Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of human translin at 2.2 Å resolution

Ikuko Sugiura,[a,b] Chiduko Sasaki,[a,b] Tsukasa Hasegawa,[a,b] Toshiyuki Kohno,[c] Shigetoshi Sugio,[a,b] Hideaki Moriyama,[d] Masataka Kasai[e] and Takao Matsuzaki[a,b]*

[a]ZoeGene, 1000 Kamoshida, Aobaku, Yokohama 227-8502, Japan, [b]Mitsubishi Chemical Corporation, 1000 Kamoshida, Aobaku, Yokohama 227-8502, Japan, [c]Mitsubishi Kagaku Institute of Life Science (MITILS), 11 Minami-ooya, Machida-shi, Tokyo 194-8511, Japan, [d]RIKEN Harima Institute/SPring8, 1-1-1 Kouto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, Japan, and [e]Department of Immunology, National Institute of Infectious Disease, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162, Japan

Correspondence e-mail: tak@rc.m-kagaku.co.jp

Received 17 September 2003
Accepted 1 February 2004

PDB Reference: translin, 1j1j, r1j1jsf.

The structure of human translin at 2.2 Å resolution is reported in space group $C222_1$. Translin forms a tetramer in the asymmetric unit. Although the monomer structure is almost the same as the crystal structure of murine translin in space group $P2_12_12$, the relative positions of the tetramers differ between the human and murine translins. This suggests that the multimerization of translin is flexible; the flexibility may be related to the binding to DNA/RNA.

1. Introduction

Translin was originally identified as a DNA-binding protein that specifically recognizes the consensus sequences found at the breakpoints of chromosomal translocation in many cases of lymphoid neoplasms (Aoki *et al.*, 1995). It was subsequently demonstrated to also bind to highly conserved Y and H elements in the 3′-untranslated region (UTR) of mRNA, thereby suppressing *in vitro* translation (Han & Hecht, 1995). Translin, which consists of 228 residues, is highly conserved among vertebrates but shows no significant sequence homology to other known proteins, except for translin-associated factor X (TRAX; Aoki *et al.*, 1997). Comparison of the amino-acid sequence of TRAX with that of translin showed 28% identity throughout the two molecules, with 38% identity in the C-terminal regions, suggesting that TRAX is a member of the translin family. Translin and TRAX form a heterodimer, but their function remains unknown. Erdemir and coworkers recently reported DNA-damage-dependent interaction of TRAX with the nuclear matrix protein C1D, an activator of the DNA-dependent protein kinase (DNA-PK) which is essential for the repair of DNA double-strand breaks (DSBs) and V(D)J recombination (Erdemir *et al.*, 2002). Although numerous functions of translin have already been reported with reference to DNA and RNA metabolism, recent results also suggest that translin participates in processes ensuring the segregation of chromosomes and cytokinesis through specific interactions with microtubules of the mitotic spindles (Ishida *et al.*, 2002). Thus, translin is a multifunctional protein and it is intriguing how translin achieves this wide variety of functions.

The crystal structure of murine translin has been solved at 2.6 Å resolution (Pascal *et al.*, 2002) and the electron-microscopic (EM) structure has been revealed (VanLoock *et al.*, 2001). The crystal structure has fourfold symmetry in the tetramer in contrast to the *C*8 symmetric structure. Both of them have a central channel that may be used for binding to the RNA or DNA target sequences.

We have solved the human translin structure at 2.2 Å resolution. This study provides the opportunity for structural

Table 1
Data-collection protocol.

Values in parentheses are for the highest resolution shell.

		SeMet			
	Native 1	Edge	Peak	Remote	Native 2
Beamline	BL24XU	BL12B2			PF6B
Space group	$C222_1$	$C222_1$			$P4_32_12$
Wavelength	0.836	0.979467	0.970312	0.973162	1.00
Unit-cell parameters					
a (Å)	129.424	129.211			97.2
b (Å)	135.272	135.391			97.2
c (Å)	134.345	131.925			283.6
Unique reflections	57689	17718	17699	17685	17357
Completeness (%)	97.3 (97.3)	99.7 (99.7)	99.7 (99.7)	99.6 (99.6)	95
Resolution(Å)	2.2 (2.32–2.20)	3.30 (3.48–3.30)	3.30 (3.48–3.30)	3.30 (3.48–3.30)	3.485
$R_{\rm merge}$ (%)	8.7 (41.4)	13.2 (33.3)	12.8 (32.2)	12.9 (32.2)	13
Multiplicity	5.1 (4.7)	6.7 (6.4)	6.8 (6.5)	6.8 (6.5)	
$I/\sigma(I)$	13.5 (3.5)	14.8 (5.7)	15.2 (5.8)	15.3 (5.9)	

comparison between human and murine translins belonging to different crystal forms and crystallized under different conditions.

