Crystal structure of the YjgF/YER057c/UK114 family protein from the hyperthermophilic archaeon Sulfolobus tokodaii strain 7

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · FEBRUARY 2005

Impact Factor: 2.63 · DOI: 10.1002/prot.20778 · Source: PubMed

8 AUTHORS, INCLUDING:

Takuya Miyakawa
The University of Tokyo
90 PUBLICATIONS **1,097** CITATIONS
SEE PROFILE

Ken-ichi Hatano
Gunma University
30 PUBLICATIONS **404** CITATIONS
SEE PROFILE

Koji Nagata
The University of Tokyo
140 PUBLICATIONS **2,007** CITATIONS
SEE PROFILE

Masaru Tanokura
The University of Tokyo
454 PUBLICATIONS **6,897** CITATIONS
SEE PROFILE

STRUCTURE NOTE

Crystal Structure of the YjgF/YER057c/UK114 Family Protein from the Hyperthermophilic Archaeon *Sulfolobus tokodaii* Strain 7

Takuya Miyakawa,[1] Woo Cheol Lee,[1] Ken-ichi Hatano,[2] Yusuke Kato,[1] Yoriko Sawano,[1] Ken-ichi Miyazono,[1] Koji Nagata,[1] and Masaru Tanokura[1]*

[1]*Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan*

[2]*Department of Biological sciences, Faculty of Engineering, Gunma University, Kiryu, Gunma, Japan*

Introduction. The YjgF/YER057c/UK114 family of proteins (referred to as the YjgF family in this article) includes over 200 prokaryotic and eukaryotic proteins consisting of ca. 130 amino acid residues.[1] The protein family members show diverse functions. In bacteria and yeasts, they play prerequisite roles in the biosynthesis of essential compounds such as isoleucine,[2–4] thiamine,[2] and purine.[5] For example, the gene disruption of a yeast member that is involved in isoleucine biosynthesis is lethal to yeast cells.[4] On the other hand, the mammalian members, whose expression is upregulated upon cellular differentiation,[6,7] inhibit translations[8,9] by degrading mRNAs.[10]

Six crystal structures have been reported thus far for the YjgF-family proteins. These proteins are homotrimeric, with each subunit having a β-barrel-like structure surrounded by two α-helices.[11–14] The overall fold and oligomeric state are quite similar to those of the chorismate mutase family.[12] The catalytic sites of YjgF proteins have not been clearly revealed, although three clefts between the two subunits are presumed to be catalytic sites by analogy to chorismate mutases.[11,12] In fact, human hp14.5 protein can entrap small molecules such as benzoic acid within the clefts.[14]

Thus far, no archaeal YjgF-family proteins have been analyzed functionally or structurally. In the present study, we determined the crystal structure of the YjgF-family protein ST0811 from the hyperthermophilic archaeon *Sulfolobus tokodaii* strain 7 [15] and compared the structure with those of bacterial and eukaryotic counterparts. Our results indicate that the archaeal YjgF protein is structurally more similar to mammalian orthologs than to bacterial ones.

Materials and Methods. *Protein production, crystallization and X-ray diffraction data collection.* Protein expression, purification, and crystallization were performed as described.[16] The crystals were obtained by the sitting-drop vapor-diffusion method at 293 K in 3 days. The reservoir solution used 16% (w/v) PEG 10,000, 0.1 *M* Bis-Tris (pH 5.3), and 0.1 *M* ammonium acetate. Crystals were transferred into the cryoprotectant solution containing 18%(w/v) PEG 10,000 as precipitant, 0.12 *M* Bis-Tris (pH 5.3), 0.12 *M* ammonium acetate, and 20% (v/v) glycerol and were then flash-cooled in a nitrogen stream. Diffraction data were collected at 100 K with an R-AXIS VII image plate detector mounted on a Rigaku FR-E rotating-anode X-ray generator (Rigaku, Japan) using the operation software CrystalClear (Rigaku/MSC) and were processed with MOSFLM[17] and SCALA.[18]

