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

Structure of a UPF0150-family protein from *Thermus thermophilus* HB8

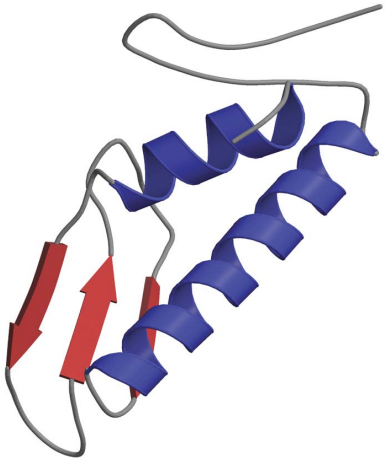
Nobuo Okazaki,[a] Maki Kumei,[a] Miho Manzoku,[a] Seiki Kuramitsu,[a,b] Mikako Shirouzu,[a,c] Akeo Shinkai[a] and Shigeyuki Yokoyama[a,c,d]*

[a]RIKEN SPring-8 Center, Harima Institute, Sayo, Hyogo 679-5148, Japan, [b]Department of Biology, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan, [c]RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan, and [d]Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Correspondence e-mail: yokoyama@biochem.s.u-tokyo.ac.jp

Received 24 November 2006
Accepted 5 February 2007

PDB Reference: TTHA0281, 2dsy, r2dsysf.



TTHA0281 is a hypothetical protein from *Thermus thermophilus* HB8 that belongs to an uncharacterized protein family, UPF0150, in the Pfam database and to COG1598 in the National Center for Biotechnology Information Database of Clusters of Orthologous Groups. The X-ray crystal structure of the protein was determined by a multiple-wavelength anomalous dispersion technique and was refined at 1.9 Å resolution to a final *R* factor of 18.5%. The TTHA0281 monomer adopts an α-β-β-β-α fold and forms a homotetramer. Based on the properties and functions of structural homologues of the TTHA0281 monomer, the TTHA0281 protein is speculated to be involved in RNA metabolism, including RNA binding and cleavage.

1. Introduction

Thermus thermophilus HB8 is an extremely thermophilic bacterium that can grow at 358 K (Ohshima & Imabori, 1974). The genome of this strain is composed of 1.85 Mbp chromosomal DNA, the 0.26 Mbp plasmid pTT27 and the 9.32 kbp plasmid pTT8. Recently, the complete genomic sequence has been determined (NCBI accession Nos. NC_006461, NC_006462 and NC_006463) and structural and functional genomic studies of this strain have been initiated (Yokoyama, Hirota *et al.*, 2000; Yokoyama, Matsuo *et al.*, 2000). About 2200 open reading frames (ORFs) have been identified within the genome and about 47.7% of them encode hypothetical proteins that are not assigned to functional categories (PEDANT database; http://pedant.gsf.de/index.jsp). One of the aims of structural genomics is to determine the three-dimensional structures of these hypothetical proteins, which could lead to the discovery of novel protein folds or to predictions of their functions.

The UPF0150 family in the Pfam database (Bateman *et al.*, 2002) is an uncharacterized protein family containing 281 bacterial, 44 archaeal and five virus proteins (http://pfam.janelia.org/cgi-bin/getdesc?name=UPF0150). In the *T. thermophilus* HB8 genome, five ORFs encoding the UPF0150-family proteins have been identified. Here, we describe the crystal structure of one of the family proteins, TTHA0281, and discuss its function based on structural homology.

2. Materials and methods

2.1. Preparation and gel-filtration analysis of the TTHA0281 protein

The TTHA0281 gene was cloned under the control of the T7 promoter on the pET-11a expression vector (Novagen, Madison, WI, USA). In order to obtain the selenomethionine-labelled TTHA0281 protein (Se-TTHA0281), the *E. coli* methionine-auxotroph strain Rosetta834(DE3), which we obtained by introducing the pRARE plasmid carrying the tRNA genes for the codons AUA, AGG, AGA, CUA, CCC and GGA (Novagen) into the B834(DE3) strain (Novagen), was used as the host. The recombinant strain was cultured in minimal medium containing 25 μg ml^{-1} L-selenomethionine, 30 μg ml^{-1} chloramphenicol and 50 μg ml^{-1} ampicillin. When an A_{600} of 0.7 was attained, isopropyl β-D-thiogalactopyranoside was added to a final concentration of 1 m*M* and the culture was incubated at 310 K for a further 5 h. The cells were collected by centrifugation and were disrupted by sonication in 20 m*M* Tris–HCl buffer pH 8.0