2. Material and methods

2.1. Expression and purification

The translin gene was cloned into the bacterial expression vector pQE9 (Qiagen Inc.). The construct was transformed into *Escherichia coli* host strain M15 [pREP4] and cells were fermented in 2×YTA culture as described previously (Kasai *et al.*, 1997). Cells were suspended in 50 m*M* Tris–HCl pH 7.0 containing 150 m*M* NaCl and 5%(*v*/*v*) glycerol and disrupted by sonication for 20 min in the presence of hen egg-white lysozyme. The supernatant was loaded onto a cobalt-based immobilized metal-affinity column previously equilibrated with the cell suspension buffer. The column was washed with the same buffer as described above and eluted with a linear gradient of imidazole (0–200 m*M*). The effluent was dialyzed against 20 m*M* Tris–HCl pH 7.0 containing 100 m*M* NaCl and 5%(*v*/*v*) glycerol and applied onto an anion-exchange column. The column was eluted with a linear gradient of 100–300 m*M* NaCl in dialysis buffer. A fraction containing most of the protein was concentrated by ultrafiltration and applied onto a gel-filtration column equilibrated with 50 m*M* Tris–HCl pH 7.0 containing 300 m*M* NaCl and 5%(*v*/*v*) glycerol. The effluent was collected and concentrated to 10–80 mg ml^{-1} protein by ultrafiltration. Protein concentration was determined by the Bradford method using bovine serum albumin as a standard. The purified protein showed a single major band on SDS–PAGE stained with Coomassie brilliant blue. For the preparation of selenomethionine-labelled translin (SeMet-translin), QuickChange (Strategene) was employed to mutate Leu183 to Met, as translin contains only one methionine residue, which is at its N-terminus. The mutated translin DNA was cloned into the pQE80 vector (Qiagen). The expression plasmid was transformed into auxotrophic (met$^-$) *E. coli* strain D41. The cells were grown in LeMaster media. When the cell density reached an absorbance of 0.578 at 600 nm, 1 m*M* isopropyl-β-D-glucopyranoside was added in order to induce protein expression. Cells were harvested 18 h after induction and processed to obtain purified protein as described above. Mass-spectroscopic analysis of purified SeMet-translin indicated that the selenomethionine residue had been incorporated into the protein.

2.2. Crystallization

The purified protein was diluted to 10 mg ml^{-1} with storage buffer (50 m*M* Tris–HCl pH 7.0, 150 m*M* NaCl, 5% glycerol). Sparse-matrix crystallization screening with Crystal Screen and Crystal Screen II (Hampton Research) then took place using the hanging-drop vapour-diffusion method. Attempts were made to optimize the conditions that yielded large single crystals. The crystals used in data collection were grown from 2.5 *M* sodium formate, 100 m*M* sodium acetate pH 4.5 at 293 K. A 1 µl aliquot of the precipitating solution was added to 1 µl of 10 mg ml^{-1} translin on a cover glass. The cover glass was placed over a well containing 500 µl precipitating solution. The typical dimensions of the crystals were 0.5 × 0.2 × 0.2 mm. The crystallization conditions and crystal size for SeMet-translin are the same as those for the native protein. Both the native and the selenomethionyl crystals belong to the orthorhombic space group $C222_1$, with unit-cell parameters a = 129, b = 135, c = 132 Å. The asymmetric unit contains four molecules of translin.

2.3. Data collection, structure determination and refinement

X-ray diffraction experiments were performed at BL24XU for native data and BL40B2 for MAD data at SPring-8. The crystals were transferred to the precipitating solution containing 30% sucrose as a cryoprotectant and were flash-cooled in a nitrogen-gas stream at 100 K.