Structure determination. The structure of the ST0811 was solved by the molecular replacement method using the program MOLREP of the CCP4 suites[19] and the coordinates of *Bacillus subtilis* YabJ (PDB code 1QD9, 50% sequence identity to ST0811).[12] MOLREP was run using the data with a resolution range of 19.0–2.0 Å in the space group *R*3. Five percent of the reflections were excluded from the total for cross-validation with the R_{free} value. The rigid-body model was initially refined with the program Refmac5,[20] and several cycles of manual model rebuilding and model refinement were then performed using XtalView[21] and Refmac5. Water molecules were automatically picked up by the ARP/wARP program,[22] and they were then confirmed based on peak heights and distance criteria in the F_o-F_c and $2F_o$-F_c maps. The quality of the model was evaluated with PROCHECK.[23] The coordinates have been deposited into the Protein Data Bank with the accession number 1X25.

Structural comparison. Structural analysis was carried out using the following computer programs: Swiss-

Grant sponsor: National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology of Japan; Grant sponsor: Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan

*Correspondence to: Masaru Tanokura, Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. E-mail: amtanok@mail.ecc.u-tokyo.ac.jp

Received 3 August 2005; Accepted 20 August 2005

Published online 1 December 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20778

PdbViewer[24] for superposition of molecules, calculation of RMS deviations and secondary structure-based alignment; CCP4[19] for calculation of trimeric-contact surface; ESPript[25] for preparation of the alignment figure; PyMol (http://pymol.sourceforge.net) for the depiction of structure; GRASP[26] for calculation and depiction of electrostatic potential at the molecular surface. Other orthologs used in the structural comparison were *B. subtilis* YabJ (PDB code 1QD9),[12] *Escherichia coli* YjgF (PDB code 1QU9),[11] yeast Yeo7 (PDB code 1JD1), goat UK114 (PDB code 1NQ3),[13] rat L-PSP (PDB code 1QAH), and human hp14.5 (PDB code 1ONI).[14]

Dynamic light scattering. Dynamic light-scattering (DLS) measurements were performed with DynaPro (Protein solutions). ST0811 was prepared to 5.0 mg/mL in 10 m*M* Tris-HCl (pH 7.1) and was then measured at 298 K. DLS data were automatically collected by Dynamics v5.1 (Protein solutions), and 20 data points with baseline levels of 0.998–1.002 were used to calculate the Stokes radius and the polydispersity of ST0811.

Results and Discussion. *Crystallographic data and refinement statistics.* The crystals of ST0811 were found to belong to the rhombohedral space group *R*3, with hexagonal unit-cell parameters $a = b = 54.98$ Å and $c = 223.16$ Å. Two molecules were contained per asymmetric unit ($V_M = 2.3$ Å^3 Da^{-1}, ca. 47% solvent).[27] The crystal structure of ST0811 was determined at 2.0 Å resolution. The final electron density allowed for positioning of all the residues except the N- and C-terminal glycine residues in chain A and 210 water molecules. R_{factor} and R_{free} were 14.8 and 18.8%, respectively. In the Ramachandran plot,[28] 92.0% of the residues fell within the most favored regions, and the rest fell within the additionally allowed regions. The data collection and refinement statistics are summarized in Table I.

TABLE I. Summary of Data Collection and Refinement Statistics

Data collection[a]	
Resolution (Å)	22.0–2.00 (2.11–2.00)
Unique reflections	16882
Redundancy	6.2 (5.9)
Data completeness (%)	99.8 (99.8)
R_{sym}[b]	0.055 (0.137)
I/σ	9.8 (5.1)
Space group	*R*3
Lattice constants (Å)	$a = b = 54.98$, $c = 221.56$
No. of molecules per asymmetric unit	2
Refinement statistics	
Resolution range (Å)	19.0–2.00
Total reflections used in refinement	16877
R_{factor} (%)[c]	14.6
R_{free} (%)[c]	18.8
No. of protein atoms (and residues)	1979 (254)
No. of water molecules	210
Mean overall *B* factor (Å^2)	18.36
Rms bond length deviations (Å)	0.020
Rms bond angle deviations (deg)	1.488
Ramachandran plot	
Most favored regions (% and no. of residues)	92.0 (206)
Additional allowed regions (% and no. of residues)	8.0 (18)
Generously allowed regions (%)	0
Disallowed regions (%)	0