containing 50 m*M* NaCl. The soluble fraction obtained after centrifugation at 15 000*g* for 20 min was heated at 343 K for 30 min. After the denatured proteins had been removed by centrifugation at 15 000*g* for 20 min, the sample was applied onto a Resource ISO column (GE Healthcare Bioscience Corp., Piscataway, NJ, USA) pre-equilibrated with 50 m*M* sodium phosphate buffer pH 7.0 containing 1.5 *M* $(NH_4)_2SO_4$. The bound proteins were eluted with a linear gradient of 1.5–0 *M* $(NH_4)_2SO_4$. The fractions containing Se-TTHA0281 were collected and applied onto a Resource Q column (GE Healthcare Bioscience Corp.), which was eluted with a linear gradient of 0–1 *M* NaCl. The fractions containing Se-TTHA0281 were collected and applied onto a BioScale CHT5-I hydroxyapatite column (Bio-Rad Laboratories Inc., Hercules, CA, USA) pre-equilibrated with 10 m*M* sodium phosphate buffer pH 7.0. The flowthrough fraction was collected, concentrated and then applied onto a HiLoad 16/60 Superdex 75 pg column (GE Healthcare Bioscience Corp.) preequilibrated with 20 m*M* Tris–HCl buffer pH 8.0 containing 0.15 *M* NaCl. Finally, the purified protein was applied onto a HiPrep 26/10 Desalting column (GE Healthcare Bioscience Corp.) pre-equilibrated with 20 m*M* Tris–HCl buffer pH 8.0. The sample was concentrated to 7.9 mg ml^{-1} with a Vivaspin 6 concentrator (5 kDa molecular-weight cutoff, Sartorius AG, Goettingen, Germany).

The TTHA0281 protein and molecular-weight standards were applied onto a Superdex 200 3.2/30 (GE Healthcare Bioscience Corp.) column, using HPLC (model 1100, Agilent Technologies, Palo Alto, CA) at a flow rate of 50 µl min^{-1}. The elution buffer was 20 m*M* Tris–HCl buffer pH 8.0 containing 1 m*M* DTT.

2.2. Crystallization

Crystals of the Se-TTHA0281 protein were obtained by the microbatch method (Chayen, 1990) at 291 K using a TERA crystallization robot (Takeda Rika Kogyo Co. Ltd, Tokyo, Japan). A 0.5 µl aliquot of crystallization reagent, consisting of 17.5%(*w*/*w*) PEG 4000, 0.1 *M* CHES buffer pH 9.4 containing 50 m*M* $MgCl_2$, was mixed with 0.5 µl 7.9 mg ml^{-1} protein solution and the mixture was then covered with 15 µl silicone and paraffin oil. Octahedral crystals grew in ten weeks to approximate dimensions of 0.2 × 0.15 × 0.15 mm.

Table 1
Summary of data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Peak	Edge	Remote
Wavelength (Å)	0.9791	0.9794	0.9815
Space group	$P2_1$		
Unit-cell parameters (Å, °)	a = 38.00, b = 62.74, c = 65.72, β = 104.3		
Resolution (Å)	1.9 (1.97–1.90)	1.9 (1.97–1.90)	1.9 (1.97–1.90)
No. of measured reflections	89118	88350	86333
No. of unique reflections	24539	24536	24497
R_{merge}† (%)	6.9 (26.3)	7.0 (33.5)	5.7 (40.4)
Completeness of data (%)	100 (99.9)	99.9 (99.5)	99.2 (93.1)
$I/\sigma(I)$ (%)	10.9 (5.2)	10.8 (4.0)	12.1 (2.8)
R factor‡ (%)	18.5		
R_{free}‡ (%)	23.2		
No. of protein atoms	2495		
No. of water molecules	174		
R.m.s.d. bonds (Å)	0.015		
R.m.s.d. angles (°)	1.487		

† $R_{merge} = \sum|I - \langle I\rangle|/\sum I$. ‡ R factor and $R_{free} = \sum||F_o| - |F_c||/\sum|F_o|$, where the free reflections (5% of total used) were held aside for calculation of R_{free} throughout refinement.