Intensity data were collected from a native crystal which diffracted to 2.0 Å resolution and were integrated using *MOSFLM* (Leslie, 1991). The data were scaled and merged with *SCALA* (Collaborative Computational Project, Number 4, 1994). The statistics are summarized in Table 1. The percentage of reflections with $I > 3\sigma(I)$ was 86.5% (39.3% in the highest resolution shell).

Crystals of SeMet-translin have the same form as the native; however, they diffracted to 3.3 Å resolution at three wavelengths. The data were processed and scaled using the *HKL* program suite (Otwinowski & Minor, 1997).

MAD calculations were performed with the program *SOLVE* (Terwilliger & Berendzen, 1999). An automated Patterson search identified eight selenium sites, which were used to calculate initial phases. The phases were then refined

Table 2
Final model-refinement statistics.

Resolution limits(Å)	15–2.2
No. reflections used in refinement	54339
No. reflections used to compute R_{free}	2899
R_{cryst} (R_{free}) (%)	22.9 (26.5)
No. non-H atoms	
Protein	7088
Solvent	414
Rm.s.d.s from ideal values†	
Bond lengths (Å)	0.007
Bond angles (°)	1.128
Improper angles (°)	0.73
Dihedral angles (°)	18.16
Ramachandran plot†	
Most favoured (%)	93.3
Additional allowed (%)	5.5
Generously allowed (%)	1.1
Mean *B* values	
Protein (Å^2)	51.40
Water (Å^2)	57.51

† Calculated using *PROCHECK* (Laskowski *et al.*, 1993).

by solvent flattening and histogram matching using the *CNX* package (Brünger *et al.*, 1998).

The resulting electron density was readily interpretable without the use of non-crystallographic symmetry averaging. The initial atomic model including four copies of the translin monomer was built using the program *O* (Jones *et al.*, 1991). The C-terminal segment from residue 218 to the end was poorly defined in each monomer and was not included in the atomic model. Refinement of the SeMet-translin model was carried out with the program *CNX*. The atomic model was subjected to rigid-body refinement followed by a round of simulated-annealing, positional refinement and overall *B*-factor refinement against the SeMet data in the resolution range 15.0–3.3 Å. At this point, the native data set at 2.2 Å resolution was employed instead of the SeMet data set and was used throughout the subsequent refinement process. An individual *B* factor was applied to each atom after several rounds of refinement. At this stage, solvent molecules were clearly visible and were incorporated into the model. The refinement procedure using the MLHL target function and manual rebuilding continued until no further decrease in the free *R* factor was observed. The final atomic model, consisting of four monomers with residues 1–217 and 414 water molecules, gives *R* and R_{free} factors of 0.229 and 0.265, respectively, against all reflections in the resolution range 15.0–2.2 Å. The final refinement statistics are summarized in Table 2. Figures were prepared with *MOLSCRIPT* (Kraulis, 1991), *XTALVIEW* (McRee, 1999) and *Raster3D* (Merritt & Bacon, 1997).

2.4. Molecular replacement with another crystal form

We have reported another translin crystal form with a tetragonal crystal lattice from different crystallization conditions: 100 m*M* MES pH 6.5, 20% PEG 4000 at 293 K (Kasai *et al.*, 1997). Despite the poor diffraction to 3.5 Å resolution, intensity data from this crystal form were measured at BL6A of the Photon Factory (Table 1) and used for structure analysis by molecular replacement. Initial phases were determined with the program *EPMR* (Kissinger *et al.*, 1999) using two molecules of the translin monomer. An enantiomorphic ambiguity in the space group of the crystal, *i.e.* $P4_32_12$, was solved during the molecular-replacement work. Refinement and interactive model building were performed using *CNX* and *O*, respectively. The final *R* and R_{free} factors against data in the resolution range 15–3.485 Å are 0.25 and 0.45, respectively. Because the R_{free} could not be decreased to below 0.45, the tetragonal structure will not be used for further discussion.