[a]Values in parentheses are for the highest resolution shell.
[b]$R_{sym} = \Sigma_{hkl}[(\Sigma_i|I_i - \langle I\rangle|)/\Sigma_i|I_i|]$.
[c]$R_{factor} = (\Sigma_{hkl}||F_o| - |F_c||)/\Sigma_{hkl}|F_o|$; 5% of the data were used for R_{free}.

Overall structure and trimer configuration. The overall structure of the ST0811 monomer shows an α + β fold with an arrangement of ββββαβαββ that is referred to as the chorismate mutase-like fold [Fig. 1(a)].[11,12] In the β-sheet, the order of strands is 1-2-3-6-4-5, where β4 and β5 are parallel and all other strands are antiparallel. The ST0811 protein forms a homotrimeric structure with a three fold axis [Fig. 1(b)], as do the bacterial and mammalian YjgF-family members whose structures were solved. The three β-sheets are arranged in a β-barrel-like manner and are surrounded by six α-helices. The contact area between any two subunits is 1890 Å^2, and 70% of it is occupied by hydrophobic residues, which contributes to stabilize the trimer [Fig. 1(b)]. ST0811 exists as a homotrimer not only in crystal but also in solution, as shown by the DLS data. The Stokes radius of ST0811 (27 Å) is consistent with the maximum radius of the ST0811 trimer in the crystal (28 Å).

All crystal structures of bacterial and mammalian YjgF proteins show a homotrimeric structure and can be fit to that of ST0811 with RMS deviations of 0.74–0.90 Å for the 125 C^α positions. As shown in Figure 2, the structural differences come from the inserted gap between α2 and β5 in *E. coli* YjgF and the additional residues at both the N- and C-termini in mammalian orthologs. Thus, the overall structure and oligomeric assembly would not influence their diversity in function.

The three clefts. Trimeric YjgF proteins have surface clefts located between subunits.[12,14] The ST0811 trimer has three clefts, each of which is 7 × 13 × 11 Å^3 in size. Each cleft is formed by the 13 conserved and 12 variable residues shown in Figure 2(a). Some conserved residues, Ser[29], Gly[30], Arg[102], and Glu[117], connect two subunits as a scaffold of the cleft [Fig. 2(b)]. The arrangement of residues in the cleft of ST0811 is quite similar to that of other YjgF family members, suggesting that the clefts of ST0811 would be suitable for binding small molecules such as benzoates, as observed in the clefts of hp14.5.[14]

The cavity. The ST0811 trimer has a hydrophilic cavity formed by residues on β3 and the β2-β3 turn. The cavity is separated from the three clefts by the hydrophobic clusters on the β-sheets. Thr[25] and Tyr[27], which are conserved in mammalian YjgF proteins as well as ST0811 but not in bacterial orthologs [Fig. 2(a)], convey hydrophilic properties on to the cavity, as shown in Figure 3(a). Because ST0811 does not have negatively charged Asp[23], the electrostatic potential in the cavity of ST0811 is different from that of mammalian orthologs [Fig. 3(a)].