2.3. X-ray diffraction and structure determination

Data were collected at the RIKEN Structural Genomics Beamline II (BL26B2) (Ueno *et al.*, 2006) at SPring-8 (Hyogo, Japan). The crystals were mounted in a nylon loop and cooled in an N_2-gas stream at 100 K using the SPring-8 Precise Automatic Cryo-sample Exchanger (SPACE), which was controlled using the beamline-scheduling software *BSS* (Ueno *et al.*, 2004, 2005). Multiple-

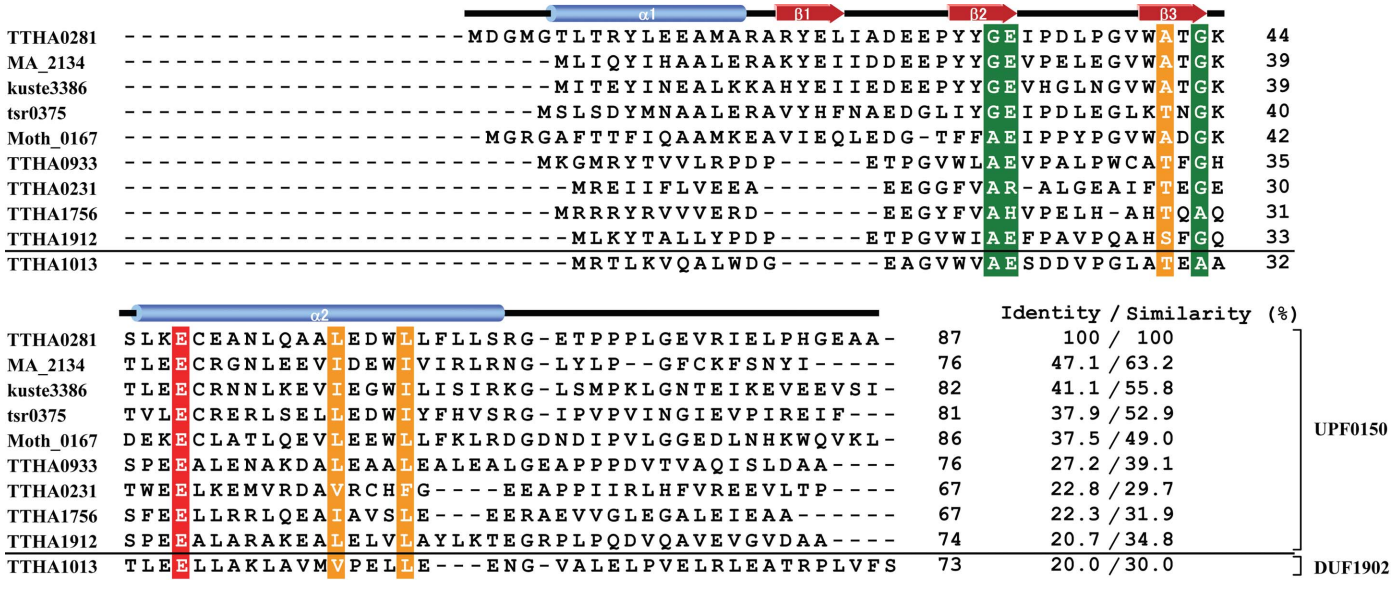


Figure 1
Amino-acid sequence alignment of TTHA0281 with homologous proteins. The amino-acid sequences of *T. thermophilus* TTHA0281, TTHA1912, TTHA1756, TTHA0933, TTHA0231 and TTHA1013, *Methanosarcina acetivorans* MA_2134, *Candidatus Kuenenia stuttgartiensis* kuste3386, *Synechococcus elongatus* tsr0375 and *Moorella thermoacetica* Moth_0167 were aligned using the program *ClustalW* (Thompson *et al.*, 1994). The Pfam database families to which these proteins belong are indicated on the right. Identical residues in all sequences, conserved substitutions and semi-conserved substitutions in the alignment are represented by white letters on red, orange and green backgrounds, respectively. The percentage identity and similarity to TTHA0281 are indicated and were obtained using the alignment program within the *European Molecular Biology Open Software Suite* (http://www.ebi.ac.uk/). Secondary-structure elements of TTHA0281, prepared using the *DSSP* program (Kabsch & Sander, 1983), are shown above the alignment. Blue, α-helix; red, β-strand.