2.5. Quality of the crystallographic model and atomic coordinates

In the orthorhombic crystal structure, the first 11 residues, including the His tag and the spacer sequence before Met1, and the last 11 resides (residues 218–228) are invisible, probably owing to disorder. The remaining 93% of the residues in the atomic model are in the core region of a Ramachandran plot (*PROCHECK*; Laskowski *et al.*, 1993). Only nine residues are located in the generously allowed region. Seven of them are in the regions 129–133 in the four monomers, which form loop regions with poor electron density. The other residues are Ala49 and Asp149 in molecule *C*, both of which are at the end of helices.

3. Results and discussion

The final model consists of 7088 protein atoms from the tetramer and 414 water molecules in the asymmetric unit. These atoms occupy about 50% of the unit-cell volume. The $2F_o - F_c$ electron-density map is continuous and well defined for both backbones and side chains, with the exception of the loop region 139–141 in each monomer. Fig. 1 shows segment 162–166 with its corresponding electron density.

The translin monomer consists of seven α-helices, six of which run nearly parallel to each other (Fig. 2*a*). Four translin monomers are related by a non-crystallographic fourfold axis in the asymmetric unit (Fig. 2*b*). The two tetramers form an octamer by crystallographic twofold symmetry (Fig. 2*c*). The overall octameric structure is approximately 90 × 90 × 110 Å

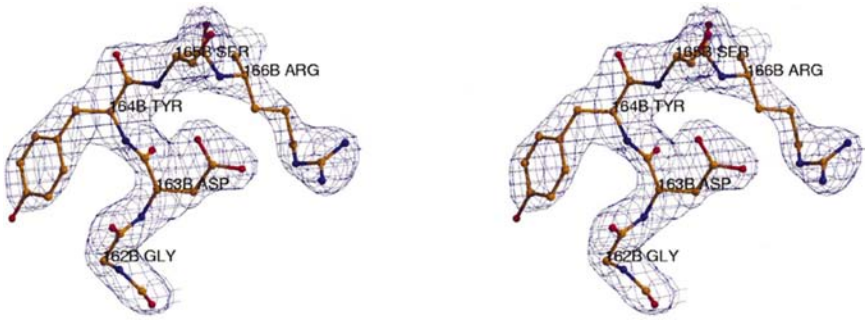

Figure 1
$2F_o - F_c$ electron-density map contoured at 1.5σ of segment 162–166 in molecule *B*.

in size. The central channel of the tetramer ring is 15 Å in diameter at its largest opening and 4.4 Å at the narrowest part.

There is no distinct dimeric structure seen in the translin crystal structure, although it has been reported that the translin dimer binds to DNA and that the smallest functional unit of translin is a dimer (Wu *et al.*, 1998). Because the dimerization of translin requires a disulfide bond between the two Cys225 residues of neighbouring molecules (Wu *et al.*, 1998), two molecules that are side-by-side, not positioned up-and-down as in Figs. 2(*b*) and 2(*c*), are needed to form a dimer. Unfortunately, however, every Cys225 cannot be seen in the crystal structure. The amino-acid sequence of translin contains a hypothetical leucine-zipper motif from residues 177 to 212, which has been also considered to contribute to dimerization; it is located in two α-helices, α6 and α7, and does not in fact form a leucine zipper. There is unlikely to be a leucine-zipper interaction between the monomers.

The translin monomers have almost the same structure and all comparisons of the monomers that make up the tetramer show C^{α} r.m.s. differences of only 0.67–1.21 Å (Fig. 3*a*); the differences are mainly confined to the loop regions 43–53 and 124–135. All comparisons between human and murine translin monomers, which have 99% sequence similarity to each other, also give small C^{α} r.m.s. differences of between 0.32 and 1.21 Å.

However, this resemblance does not apply to the tetramers; two combinations of the human tetramers such as molecules *ABCD versus BCDA* (identical to *ABCD versus DABC*) and *ABCD versus CDAB* show C^{α} r.m.s. differences of 2.4 and 2.3 Å, which means the local fourfold symmetry is not accurately a fourfold symmetry. Comparisons between the human and murine tetramers reveal that the C^{α} r.m.s. differences are much larger (3.1, 3.7, 3.8 and 4.0 Å; Fig. 3*b*), which means the relative positions of the tetramers differ between the human and murine translins. All four superpositions show that the difference primarily arises from molecule *D* and secondarily from molecule *C* of the murine structure and is not distributed over the whole tetramer (Fig. 3*b*). In particular, α1 of molecule *C* has a large deviation of 4–6 Å from the corresponding helix of human translin (Fig. 3*c*). Consequently, the murine translin octamer has an approximately 4 Å wider cleft between the N-terminal helices.