The surface potential of the cavity side of the ST0811 is positive and is more similar to those of mammalian rather than of bacterial and yeast orthologs [Fig. 3(b)]. The surface potentials are reflective of the isoelectric points calculated from the amino acid sequences of ST0811, L-PSP, and YjgF: 8.5, 8.5, and 5.2, respectively. In addition, the structure and

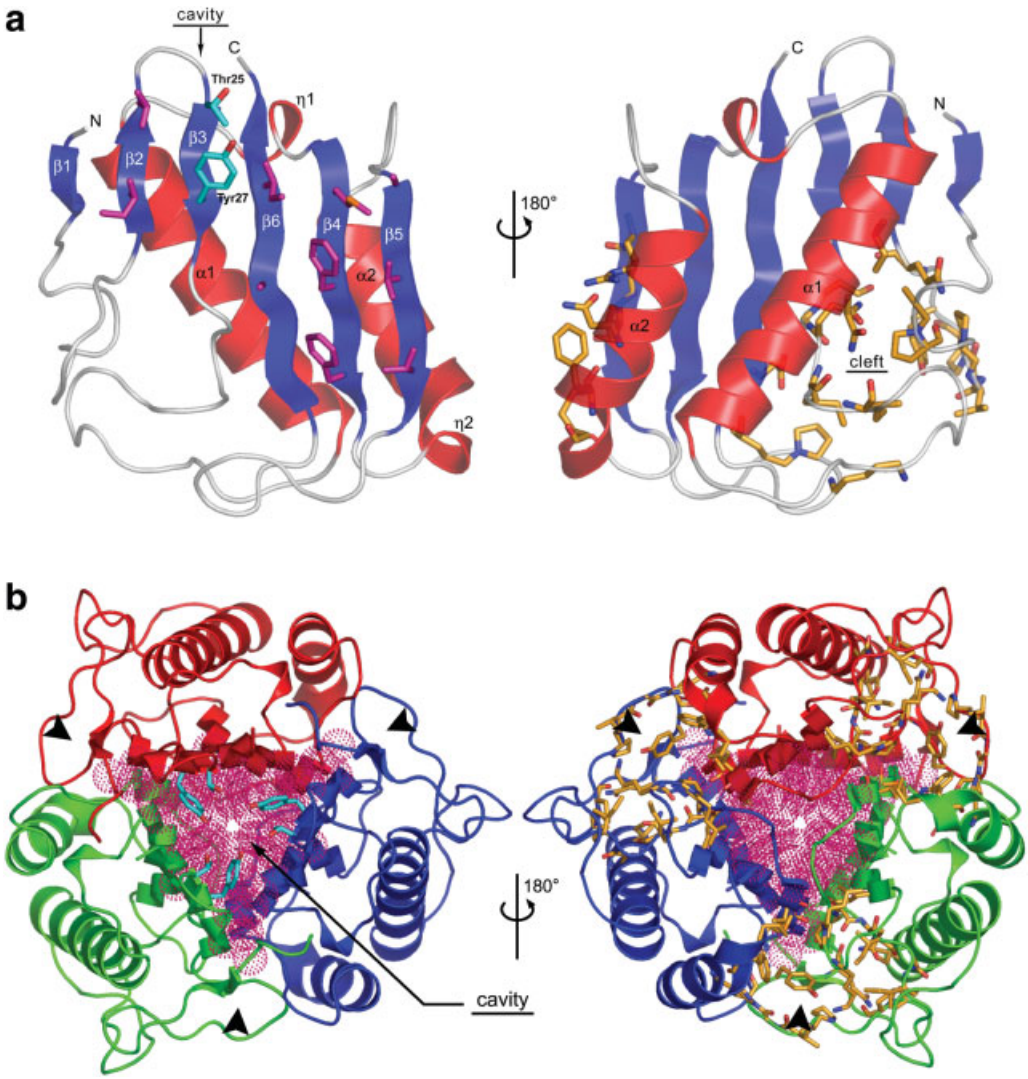

Fig. 1. (a) Ribbon diagram of the monomeric structure of ST0811. α- and 3_{10}-helices are colored red and β-strands blue. Side-chains are represented by stick models (magenta, hydrophobic residues on β-strands; cyan, hydrophilic residues in the cavity; orange, residues in the cleft). (b) Ribbon diagram of the homotrimeric structure of ST0811 viewed from the cavity side (left) and the bottom side (right). Each monomer is colored red, blue, or green. Hydrophobic cluster, represented by purple mesh, is formed by the hydrophobic side chains colored magenta in (a). The hydrophilic side-chains of Thr^{25} and Tyr^{27} protrude over the hydrophobic cluster in the cavity. The positions of the three clefts are shown with allow heads.