wavelength anomalous dispersion (MAD) data were collected utilizing the anomalous scattering from Se atoms. The data sets were collected to 1.9 Å resolution using a Jupiter210 CCD detector (Rigaku MSC Co., Tokyo, Japan). The data were processed using the *HKL*-2000 program suite (Otwinowski & Minor, 1996). The atomic positions of the Se atoms in the unit cell were determined using the program *SOLVE* (Terwilliger & Berendzen, 1999) and density modification was then performed using the program *RESOLVE* (Terwilliger, 2001). The automatic tracing procedure in the program *ARP*/*wARP* (Morris *et al.*, 2003) was utilized to build the initial model. Model refinement was carried out using the programs *REFMAC* (Murshudov *et al.*, 1997) and *Coot* (Emsley & Cowtan, 2004) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Initial picking and manual verifying of water molecules were performed with the program *Coot* (Emsley & Cowtan, 2004). According to *PROCHECK* (Laskowski *et al.*, 1993), the final model has 93.8% of the residues in the most favoured conformation of the Ramachandran plot and no residues in the disallowed regions. Data-collection statistics and processed data statistics are shown in Table 1.

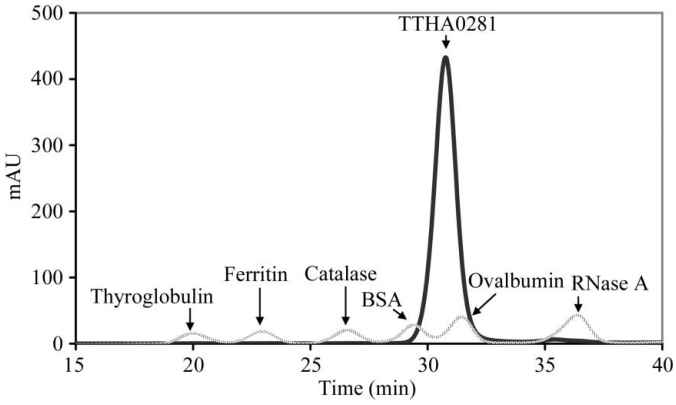


Figure 2
Gel-filtration analysis of the recombinant TTHA0281 protein. The purified TTHA0281 protein (1.9 µg per 0.2 µl, black line) and 0.6 µl molecular-weight standards (grey line), containing 0.2 µg thyroglobulin (669 kDa), 0.4 µg ferritin (440 kDa), 0.1 µg catalase (158 kDa) and 0.3 µg each of RNase A (13.7 kDa), ovalbumin (43 kDa) and bovine serum albumin (BSA) (67 kDa), were analyzed by gel-filtration column chromatography. The elution profile of the sample was detected by measuring A_{215}. The peak for each protein is indicated by an arrow.

3. Results and discussion

3.1. Primary structure and oligomerization of TTHA0281

The open reading frame (ORF) named TTHA0281 encodes a hypothetical protein composed of 87 amino-acid residues, with a predicted molecular weight of 9.68 kDa and a potential pI of 4.16. A database search showed that the TTHA0281 protein is included in an uncharacterized protein family, UPF0150, in the Pfam database