The multimerization is mainly caused by hydrophobic interactions rather than site-specific hydrogen bonds. Several intermolecular hydrogen bonds are found between Glu207 in molecule *A* and Lys204 in molecule *B*, between Lys203 in molecule *B* and Asp211 in molecule *C*, between Tyr210 in molecule *C* and Asp211 in molecule *D*, and between Glu207 in molecule *D* and Lys204 in molecule *A*, but these hydrogen bonds are not conserved among the four monomer interfaces.

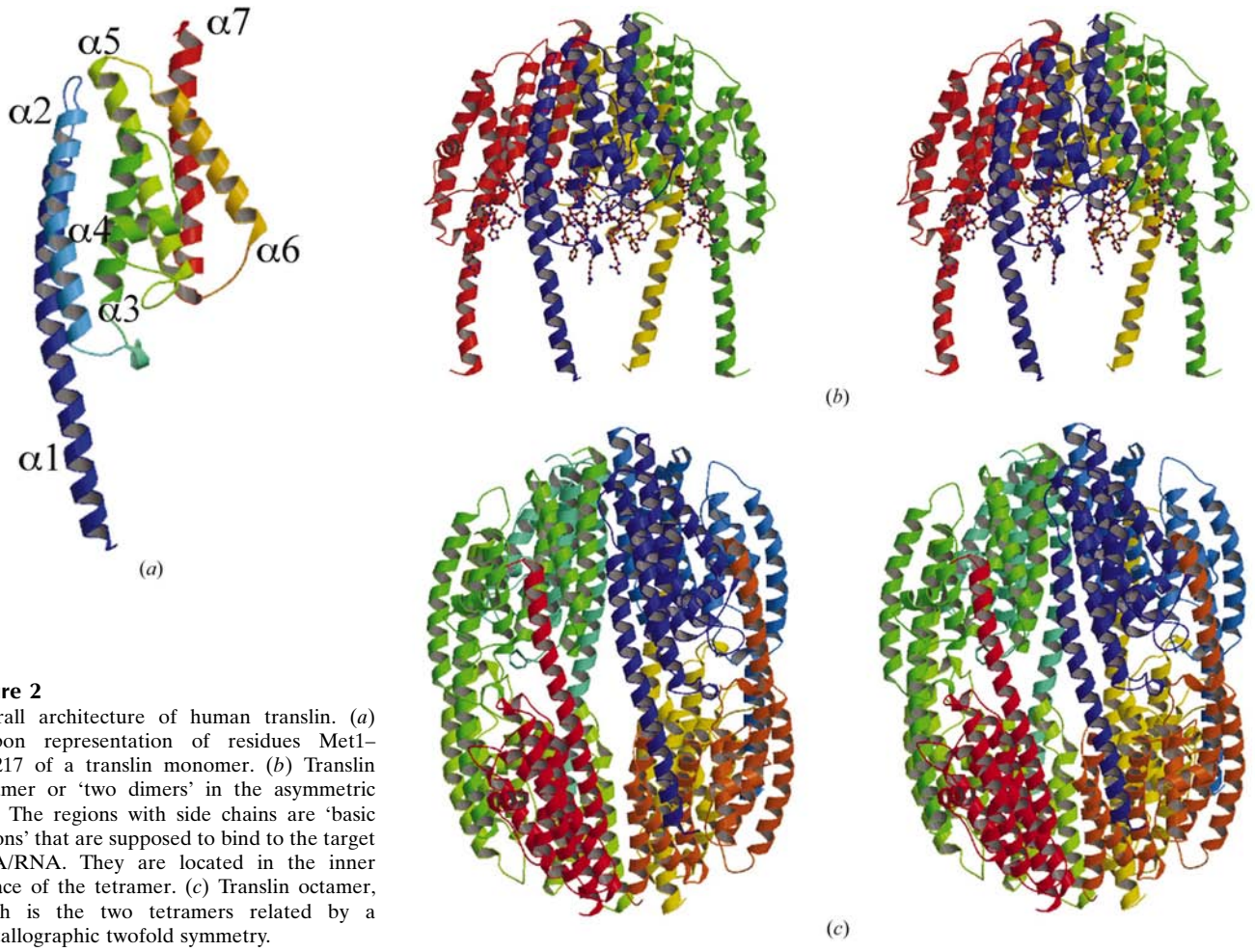

Figure 2
Overall architecture of human translin. (*a*) Ribbon representation of residues Met1–Phe217 of a translin monomer. (*b*) Translin tetramer or 'two dimers' in the asymmetric unit. The regions with side chains are 'basic regions' that are supposed to bind to the target DNA/RNA. They are located in the inner surface of the tetramer. (*c*) Translin octamer, which is the two tetramers related by a crystallographic twofold symmetry.

The octamer is formed similarly; there is no conserved hydrogen bond between the two tetramers, except that three of four pairs of Leu183 and Gln11 form hydrogen bonds and hydrophobic interaction is also dominant in the interaction between the two tetramers.