trimeric assembly of ST0811 was found to be stable in the presence of strong acid (5% perchloric acid) as monitored by far-UV CD spectroscopy and DLS (data not shown). This property is similar to that of mammalian YjgF members and the eukaryotic histone proteins.[6,29] Unlike bacterial orthologs, rat L-PSP and human hp14.5 are reported to possess RNase activity.[10,30] The positively charged surface of these mammalian YjgF proteins would be important to attracting negatively charged RNA molecules. The positively charged surface on the cavity side of ST0811 would be able to interact with RNA molecules in a similar manner. Although additional experimental evidence must be obtained, we propose here that ST0811 and other YjgF proteins from hyperthermophilic archaea may bind and regulate the stability of RNA molecules.

Acknowledgments. This work was supported in part by the National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology of Japan and by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

1. Parsons L, Bonander N, Eisenstein E, Gilson M, Kairys V, Orban J. Solution structure and functional ligand screening of HI0719, a highly conserved protein from bacteria to humans in the YjgF/YER057c/UK114 family. Biochemistry 2003;42:80–89.
2. Enos-Berlage JL, Langendorf MJ, Downs DM. Complex metabolic phenotypes caused by a mutation in *yjgF*, encoding a member of the highly conserved YER057c/YjgF family of proteins. J Bacteriol 1998;180:6519–6528.
3. Schmitz G, Downs DM. Reduced transaminase B (IlvE) activity