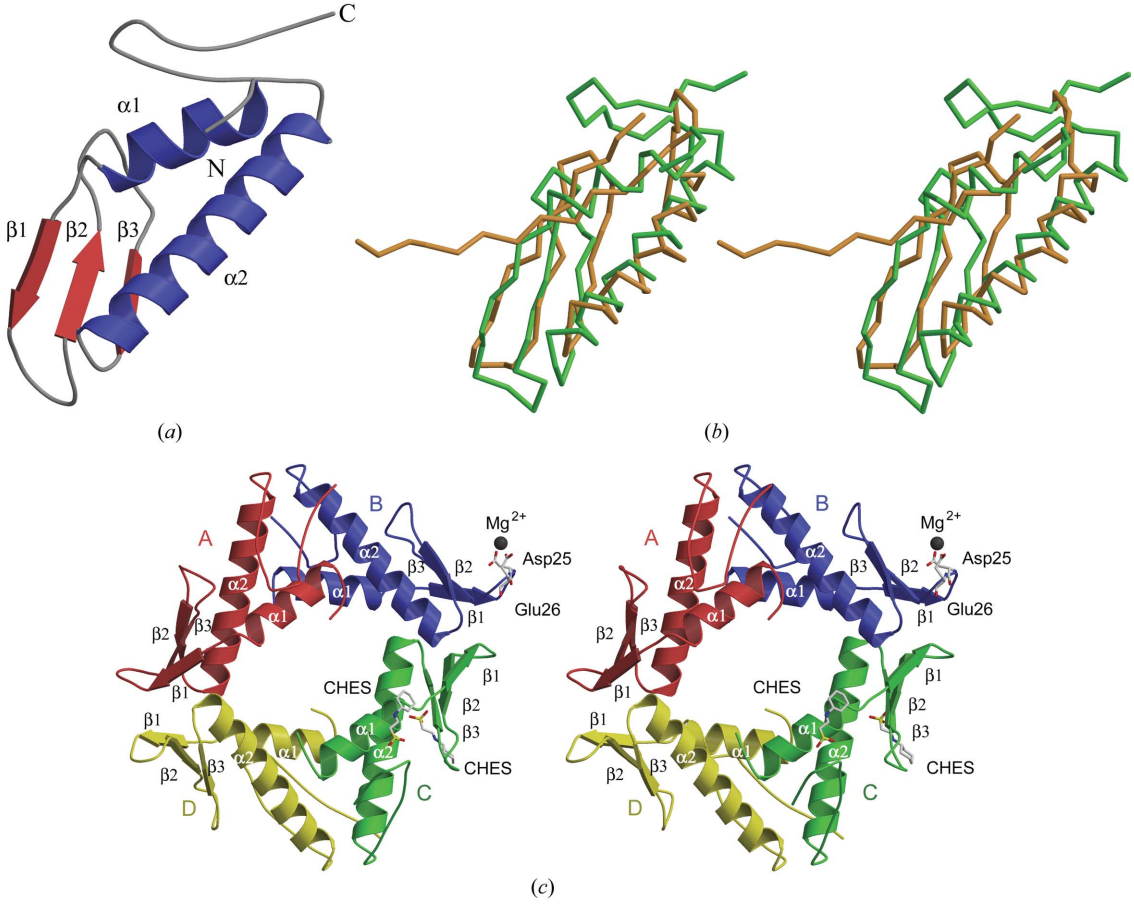


Figure 3
Schematic representations of the overall fold of TTHA0281. (*a*) The monomer structure (subunit *A*) of TTHA0281. Blue and red represent α-helix and β-strand, respectively. (*b*) Superposed C^{α} traces of TTHA0281 (green) and TTHA1013 (brown) (Hattori *et al.*, 2005). (*c*) The tetrameric structure of TTHA0281. The subunits *A*–*D*, CHES and Mg^{2+} are indicated. The residues that interact with Mg^{2+} are shown as stick models. These figures were prepared using *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Murphy, 1994).

(Bateman *et al.*, 2002) and in COG1598 in the COG database (Tatusov *et al.*, 2003). In addition to the TTHA0281 protein, the *T. thermophilus* HB8 genome contains four ORFs encoding the UPF0150-family proteins TTHA1912, TTHA1756, TTHA0933 and TTHA0231 (http://pfam.janelia.org/cgi-bin/getdesc?name=UPF0150). An amino-acid sequence alignment of the representative UPF0150-family proteins, including those from *T. thermophilus* HB8, is shown in Fig. 1.