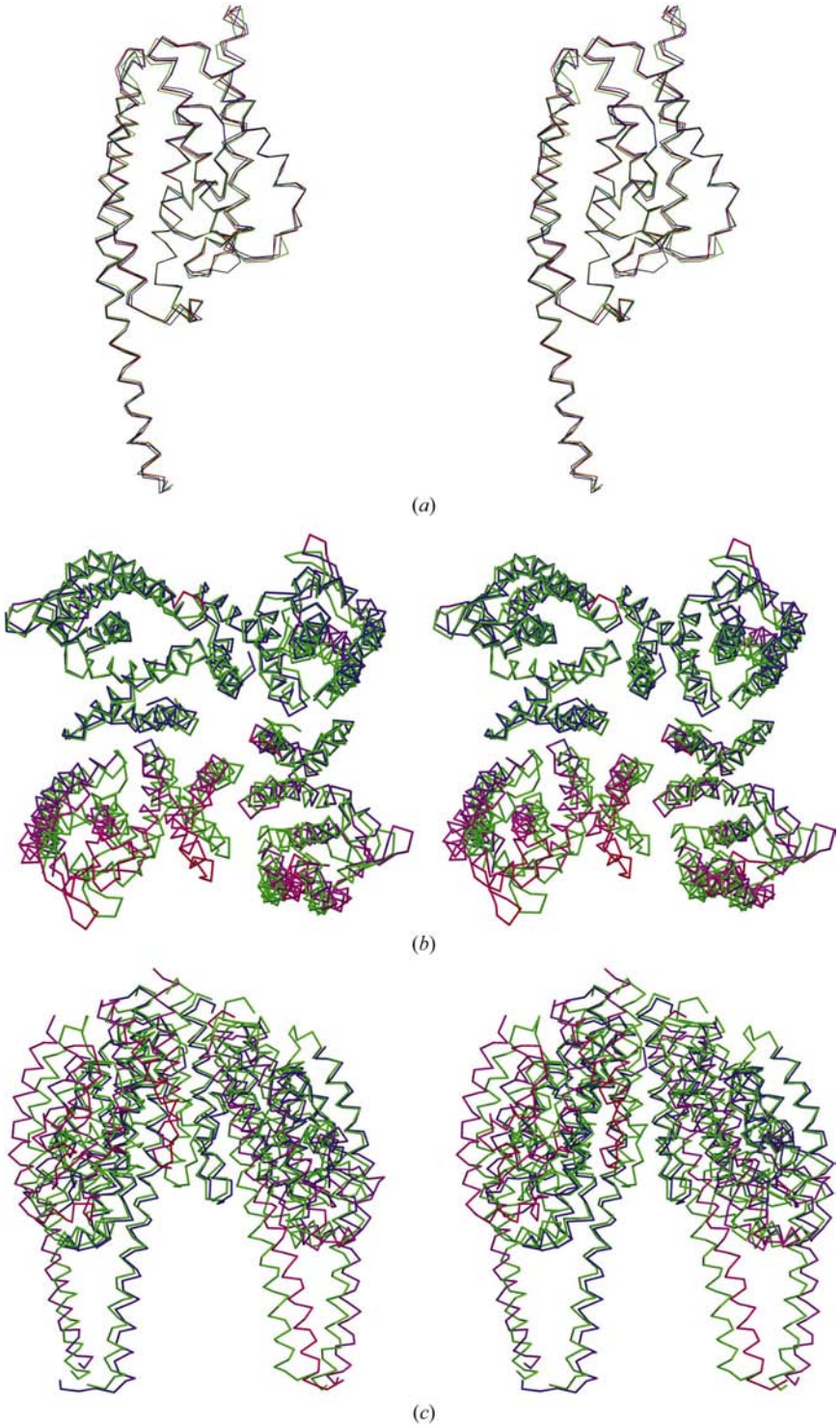

Figure 3
(*a*) Superimposition of the four monomers of human translin. (*b*) Superimposed C^{α} traces of human (green; the upper right molecule is *A*, then *D*, *C* and *B* clockwise) and murine (from blue to red; the upper right molecule is *A*, then *C*, *D* and *B* clockwise; Pascal *et al.*, 2002; PDB code 1key) translin tetramers. This superposition gives aC^{α} r.m.s. difference of 3.1 Å. The colours of the murine structure represent how different it is from the human translin tetramer; the colour is ramped from 0 Å (blue) to 8.5 Å (red). (*c*) Rotated by about 90° with respect to (*b*). The front left molecule of murine translin is molecule *D* and the front right is *C*.

These results suggest that the multimerization of translin is flexible and not based on strong specific intermolecular interactions and that the relative positions of monomers can easily be changed.

On the other hand, this octameric structure in the crystal is not consistent with the electron-microscopic (EM) structure. Although the EM structure is also ring-shaped, it does not have a twofold symmetry perpendicular to the ring and DNA/RNA binds to one side of the ring. It looks more like the tetramer in the crystallographic asymmetric unit. In addition, the basic region, which consists of residues 86–92 and is supposed to bind DNA/RNA, is located on the inside of the crystallographic octamer (Figs. 2*b* and 2*c*). There are several possibilities for DNA/RNA binding: (i) the structure of translin is in equilibrium between a tetramer and an octamer (and possibly a dimer) and the tetramer binds to DNA/RNA, (ii) the present octamer is the binding form, in which the wide cleft observed in the murine translin octamer indicates an entrance for DNA/RNA, or (iii) the octamer of translin could drastically change to another form of octamer which is seen in the EM structure.