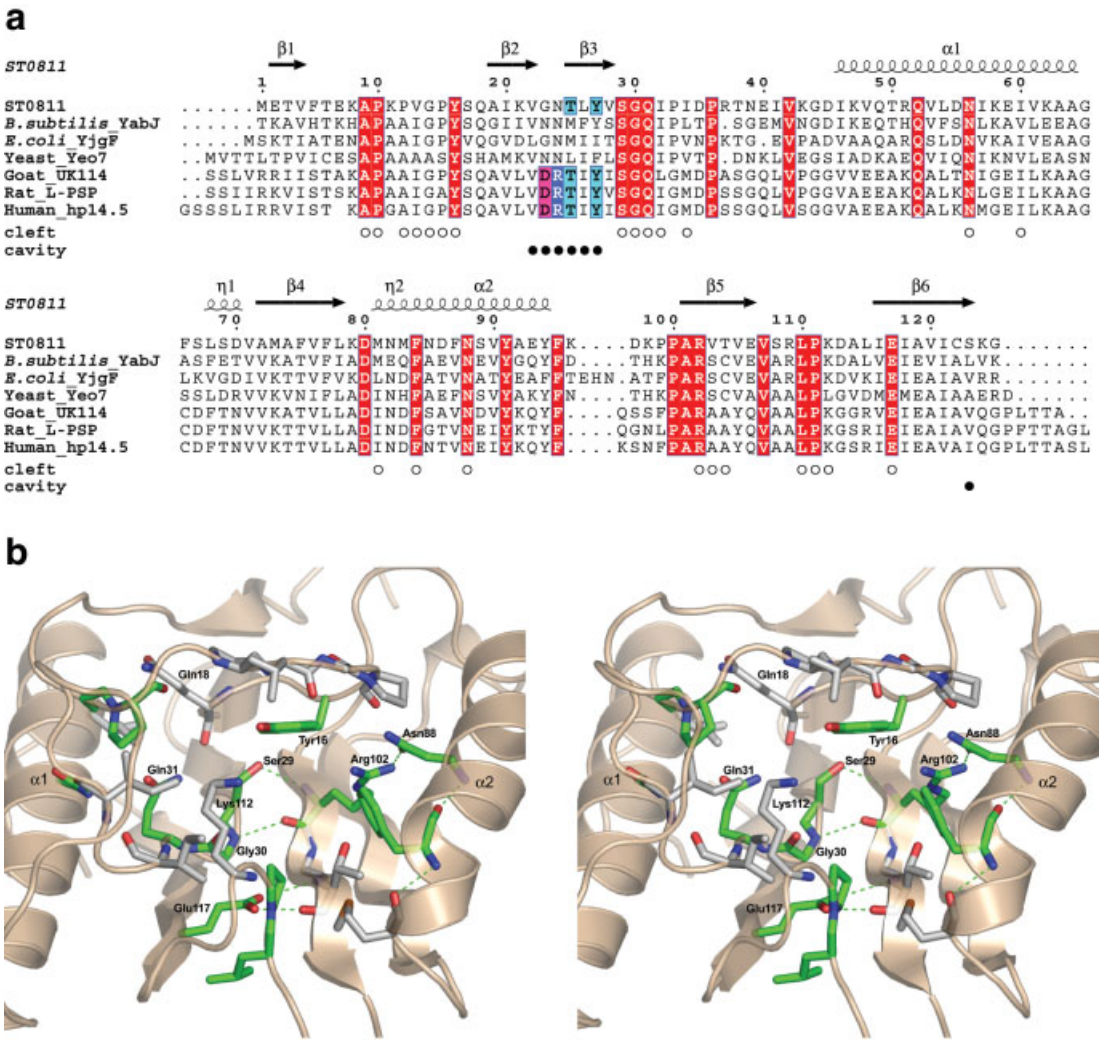

Fig. 2. (a) Structure-based sequential alignment of ST0811 and other YjgF proteins with available crystal structure. The secondary structure assignment for ST0811 is indicated by helices(α- and 3_{10}-helices) and arrows (β-strands). Residues forming the cleft and the cavity are indicated by open circles (○) and filled circles (●), respectively. Conserved residues are boxed. Hydrophilic residues in cavity, Thr^{25} and Tyr^{27}, are colored cyan, whereas charged residues around the cavity in mammalian YjgF orthologs, Asp^{23} and Arg^{24}, are red and blue, respectively. (**b**) A stereo diagram of the cleft in ST0811. All the side chains forming the cleft are shown by a stick model. Conserved residues are colored green, and hydrogen bonds are shown by broken lines.

