface delineated by HA1 and HB1 that we propose to be the major contact area for protein–RNA interaction (Fig. 7). The side chains of residues on chain HA1 (residues 3–26), HB1 (residues 3–26) and the loop (residues 27–34) are involved in many polar and charged amino acids that accommodate the side chains of a symmetry-related molecule and hydrating waters. In contrast, the contact zones in Rop appear to cover a smaller surface (Table 3 and Fig. 5*b*). Given that the diffraction properties are associated with this form, the stability of the packing can be judged in terms of surface dimensions. The helix forms an intermolecular hydrophobic core together with the protein C-terminus and stabilizes the intersubunit interface, making all conserved hydrogen bonds between subunits inaccessible to solvent (Banner *et al.*, 1987). In this structure, the tops of two adjacent subunits are brought together around the solvent-inaccessible hydrogen bond formed by His44 ND2 and Val69 O. Since the dimeric Rop motif is the conserved structure, the degree of compaction observed in the crystal packing could be related to variations in side-chain interactions not involved in the hydrogen-bonding network of the α-helices.

5. Conclusions

In the present study, we report the crystal structure of Rom with a novel structural packing. Strong packing interactions with low solvent content and accessibility are evident in the crystal structure. An optimized stabilization of the dimer structure can be assessed by the formation of highly ordered crystal packing. Electron-density maps show altered side-chain conformations for the residues of Rom involved in RNA binding. We propose that conformational changes of side chains involved in RNA recognition are coupled with formation of this new crystal form. The difference in conformation of the side chains involved in RNA binding may be a function of their inherent flexibility in the absence of kissing hairpin RNA.

This appears to be an unusual example of crystallization: changes in crystal packing usually arise either from variations of the crystal-growth parameters (temperature, pH, precipitant salts or crystallizing reagents) or from important conformational changes in modular proteins in response to RNA binding. This packing modification may be seen as a response to addition of the RNA polyanions in the crystallization solution, promoting a tighter protein–protein association in initiation and propagation of crystal growth. In our case, the active site at the dimeric surface is connected to a tunnel built by the arrangement of all subunits. The conformational flexibility and variability of the side chains involved in RNA binding may explain the variety of Rop crystals observed. In summary, the results of this structural study provide additional information toward a detailed picture of Rom protein–RNA recognition and insights into protein crystallization in general.

This work was supported by NIH grant GM 4921501 to SRH. Facilities and equipment were provided through support of the Office of Energy Research, Office of Health and Environmental Research, Health Effects Research Division of the US Department of Energy. This work was also supported by a Korea Research Foundation Grant (KRF-2005-041-E00510) to SBJ. Melting studies were performed by LRC in the laboratory of Professor Ignacio Tinoco at UC Berkeley.

References

Banner, D. W., Kokkinidis, M. & Tsernoglou, D. (1987). *J. Mol. Biol.* **196**, 657–675.

Brünger, A. T. (1992*a*). *Nature (London)*, **355**, 472–475.

Brünger, A. T. (1992*b*). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*. New Haven, CT, USA: Yale University Press.

Castagnoli, L., Scarpa, M., Kokkinidis, M., Banner, D. W., Tsernoglou, D. & Cesareni, G. (1989). *EMBO J.* **8**, 621–629.

Castagnoli, L., Vetriani, C. & Cesareni, G. (1994). *J. Mol. Biol.* **237**, 378–387.

Chou, K.-C., Maggiora, G. M. & Scheraga, H. A. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 7315–7319.
Comolli, L. R., Pelton, J. G. & Tinoco, I. Jr (1998). *Nucleic Acids Res.* **26**, 4688–4695.
Cowtan, K. D. & Main, P. (1993). *Acta Cryst.* D**49**, 148–157.
Eberle, W., Klaus, W., Cesareni, G., Sander, C. & Rosch, P. (1990). *Biochemistry*, **29**, 7402–7407.
Eberle, W., Pastore, A., Sander, C. & Rosch, P. (1991). *J. Biomol. NMR*, **1**, 71–82.
Eguchi, Y. & Tomizawa, J. (1990). *Cell*, **60**, 199–209.
Eguchi, Y. & Tomizawa, J. (1991). *J. Mol. Biol.* **220**, 831–842.
Glykos, N. M., Cesareni, G. & Kokkinidis, M. (1999). *Structure Fold. Des.* **7**, 597–603.
Itoh, T. & Tomizawa, J. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 2450–2454.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
Klosterman, P. S., Shah, S. A. & Steitz, T. A. (1999). *Biochemistry*, **38**, 14784–14792.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.
Milligan, J. F. & Uhlenbeck, O. C. (1989). *Methods Enzymol.* **180**, 51–62.
Nagi, A. D. & Regan, L. (1997). *Biol. Chem.* **378**, 1141–1152.
Nichols, B. P. & Yanofsky, C. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 5244–5248.
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.
Papanikolau, Y., Kotsifaki, D., Fadouloglou, V. E., Gazi, A. D., Glykos, N. M., Cesareni, G. & Kokkinidis, M. (2004). *Acta Cryst.* D**60**, 1334–1337.
Polisky, B. (1988). *Cell*, **55**, 929–932.
Predki, P. F., Agrawal, V., Brünger, A. T. & Regan, L. (1996). *Nature Struct. Biol.* **3**, 54–58.
Predki, P. F., Nayak, L. M., Gottlieb, M. B. C. & Regan, L. (1995). *Cell*, **80**, 41–50.
Puglisi, J. D. & Tinoco, I. (1989). *Methods Enzymol.* **180**, 304–325.
Spyridaki, A., Glykos, N. M., Kotsifaki, D., Fadouloglou, V. E. & Kokkinidis, M. (2000). *Acta Cryst.* D**56**, 1015–1016.
Tomizawa, J. (1984). *Cell*, **38**, 861–870.
Tomizawa, J. & Itoh, T. (1981). *Cell*, **31**, 575–583.
Tomizawa, J. & Som, T. (1984). *Cell*, **38**, 871–878.
Vlassi, M., Steif, C., Weber, P., Tsernoglou, D., Wilson, K. S., Hinz, H. J. & Kokkinidis, M. (1994). *Nature Struct. Biol.*, **1**, 706–716.