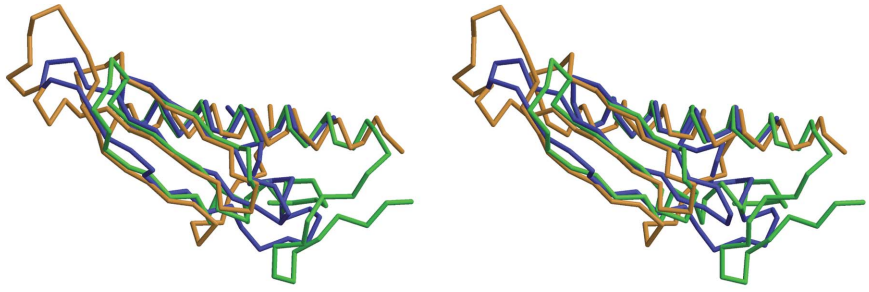
The recombinant Se-TTHA0281 protein expressed in *E. coli* cells was purified from the heat-treated cell lysate by hydrophobic, anion-exchange, hydroxyapatite and gel-filtration column-chromatography steps. The purified protein (more than 95% purity; data not shown) was analyzed by gel filtration to determine its state in solution. The retention time of the protein was around 31 min and the apparent molecular weight of the TTHA0281 protein was estimated to be around 49 kDa (Fig. 2). These results indicate that the TTHA0281 protein forms an oligomer in solution.

3.2. Overall structure

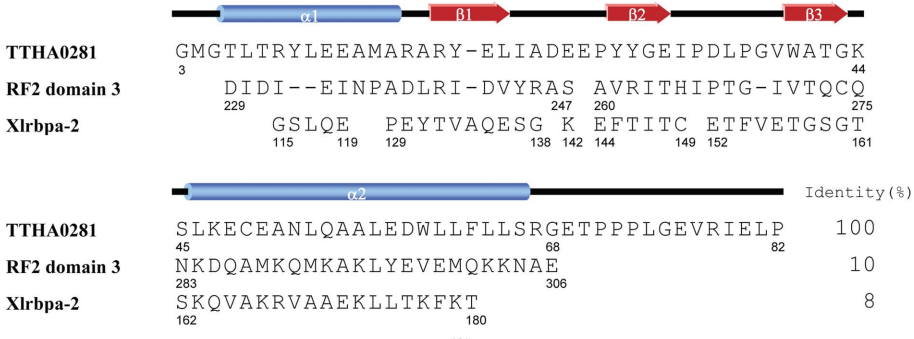
TTHA0281 exists as a single-domain structure of 87 residues with two α-helices, $\alpha1$ and $\alpha2$, and three β-strands, $\beta1$, $\beta2$ and $\beta3$ (Fig. 3*a*). This arrangement forms a three-stranded twisted antiparallel β-sheet flanked by $\alpha2$, known as an α-β-β-β-α fold, which is also present in several double-helical nucleic acid-binding proteins, such as ribosomal protein S5 and RNase III (Ramakrishnan & White, 1992; Bycroft *et al.*, 1995; Kharrat *et al.*, 1995). The TTHA0281 protein formed a tetramer in the asymmetric unit of the crystal (Fig. 3*c*). This is consistent with the fact that the TTHA0281 protein forms an oligomer in solution, although the molecular weight in solution is estimated to be slightly higher than that of the tetrameric state of this protein, as described above. The interactions between subunits *A* and *B* and between subunits *C* and *D* are stabilized by hydrogen-bonding interactions between the N atom of Thr8 in $\alpha1$ and the side-chain O atom of Glu58 in $\alpha2$. The amido O atoms of Lys44 in the loop between $\beta3$ and $\alpha2$, the side-chain N atom of Lys47 in $\alpha2$ and the side-chain hydroxyl group of Tyr29 in $\beta2$ of subunit *A* form hydrogen bonds to the side-chain amido group of Arg19 in $\beta1$, the amido O atom of Glu12 in $\alpha1$ and the N atom of Tyr20 in $\beta1$ of subunit *D*, respectively; furthermore, the side-chain N atom of Lys47 in $\alpha2$, the side-chain hydroxyl group of Tyr29 in $\beta2$ and the N atom of Tyr20 in $\beta1$ of subunit *B* form hydrogen bonds to the amido O atom of Ala18 in the loop between $\alpha1$ and $\beta1$, the side-chain carboxyl group of Glu21 in $\beta1$ and the side-chain hydroxyl group of Tyr29 in $\beta2$ of subunit *C*, respectively. The interaction surface areas at the interfaces of subunits *AB* and *AD*, calculated with the *AREAIMOL* program (Collaborative Computational Project, Number 4, 1994), are ~800 and ~600 Å^2, which are 15% and 11% of the total surface area of subunit *A*, respectively (Fig. 3*c*). In the crystal, CHES molecules, which may be derived from the crystallization reagent, were found around Glu13 and Trp40 of subunit *C* in the difference Fourier map, at 2σ and 4σ (Fig. 3*c*). A magnesium ion was located around the side chains of Asp25 and Glu26 in subunit *B* (Fig. 3*c*).