caused by the lack of *yjgF* is dependent on the status of threonine deaminase (IlvA) in *Salmonella enterica* serovar typhimurium. J Bacteriol 2004;186:803–810.
4. Kim JM, Yoshikawa H, Shirahige K. A member of the YER057c/yjgf/Uk114 family links isoleucine biosynthesis and intact mitochondoria maintenance in *Saccharomyces cerevisiae*. Genes Cells 2001;6:507–517.
5. Rappu P, Shin BS, Zalkin H, Mäntsälä P. A role for a highly conserved protein of unknown function in regulation of *Bacillus subtilis purA* by the purine repressor. J Bacteriol 1999;181:3810–3815.
6. Asagi K, Oka T, Arao H, Suzuki I, Thakur MK, Izumi K, Natori Y. Purification, characterization and differentiation-dependent expression of a perchloric acid soluble protein from rat kidney. Nephron 1998;79:80–90.
7. Kaneki K, Kanouchi H, Matsumoto M, Kawasaki Y, Akuzawa M, Oka T. Down regulation of a novel protein, PSP, in rat hepatoma cell dRLh 84-beared tumor. J Vet Med Sci 2003;65:781–785.
8. Oka T, Tsuji H, Noda C, Sakai K, Hong YM, Suzuki I, Muñoz S, Natori Y. Isolation and characterization of a novel perchloric acid-soluble protein inhibiting cell-free protein synthesis. J Biol Chem 1995;270:30060–30067.
9. Schmiedeknecht G, Kerkhoff C, Orso E, Stohr J, Aslanidis C, Nagy GM, Knuechel R, Schmits G. Isolation and characterization of a 14.5-kDa trichloroacetic-acid-soluble translational inhibitor protein from human monocytes that is upregulated upon cellular differentiation. Eur J Biochem 1996;242:339–351.
10. Morishita R, Kawagoshi A, Sawasaki T, Madin K, Ogasawara T, Oka T, Endo Y. Ribonuclease activity of rat liver perchloric acid-soluble protein, a potent inhibitor of protein synthesis. J Biol Chem 1999;274:20688–20692.
11. Volz K. A test case for structure-based functional assignment: the 1.2 Å crystal structure of the yjgF gene product from *Escherichia coli*. Protein Sci 1999;8:2428–2437.
12. Sinha S, Rappu P, Lange SC, Mäntsälä P, Zalkin H, Smith JL. Crystal structure of *Bacillus subtilis* YabJ, a purine regulatory protein and member of the highly conserved YjgF family. Proc Natl Acad Aci USA 1999;96:13074–13079.
13. Deriu D, Briand C, Mistiniene E, Naktinis V, Grütter MG. Structure and oligomeric state of the mammalian tumour-associated antigen Uk114. Acta Crystallogr D 2003;59:1676–1678.
14. Manjasetty BA, Delbrück H, Pham DT, Mueller U, Fieber-Erdmann M, Scheich C, Sievert V, Büssow K, Neisen FH, Wilhelm W, Loll B, Saenger W, Heinemann U. Crystal structure of *Homo sapiens* protein hp14.5. Proteins 2004;54:797–800.
15. Kawarabayasi Y, Hino Y, Horikawa H, Jin-no K, Takahashi M, Sekine M, Baba S, Ankai A, Kosugi H, Hosoyama A, Fukui S, Nagai Y, Nishijima K, Otsuka R, Nakazawa H, Takamiya M, Kato