To clarify the situation, further crystallographic and electron-microscopic studies on translin complexed with the target DNA/RNA are under way.

References

Aoki, K., Ishida, R. & Kasai, M. (1997). *FEBS Lett.* **401**, 109–112.
Aoki, K., Suzuki, K., Sugano, T., Tasaka, T., Nakahara, K., Kuge, O., Omori, A. & Kasai, M. (1995). *Nature Genet.* **10**, 167–174.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Erdemir, T., Bilican, B., Oncel, D., Goding, C. R. & Yavuzer, U. (2002). *J. Cell. Sci.* **115**, 207–216.
Han, J. R. & Hecht, N. B. (1995). *Biol. Reprod.* **53**, 707–717.
Ishida, R., Okado, H., Sato, H., Shionoiri, C., Aoki, K. & Kasai, M. (2002). *FEBS Lett.* **525**, 105–110.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kasai, M., Matsuzaki, T., Katayanagi, K., Omori, A., Maziarz, R. T., Strominger, J. L., Aoki, K. & Suzuki, K. (1997). *J. Biol. Chem.* **272**, 11402–11407.
Kissinger, C. E., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leslie, A. G. W. (1991). *Crystallographic Computing V*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 27–38. Oxford University Press.
McRee, D. E. (1999). *Practical Protein Crystallography*, 2nd ed. San Diego: Academic Press.
Merritt, E. A. & Bacon, M. (1997). *Methods Enzymol.* **277**, 505–524.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pascal, J. M., Hart, P. J., Hecht, N. B. & Robertus, J. D. (2002). *J. Mol. Biol.* **319**, 1049–1057.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
VanLoock, M., Yum X., Kasai, M. & Egelman, E. H. (2001). *J. Struct. Biol.* **135**, 58–66.
Wu, K.-Q., Xu, L. & Hecht, N. B. (1998). *Nucleic Acids. Res.* **26**, 1675–1680.