3.3. Structure comparison and functional prediction

The overall folding of the hypothetical protein TTHA1013, which is included in the DUF1902 family of the Pfam database, is arranged as a $\beta1$-$\beta2$-$\beta3$-$\alpha1$-$\beta4$ fold and it forms a dimer with a symmetry-related molecule (PDB code 1wv8; Hattori *et al.*, 2005). The dimer interface is composed of the $\beta4$-strands. The $\beta1$-$\beta2$-$\beta3$-$\alpha1$ fold of TTHA1013 is suggested to be a structural homologue of the partially degraded RNase H fold of the HicB-family protein, which belongs to the UPF0150 and COG1598/COG4226 families (Makarova *et al.*, 2006). The $\beta1$-$\beta2$-$\beta3$-$\alpha1$ fold of TTHA1013 is superimposable on $\beta1$-$\beta2$-$\beta3$-$\alpha2$ of TTHA0281 (subunit *A*) with a *Z* score of 3.9 and a



(*a*)



(*b*)

Figure 4
(*a*) Superposed C$^\alpha$ traces of TTHA0281 (green), RF2 domain 3 (brown) and Xlrbpa-2 (blue). This figure was prepared using *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Murphy, 1994). (*b*) Structure-based alignment of TTHA0281, RF2 domain 3 and Xlrbpa-2. The numbers on the residues represent the position from the N-terminus of the entire protein. Secondary-structure elements of TTHA0281, prepared as described in the legend to Fig. 1, are shown above the alignment. The sequence identity was calculated using the *DALI* server (http://www.ebi.ac.uk/dali/; Holm & Sander, 1998).

root-mean-square deviation (r.m.s.d.) of 2.6 Å, although the amino-acid sequence identity is 20% (Figs. 1 and 3*b*) as calculated using the *DALI* server (http://www.ebi.ac.uk/dali/; Holm & Sander, 1998). Thus, the β-β-β-α folds are similar to each other, although the two proteins belong to different Pfam families. This result suggests that when the β-β-β-α fold is preceded by an α-helix it forms a tetramer, as in the case of TTHA0281. In contrast, when the β-β-β-α fold is followed by a β-strand it forms a dimer at the C-terminal β4 strand, as in the case of TTHA1013. Interestingly, Glu36 in the α-helix of the TTHA1013 protein, which corresponds to one of the essential residues for catalysis in several nuclease superfamilies of the RNase H fold (Makarova *et al.*, 2006), is conserved in the TTHA0281 protein (Glu48 in the α2-helix; Fig. 1).

Furthermore, two structural homologues of the TTHA0281 monomer were identified with *Z* scores of >4 (Fig. 4*a*) in the PDB using the *DALI* server. One is domain 3 of the bacterial polypeptide release factor RF2 (PDB code 1gqe chain *A*, $Z = 4.3$, r.m.s.d. = 3.1 Å), which promotes the termination of protein synthesis by recognizing the stop codon UAA or UAG (Vestergaard *et al.*, 2001). Based on a structure analysis, domain 3 may interact with the peptidyl-transferase centre of the 50S ribosomal subunit and its distal end lies close to the 23S rRNA (Rawat *et al.*, 2003; Klaholz *et al.*, 2003). Another structural homologue is the second double-stranded RNA-binding domain of *Xenopus laevis* RNA-binding protein A (Xlrbpa-2; PDB code 1di2 chain *A*, $Z = 4.1$, r.m.s.d. = 3.0 Å; Ryter & Schultz, 1998). Interestingly, the primary structures of RF2 domain 3 and Xlrbpa-2 exhibited low sequence similarity to that of the TTHA0281 protein (Fig. 4*b*).