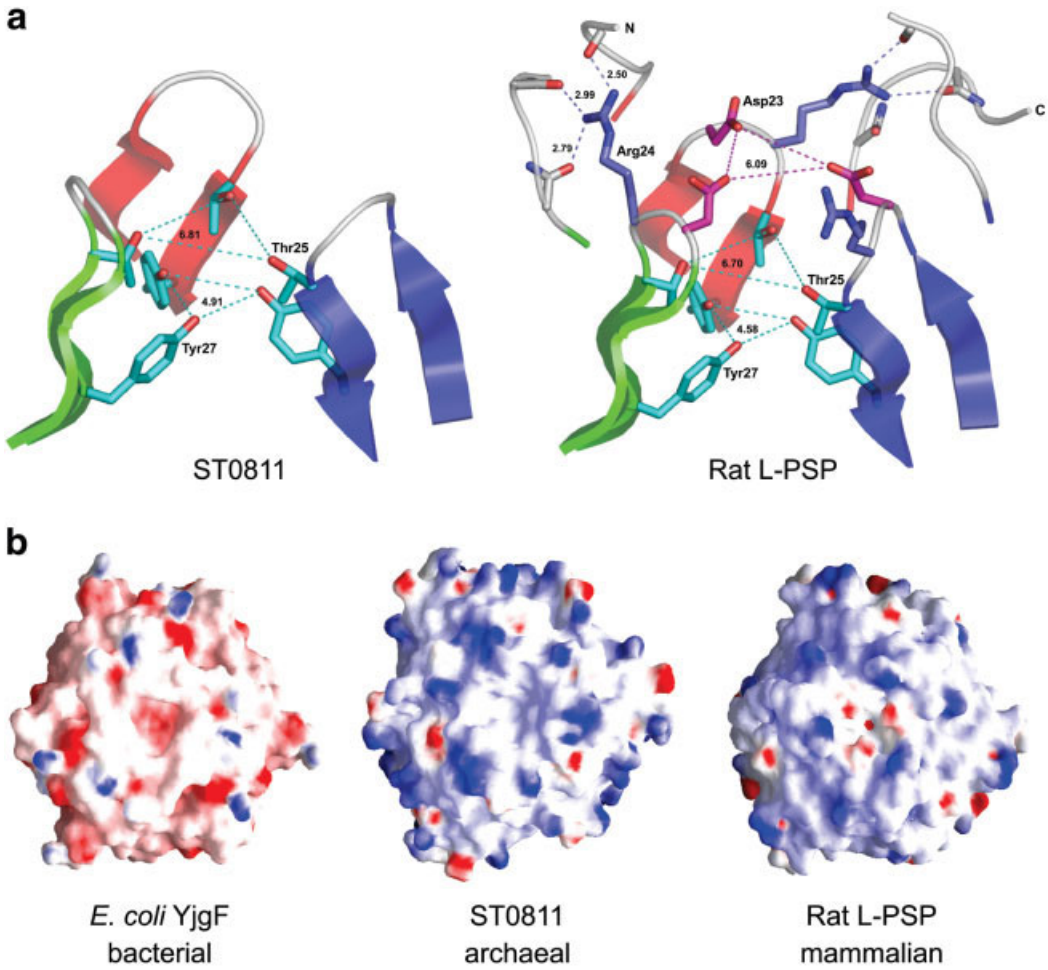

Fig. 3. (a) Ribbon diagram showing the cavity in ST0811 and rat L-PSP. The side chains forming the cavity are shown by stick model and are colored as in Figure 2(a). ST0811 does not have the regions corresponding to the N- and C-terminal extended-loop as observed in mammalian orthologs such as rat L-PSP. (b) Electrostatic potential diagram of the cavity side of ST0811, *E. coli* YjgF, and rat L-PSP. Positive and negative potentials are represented by blue and red with thresholds of ±10 kT/e, respectively.

Y, Yoshizawa T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Aoki K, Masuda S, Yanagii M, Nishimura M, Yamagishi A, Oshima T, Kikuchi H. Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7. DNA Res 2001;8:123–140.
16. Miyakawa T, Hatano K, Lee WC, Kato Y, Sawano Y, Yumoto F, Nagata K, Tanokura M. Crystallization and preliminary X-ray analysis of the YjgF/YER057c/UK114-family protein ST0811 from *Sulfolobus tokodaii* strain 7. Acta Crystallogr F 2005;61:828–830.
17. Leslie AGW. Jnt CCP4/ESF-EAMCB Newsl Protein Crystallogr 1992;26:27–33.
18. CCP4. The CCP4 Suite: programs for protein crystallography. Acta Crystallogr D 1994;50:760–763.
19. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D 1994;50:760–763.
20. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D 1997;53:240–255.
21. McRee DE. XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. J Struct Biol 1999;125:156–165.
22. Lamzin VS, Wilson KS. Automated refinement for protein crystallography. Methods Enzymol 1997;277:269–305.
23. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemistry quality of protein structures. J Appl Crystallog 1993;26:283–291.
24. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 1997;18:2714–2723.
25. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999; 14:305–308.
26. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
27. Matthews BW. Solvent content of protein crystals. J Mol Biol 1968;33:491–497.
28. Ramachandran GN, Sasisekharan V. Conformation of polypeptides and proteins. Adv Protein Chem 1968;23:283–438.
29. Ibrahim MA, Hamed RR, Rasched I. Purification and characterization of a novel acid-soluble nuclear protein from developing embryos of the camel tick *Hyalomma dromedarii* (Acarina: Ixodidae). Biochim Biophys Acta 1995;1249:79–85.
30. Schmiedeknecht G, Buchler C, Schmitz G. A bidirectional promoter connects the p14.5 gene to the gene for RNase P and RNase MRP protein subunit hPOP1. Biochem Biophys Res Commun 1997;241:59–67.