Considering the properties and functions of the aforementioned structural homologues of the TTHA0281 protein, we speculate that it is involved in RNA metabolism, including RNA binding and cleavage.

We thank Dr Koji Takio for the gel-filtration analysis of the TTHA0281 protein. This work was supported by the RIKEN Structural Genomic/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M. & Sonnhammer, E. L. (2002). *Nucleic Acids Res.* **30**, 276–280.

Bycroft, M., Grunert, S., Murzin, A. G., Proctor, M. & St Johnston, D. (1995). *EMBO J.* **14**, 3563–3571.

Chayen, E. (1990). *J. Appl. Cryst.* **23**, 297–302.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Hattori, M., Mizohata, E., Manzoku, M., Bessho, Y., Murayama, K., Terada, T., Kuramitsu, S., Shirouzu, M. & Yokoyama, S. (2005). *Proteins*, **61**, 1117–1120.

Holm, L. & Sander, C. (1998). *Nucleic Acids Res.* **26**, 316–319.

Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.

Kharrat, A., Macias, M. J., Gibson, T. J., Nilges, M. & Pastore, A. (1995). *EMBO J.* **14**, 3572–3584.

Klaholz, B. P., Pape, T., Zavialov, A. V., Myasnikov, A. G., Orlova, E. V., Vestergaard, B., Ehrenberg, M. & van Heel, M. (2003). *Nature (London)*, **421**, 90–94.

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.

Makarova, K. S., Grishin, N. V. & Koonin, E. V. (2006). *Bioinformatics*, **22**, 2581–2584.

Merritt, E. A. & Murphy, M. E. (1994). *Acta Cryst.* D**50**, 869–873.

Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Ohshima, T. & Imabori, K. (1974). *Int. J. Syst. Bacteriol.* **24**, 102–112.

Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.

Ramakrishnan, V. & White, S. W. (1992). *Nature (London)*, **358**, 768–771.

Rawat, U. B., Zavialov, A. V., Sengupta, J., Valle, M., Grassucci, R. A., Linde, J., Vestergaard, B., Ehrenberg, M. & Frank, J. (2003). *Nature (London)*, **421**, 87–90.

Ryter, J. M. & Schultz, S. C. (1998). *EMBO J.* **17**, 7505–7513.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J. & Natale, D. A. (2003). *BMC Bioinformatics*, **4**, 41.

Terwilliger, T. C. (2001). *Acta Cryst.* D**57**, 1763–1775.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Ueno, G., Hirose, R., Ida, K., Kumasaka, T. & Yamamoto, M. (2004). *J. Appl. Cryst.* **37**, 867–873.

Ueno, G., Kanda, H., Hirose, R., Ida, K., Kumasaka, T. & Yamamoto, M. (2006). *J. Struct. Funct. Genomics*, **7**, 15–22.

Ueno, G., Kanda, H., Kumasaka, T. & Yamamoto, M. (2005). *J. Synchrotron Radiat.* **12**, 380–384.

Vestergaard, B., Van, L. B., Andersen, G. R., Nyborg, J., Buckingham, R. H. & Kjeldgaard, M. (2001). *Mol. Cell*, **8**, 1375–1382.

Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., Kyogoku, Y., Miki, K., Masui, R. & Kuramitsu, S. (2000). *Nature Struct. Biol.* **7**, *Suppl.*, 943–945.

Yokoyama, S., Matsuo, Y., Hirota, H., Kigawa, T., Shirouzu, M., Kuroda, Y., Kurumizaka, H., Kawaguchi, S., Ito, Y., Shibata, T., Kainosho, M., Nishimura, Y., Inoue, Y. & Kuramitsu, S. (2000). *Prog. Biophys. Mol. Biol.* **73**, 363–376.